Uma maneira fácil e divertida de **entender estatística!**

Estatística

PARA LEIGOS®

FOR DUMMIES®

Tornando tudo mais fácil!

- *Aprenda Estatística de forma fácil e realista.*
- *Aprenda passo a passo e como conceitos são aplicados na vida real.*

Deborah Rumsey, PhD

Diretora do Mathematics and Statistics Laerning Center, na Ohio State University

Estatística Para Leigos®

Folha de Cola

Estatísticas Comuns

A seguir, veja as estatísticas mais comuns junto com suas fórmulas e uma breve descrição do que cada uma mede.

Estatística	Fórmula	Usada para
Média amostral (média)	$\bar{x} = \frac{\sum x}{n}$	Centro dos dados; influenciada por valores discrepantes
Mediana	valor central em um conjunto de dados ordenados	Centro numérico dos dados; não é influenciada por valores discrepantes
Desvio padrão de amostra	$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$	Medida da variabilidade, distancia típica para a média
Coeficiente de correlação	$r = \frac{1}{n-1} \sum \frac{(x-\bar{x})(y-\bar{y})}{s_x s_y}$	Mede o grau e a direção da relação linear entre x e y

Intervalos de Confiança

Um intervalo de confiança é uma suposição sobre alguma característica da população (por exemplo, qual a porcentagem de pessoas nos Estados Unidos que possuem um telefone celular?). Um intervalo de confiança contém uma estimativa inicial (digamos, o salário inicial médio baseado em uma amostra de 1.000 recém-formados) mais uma margem de erro para mais ou para menos (o valor pelo qual se espera que os resultados variem, caso uma amostra diferente seja analisada). A seguir, veja as fórmulas para os intervalos de confiança mais comuns. Leia o Capítulo 13 para mais detalhes.

Para	Estatística	Margem de Erro	Quando usar
Média populacional (μ)	$\bar{x}$	$\pm Z \times \frac{s}{\sqrt{n}}$	n é no mínimo igual a 30
Média populacional (μ)	$\bar{x}$	$\pm t_{n-1} \times \frac{s}{\sqrt{n}}$	n é menor do que 30
Proporção populacional (p)	$\hat{p}$	$\pm Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$n \times \hat{p}$ e $n(1 - \hat{p})$ é no mínimo igual a 5
Diferença de duas médias populacionais ($\mu_x - \mu_y$)	$(\bar{x} - \bar{y})$	$\pm Z \times \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	n_1 e n_2 são iguais a no mínimo 30
Diferença de duas proporções populacionais ($p_1 - p_2$)	$(\hat{p}_1 - \hat{p}_2)$	$\pm Z \times \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$1 \times \hat{p}$ e $n(1 - \hat{p})$ devem ser igual a no mínimo 5 para cada grupo.

Coeficiente de Confiança (Valores Z)

Os coeficientes de Confiança são um importante componente dos intervalos de confiança. O valor Z é parte da margem de erro — o valor que deve ser somado ou subtraído, de modo a obter-se um certo nível de confiança para os resultados. Para se obter maior confiança em seus resultados, é necessário um valor Z maior. Veja o Capítulo 9 para mais detalhes.

Nível de confiança	Valor Z
80%	1.28
85%	1.44
90%	1.64

Nível de Confiança	Valor Z
95%	1.96
98%	2.33
99%	2.58

Estatística Para Leigos®

Folha de Cola

Testes de hipótese

Os testes de hipótese são utilizados para verificar se alguma afirmação sobre uma população é verdadeira (por exemplo, alguém pode afirmar que 40% dos americanos possuem um telefone celular. Mas será realmente verdade?). Para verificar uma hipótese é necessário obter uma amostragem, coletar dados, formar uma estatística, padronizá-la, de modo a formar uma estatística de teste (para que possa ser interpretada dentro de uma escala padrão) e, por fim, decidir se a estatística de teste valida a afirmação inicial (veja os Capítulos 14 e 15 para mais detalhes). A seguir, você encontrará fórmulas para os testes de hipótese mais comuns

Testes para	*Hipótese nula (Ho)*	*Estatística de teste*	*Distribuição*	*Quando usar*
Média populacional (μ)	$\mu = \mu_0$	$\dfrac{(\bar{x} - \mu_0)}{s/\sqrt{n}}$	Normal padrão (Z)	n é igual a 30 no mínimo
Média populacional (μ)	$\mu = \mu_0$	$\dfrac{(\bar{x} - \mu_0)}{s/\sqrt{n}}$	t_{n-1}	n é menor do que 30
Proporção populacional (p)	$p = p_0$	$\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	Normal Padrão (Z)	$n \times p_0$ e $n(1-p_0)$ é igual a 5 no mínimo
Diferença de duas médias populacionais ($\mu_x - \mu_y$)	$\mu_x - \mu_y = 0$	$\dfrac{(\bar{x} - \bar{y}) - 0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$	Normal Padrão (Z)	n_1 e n_2 igual no mínimo 30
Média de diferença (antes e depois)	$\mu_d = 0$	$\dfrac{\bar{d} - 0}{s/\sqrt{n}}$	Normal Padrão (Z)	30 ou mais pares de dados
Média de diferença (antes e depois)	$\mu_d = 0$	$\dfrac{\bar{d} - 0}{s/\sqrt{n}}$	t_{n-1}	Menor do que 30 pares de dados
Diferença de duas proporções populacionais ($p_1 - p_2$)	$p_1 - p_2 = 0$	$\dfrac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$	Normal Padrão (Z)	$n \times \hat{p}$ e $n(1-\hat{p})$ é igual a 5 para cada grupo no mínimo.

Estatística

PARA

LEIGOS™

Estatística PARA LEIGOS™

Por Deborah Rumsey, Ph.D.

ALTA BOOKS
EDITORA
Rio de Janeiro 2010

ISBN 978-85-7608-387-0

Produção Editorial:
Starlin Alta Con. Com. Ltda

Coordenação Editorial:
Marcelo Utrine

Coordenação Administrativa:
Anderson Câmara

Tradução:
Larissa Franzin

Revisão Gramatical:
Elton Nunes

Revisão Técnica:
Beethoven Leite

Diagramação:
Mauro da Silva

Colaborou na Revisão Final dos Capítulos 01 ao 05:
Weuler Gonçalves

Fechamento:
Andréa Alves

Este livro não contém CD-ROM, disquete ou qualquer outra mídia.

NOTA DESTA EDIÇÃO: OS PREÇOS APRESENTADOS NESTE LIVRO ENCONTRAM-SE EM DÓLAR, PORTANTO, SUJEITOS A ALTERAÇÃO.

Erratas e atualizações: Sempre nos esforçamos para entregar ao leitor, um livro livre de erros técnicos ou de conteúdo; porém, nem sempre isso é conseguido, seja por motivo de alteração de software, interpretação ou mesmo quando alguns deslizes que constam na versão original de alguns livros que traduzimos. Sendo assim, criamos em nosso site, www.altabooks.com.br, a seção *Erratas*, onde relataremos, com a devida correção, qualquer erro encontrado em nossos livros.

Impresso no Brasil

ALTA BOOKS
EDITORA

Rua Viúva Cláudio, 291 - Bairro Industrial do Jacaré
CEP: 20970-031 - Rio de Janeiro – Tel: 21 3278-8069/8419 Fax: 21 3277-1253
www.altabooks.com.br – e-mail: altabooks@altabooks.com.br

Sobre o Autor

Deborah Rumsey obteve seu Ph.D em estatística na Universidade do Estado de Ohio (OSU) em 1993. Depois de sua graduação, ela ingressou no Departamento de Estatística da Universidade do Estado de Kansas, onde, em 1998, ganhou o distinto prêmio Presidential Teaching Award, o que garantiu o direito de posse ao cargo e uma promoção. Em 2000, ela retornou à OSU como Diretora do Centro de Aprendizagem de Matemática e Estatística, onde continua até hoje. Deb é a editora do "Teaching Bits" do Journal of Statistics Education; ela também publica artigos e realiza apresentações profissionais sobre ensino da estatística, com especial ênfase em literacia estatística (habilidade de compreender a estatística dentro do cotidiano pessoal e profissional) e em ambientes imersivos de aprendizagem (ambientes que permitem aos estudantes descobrirem e aprenderem por si mesmos). Entre suas paixões, incluem-se a pesca, a observação de pássaros e o time de futebol americano Buckeye da Universidade do Estado de Ohio (não necessariamente nessa ordem).

Dedicatória

A meu marido, Eric, e meu filho, Clint Eric. Vocês são meus maiores professores.

Agradecimentos do Autor

Eu gostaria de agradecer às pessoas que tornaram esse livro uma realidade: Kathy Cox, pelo convite de escrever o livro com o qual sempre sonhei; Tere Drenth, por dar seu toque especial ao meu sonho, garantindo que ele ficasse pronto a tempo e no formato correto; Janet Dunn, pela sua minuciosa revisão; John Gabrosek, Universidade do Estado de Grand Valley, por sua detalhada revisão técnica. Também sou muito grata ao Departamento de Serviços de Composição da Wiley Publishing, Inc., que criou todas as equações deste livro e lidou brilhantemente com as questões de layout mais complexas.

A Peg Steigerwald, Mike O'Leary e meu colega Jim Higgins pelas nossas muitas conversas sobre estatísticas anotadas em guardanapos, que me ajudaram a moldar a maneira como eu penso e falo sobre esse assunto. Obrigada por tornar esse sonho real e pelas palavras de incentivo. Obrigada a Tony Barkauskas, UW- LaCrosse, o primeiro e melhor professor que já tive, por inspirar-me. Sou muito grata a meus amigos do Departamento de Estatística e Matemática e do Centro de Aprendizagem de Estatística na Ohio State que me deram apoio e me encorajaram de maneira constante. E, por fim, muito obrigada a minha família, por seu amor e por acreditarem em mim.

Sumário Resumido

Sumário

Introdução

O objetivo deste livro é ajudá-lo apto a entender e avaliar a incrível quantidade de informação estatística a qual somos expostos diariamente. (você vai aprender sobre gráficos, tabelas, manchetes que falam sobre os resultados das últimas pesquisas de opinião, experimentos e outros estudos científicos). Este livro lhe dará os recursos necessários para decifrar e tomar importantes decisões com relação aos resultados estatísticos (por exemplo, as recentes descobertas médicas), além de tornar-lhe consciente de como você pode ser iludido pela estatística e como lidar com isso.

Este livro está recheado de exemplos reais, retirados de fontes reais, relevantes para seu dia-a-dia: desde as últimas descobertas da medicina, estudos criminalísticos e tendências da população até as pesquisas de opinião sobre namoros na Internet, o uso de telefones celulares e os piores carros do milênio. Ao ler os capítulos deste livro, você começará a entender como utilizar gráficos, tabelas, e, também aprenderá como examinar os resultados das últimas pesquisas de opinião, enquetes, experimentos e outros estudos. Você irá até mesmo descobrir como usar grilos para medir a temperatura e como ter mais chances de acertar na loteria.

Você também irá se divertir tirando sarro dos estatísticos (que, às vezes, levam-se muito a sério). E tudo porque você não precisa ser um estatístico para entender de estatística.

Sobre Este Livro

Este livro diferencia-se dos tradicionais livros, materiais de referência e manuais de estatísticas, pois possui:

- Explicações intuitivas e práticas sobre conceitos estatísticos, ideias, técnicas, fórmulas e cálculos.
- Passo a passo conciso e claro de procedimentos que intuitivamente explicam como lidar com problemas estatísticos.
- Exemplos interessantes do mundo real relacionados ao cotidiano pessoal e profissional.
- Respostas honestas e sinceras para perguntas como “O que isso realmente significa?” e “Quando e como eu vou usar isso?”

Convenções Usadas neste Livro

Há duas convenções que você deve estar atento enquanto estiver lendo este livro.

- **Definição do tamanho amostral (*n*):** quando me refiro ao tamanho de uma amostra, geralmente quero falar do número de indivíduos selecionados para participar de uma pesquisa de opinião, estudo ou experimento,(o símbolo para tamanho amostral é n). Suponha, no entanto, que 100 pessoas tenham sido selecionadas para participar de uma pesquisa de opinião e apenas 80 responderam o questionário: quais desses dois números devem ser considerados como o valor n: 100 ou 80? Eu considero o 80, ou seja, o número de pessoas que realmente responderam o questionário da pesquisa, e este número pode ser menor do que o número de pessoas selecionadas para participar. Portanto, todas as vezes em que você ler "tamanho amostral (tamanho da amostra)", entenda como o número final de indivíduos que participaram e forneceram informação para o estudo em questão.

- **A ambiguidade do termo "estatística":** em algumas situações, eu me refiro a estatística como o assunto de um estudo, ou como um campo de pesquisa. Por exemplo, "A estatística é um assunto realmente interessante". Em outras situações, eu me refiro à estatística como um número ou valor. Por exemplo, "As estatísticas mais comuns são a média e o desvio padrão".

- **Uso da palavra "dados":** Talvez você desconheça o acirrado debate entre estatísticos sobre se esta palavra deva estar no singular (dados é...) ou no plural (dados são...). O debate foi tão forte que um grupo de estatístico criou dois tipos de blusa, uma com dizeres no plural e outra com dizeres no singular. Correndo os risco de ofender alguns dos meus colegas, optei pelo plural neste livro, isto porque a palavra "dados" é, sempre será, um substantivo (pelo menos, é o que o editor me disse).

Suposições sem Importância

Não considerei que você tenha tido qualquer outro tipo de contato com a estatística a não ser pelo fato de que, como todos nós, você tem sido bombardeado todos os dias com estatística em forma de números, porcentagens, tabelas, gráficos, resultados "estatisticamente significativos", estudos "científicos", pesquisas de opinião, enquetes, experimentos e outros.

O que eu realmente levei em consideração é que você conhece as operações matemáticas básicas e entende algumas das noções básicas utilizadas em álgebra, tais como as variáveis x e y, os sinais de somatória, a extração da raiz quadrada, potenciação e outras. Caso você necessite afiar seus conhecimentos em álgebra, confira o livro Álgebra Para Leigos, de Mary Jane Sterling (Editora Alta Books).

Tenha em mente, entretanto, que estatística é bem diferente de matemática. Estatística é, antes de qualquer coisa, um método científico que determina questões de pesquisa; projeta estudos e experimentos; coleta, organiza, resume e analisa dados; interpreta resultados e esboça conclusões. Ou seja, você começará a utilizar dados como evidências para responder a interessantes questões sobre o mundo. A matemática só é utilizada para calcular a estatística sumária e realizar algumas das análises, mais isso é apenas uma pequena parte do que realmente é a estatística.

Não vou iludi-lo: você irá encontrar fórmulas neste livro, pois a estatística realmente exige alguns cálculos numéricos. Mas não deixe isso lhe abater. Vou mostrar pacientemente cada passo de qualquer cálculo que seja necessário. Também forneço no livro exemplos para que você se familiarize com os cálculos e possa fazê-los sozinho.

Como Este Livro está Organizado

Este livro divide-se em sete partes principais que exploram os objetivos principais deste livro, juntamente com a parte final, que oferece uma rápida referência para você usar. Cada parte contém capítulos que se subdividem, de modo a tornar cada objetivo mais compreensível.

Parte I: O Que Você Precisa Saber Sobre Estatística

Esta parte lhe ajudará a conscientizar-se da quantidade e da qualidade da estatística que você encontra todos os dias, em seu trabalho e em sua vida pessoal. Você também irá descobrir que grande parte dessas informações está incorreta, tanto por acidente quanto propositalmente. Você também dará o primeiro passo para se tornar estatisticamente consciente, ao reconhecer alguns dos segredos do ofício, desenvolver uma visão geral da estatística como um processo para conseguir e interpretar informações e familiarizar-se com alguns jargões técnicos.

Parte II: Fundamentos de Cálculos numéricos

Esta parte lhe ajudará a tornar-se mais familiar e confortável com os mostradores de dados (ou seja, gráficos, tabelas e outros). Esta parte ainda dará dicas de como interpretar tais tabelas e gráficos, além de identificar de imediato os gráficos mentirosos. Você também descobrirá como resumir dados utilizando algumas das funções estatísticas mais comuns.

Parte III: Determinando Probabilidades

Esta parte revelará os fundamentos da probabilidade: como usá-la; o que você precisa saber; e o que está contra você em um jogo de azar. Um conselho? Probabilidade e intuição não se misturam!

Você irá descobrir como a probabilidade influencia sua vida diária e aprenderá algumas regras básicas de probabilidade. Também saberá como realmente funcionam os jogos de azar: como os cassinos trabalham e por que a casa sempre ganha no final.

Parte IV: Decifrando Resultados

Nesta parte, você conhecerá os pilares da estatística, incluindo a distribuição amostral, a precisão, as margens de erro, o percentil e os escores padrão. Você entenderá como calcular duas medidas relativas: escores padrão e percentil. Também compreenderá o que os estatísticos descrevem como "a joia da coroa de toda a estatística" (Teorema do Limite Central) e como ela facilita a interpretação da estatística. Por fim, você começará a entender como os estatísticos medem a variabilidade de uma amostra para outra e o porquê de isso ser tão importante. Nesta parte, você também descobrirá o que é margem de erro, termo comumente utilizado.

Parte V: Estimando com Certeza

Esta parte focará em como fazer uma boa estimativa para uma média ou proporção populacional quando você não souber o número real (por exemplo, o número médio de horas que um adulto assiste televisão por semana ou a porcentagem de pessoas nos Estados Unidos que possuem pelo menos um adesivo de parachoque em seus carros. Você também irá descobrir como fazer uma ótima estimativa com uma amostra relativamente pequena (comparada com o tamanho populacional). Além disso, você aprenderá os conceitos gerais sobre os intervalos de confiança, saberá para que são utilizados, entenderá como são formados e ficará por dentro dos elementos básicos de um intervalo de confiança (uma estimativa acima ou abaixo da margem de erro). Você também irá explorar os fatores que influenciam o tamanho de um intervalo de confiança (tais como o tamanho amostral) e descobrirá fórmulas, cálculos passo a passo e exemplos dos intervalos de confiança mais comuns.

Parte VI: Testando uma Hipótese

Esta parte falará sobre o enorme papel desempenhado pela estatística no processo de tomada de decisões. Você irá saber como os pesquisadores deveriam proceder com relação à formação e ao teste de suas hipóteses e como você pode avaliar esses resultados a fim de garantir que a estatística tenha sido conduzida de maneira correta, providenciando conclusões confiáveis. Você também irá revisar instruções passo a passo para realizar cálculos comumente utilizados em testes de hipóteses e na interpretação apropriada dos resultados.

Parte VII: Estudos Estatísticos: Fique por dentro

Esta parte fornecerá uma visão geral a respeito de pesquisas de opinião, experimentos, estudos observacionais e processos de controle de qualidade. Você descobrirá o que esses estudos fazem, como são conduzidos, quais são suas limitações e como avaliá-los, para, assim, você poder determinar se seus resultados são dignos de confiança.

Parte VIII: A Parte dos 10

Esta parte, muito breve e prática, irá compartilhar dez critérios para a realização de uma boa pesquisa de opinião e dez maneiras muito comuns de uso indevido da estatística por parte de pesquisadores, da mídia e do público.

Apêndice

Um dos principais objetivos deste livro é motivá-lo e encorajá-lo a tornar-se um detetive estatístico, indo a fundo para encontrar informações reais necessárias para que você possa tomar decisões bem fundamentadas sobre as estatísticas com as quais se deparar. O apêndice contém todas as fontes que utilizei nos exemplos citados por todo o livro, caso você queira seguir a pista de algum deles.

Ícones Usados neste livro

Alguns ícones serão usados neste livro para chamar sua atenção para certos aspectos que ocorrem com frequência. A seguir, veja o que cada um deles significa:

Esse ícone refere-se a dicas, ideias ou atalhos úteis que você pode utilizar para economizar tempo. Ele também destaca modos alternativos de entender um determinado conceito.

Esse ícone contém determinadas ideias que você deve recordar mesmo depois de ter lido este livro.

Esse ícone refere-se às maneiras específicas de como os pesquisadores e a mídia podem lhe enganar com o uso da estatística e dar-lhe dicas do que você pode fazer a respeito.

Você pode apostar nesse ícone caso tenho especial interesse em entender os aspectos mais técnicos das questões estatísticas. Entretanto, você pode deixar de lê-lo caso não queira entrar em maiores detalhes.

Por Onde Começar

Este livro foi escrito de tal maneira que se pode começar a leitura de qualquer ponto e ainda continuar a entender o que se passa. Portanto, dê uma olhada na tabela de conteúdos ou no índice remissivo, procure pela informação interessante a você e vá para a página indicada.

Ou, se você não sabe ao certo por onde quer começar, considere começar sua leitura pelo Capítulo 1 e siga em frente.

Parte I

O Que Você Precisa Saber Sobre Estatística

A 5ª Onda por Rich Tennant

Nesta Parte...

Quando você liga a TV ou abre o jornal, recebe um bombardeio de números, tabelas, gráficos e resultados estatísticos. Da pesquisa de opinião de hoje até a mais recente descoberta científica, os números não param de chegar. Ainda que, na verdade, muitas dessas informações estatísticas que você acaba consumindo estejam erradas por acidente – ou, até mesmo, de propósito. Mas, como descobrir no que acreditar? Sendo um bom detetive.

Esta parte ajudará a despertar o espírito estatístico investigativo que adormece em você, ao analisar como a estatística afeta seu dia-a-dia e seu trabalho, a real qualidade da informação que nos é apresentada e o que você pode fazer com relação a isso. Além disso, esta parte lhe ajudará a entender alguns termos muito úteis do jargão estatístico.

Capítulo 1

A Estatística da Vida Diária

Neste Capítulo

- Estatística no seu dia-a-dia: como e com que frequência você a vê
- Descubra como a estatística é utilizada no local de trabalho

Atualmente, a sociedade está completamente tomada pelos números. Eles aparecem em todos os lugares para onde você olha, de outdoors mostrando as últimas estatísticas sobre aborto, passando pelos programas de esporte que discutem as chances de um time de futebol chegar à final do campeonato, até o noticiário da noite, com reportagens focadas no índice de criminalidade, na expectativa de vida de uma pessoa que não come alimentos saudáveis e no índice de aprovação do presidente. Em um dia comum, você pode se deparar com cinco, dez ou, até mesmo, vinte diferentes estatísticas (ou até muito mais em um dia de eleição). Se você ler todo o jornal de domingo, irá se deparar com centenas de estatísticas em reportagens, propagandas e artigos sobre todo tipo de assunto: desde sopa (quanto em média uma pessoa consome por ano?) até castanhas (quantas castanhas você precisa comer para aumentar seu QI?).

O objetivo deste capítulo é mostrar a frequência com que a estatística aparece em sua vida pessoal e profissional e como ela é apresentada para o público em geral. Depois de ler este capítulo, você começará a perceber a frequência com que a mídia nos bombardeia com números e o quanto é importante estar apto a decifrá-los. Portanto, goste você ou não, a estatística faz parte de sua vida. Então, se você não pode com ela e não quer se juntar a ela, pelo menos, tente entendê-la.

Estatística e o Bombardeio da Mídia: Mais Perguntas do que Respostas?

Abra o jornal e comece a procurar por exemplos de artigos e notícias que envolvam números. Não vai levar muito tempo até que os números comecem a se tumultuar. Os leitores são inundados com resultados de estudos, anúncios de avanços tecnológicos, relatórios estatísticos, previsões, projeções, tabelas, gráficos e resumos. A maneira como a estatística aparece na mídia é espantosa. Você pode nem se dar conta de quantas vezes foi bombardeado com números na atual era da informação. Aqui vão apenas alguns exemplos de um jornal de domingo. Enquanto lê, você poderá se sentir nervoso, perguntando-se no que pode ou não confiar. Relaxe! É por isso que este livro está aqui: para ajudá-lo a separar o joio do trigo (Capítulos de 2 a 5 lhe darão um ótimo começo).

Briga no campo

O primeiro artigo com o qual me deparei e que lidava com números apresentava a seguinte manchete: "Lavoura de milho enfrenta saúde pública" e o subtítulo era: "Trabalhadores doentes dizem que palatabilizantes químicos causaram problemas pulmonares". O artigo descrevia como o Centro de Controle de Doenças mostrava-se preocupado com a possível relação entre a exposição aos palatabilizantes químicos utilizados no milho para pipoca de microondas e alguns casos de doença pulmonar obstrutiva. Oito trabalhadores de uma única lavoura de milho apresentaram a doença e quatro deles estavam aguardando por um transplante. De acordo com o artigo, casos semelhantes também haviam sido relatados em outras indústrias de milho. Agora, você deve estar se perguntando: "E as pessoas que comeram a pipoca de microondas?". Segundo o artigo, o Centro de Controle de Doenças disse "que as pessoas que comeram pipoca de microondas não têm motivos para se preocupar" (Fique atento). Eles ainda disseram que seu próximo passo seria fazer uma avaliação mais detalhada desses funcionários, incluindo pesquisas para determinar as condições de saúde e a possível exposição aos produtos químicos mencionados, para verificação da capacidade pulmonar e para coleta de amostras do ar. A questão aqui é: Quantos casos dessa doença pulmonar realmente constituem um padrão, comparado à mera casualidade ou a uma anomalia estatística? (Mais sobre o assunto no Capítulo 14).

Virose na Rede

O segundo artigo que encontrei discutia o mais recente ataque na rede – um vírus tipo traça que tem passeado pela Internet, deixando mais lenta a navegação pela rede e a troca de e-mails no mundo todo. Quantos computadores foram infectados? Os estudiosos mencionados no artigo disseram que 39.000 computadores haviam sido infectados, afetando centenas de milhares de outros sistemas. Mas, como eles chegaram a esse número? Esse não seria um número muito difícil de ser calculado? Será que eles verificaram todos esses computares para ver se eles realmente estavam infectados? O fato de esse artigo ter sido escrito em menos de 24 horas depois da ocorrência do ataque sugere que esse número é apenas uma suposição. Então, por que dizer 39.000 e não 40.000? Para saber mais sobre como estimar com mais certeza (e como avaliar a estimativa feita por outros), veja o Capítulo 11.

Acidentes de percurso

Logo em seguida, no jornal, apareceu um alerta sobre o elevado número de acidentes com motos. Alguns estudiosos disseram que esses acidentes têm crescido mais do que 50% desde 1997, e ninguém consegue explicar o motivo. As estatísticas mostram um fato curioso. Em 1997, morreram 2.116 motociclistas nos Estados Unidos; em 2001, o número aumentou para 3.181, segundo os registros do National Highway Traffic Safety Administration (NHTSA) – órgão gestor do tráfico nos Estados Unidos. Nesse artigo, discutiram-se muitas possíveis causas para o aumento das taxas de mortalidade em acidentes envolvendo motociclistas, incluindo o fato de que atualmente os motociclistas tendem a ser mais velhos (a idade média dos motociclistas mortos em acidentes saltou de 29,3 anos em 1990 para 36,3 anos em 2001).

O tamanho das motos, cada vez maiores, também foi apontado como outra possibilidade. O tamanho do motor de uma motocicleta média aumentou 25% - de 769 centímetros cúbicos em 1990 para 959 centímetros cúbicos em 2001. Outra possibilidade pode ter sido o fato de alguns estados estarem afrouxando suas leis com relação ao uso do capacete. Os estudiosos entrevistados disseram, no artigo, que é necessário um estudo de causa mais abrangente, mas um estudo como esse provavelmente não seria realizado, pois custaria de 2 a 3 milhões de dólares. Uma questão não mencionada no artigo foi o número de motociclistas em 2001 comparado com o número em 1990. Um número maior de pessoas nas estradas geralmente ocasiona um número maior de acidentes, mesmo quando todos os outros fatores permanecem os mesmos. Entretanto, juntamente com o artigo, havia um gráfico mostrando as mortes causadas em acidentes com motos a cada 100 milhões de milhas rodadas nos Estados Unidos de 1997 até 2001; mas ele levava em consideração o aumento do número de pessoas nas estradas? Também foi incluído um gráfico de barras comparando as mortes causadas por acidentes com motos versus as mortes causadas por acidentes com outros tipos de veículos. Esse gráfico mostrava que as mortes com moto ocorreram a uma taxa de 34,4 mortes a cada 100 milhões de milhas rodadas, enquanto as mortes por acidentes envolvendo carros ocorreram a uma taxa de apenas 1,7 mortes para a mesma quantidade de milhas rodadas. Esse artigo apresentava muitos números e estatísticas, mas o que isso tudo realmente significa? A quantidade e o tipo das estatísticas podem se confundir rapidamente. O Capítulo 4 o ajudará a decifrar gráficos, tabelas e as estatísticas que os acompanha.

Saúde em crise

Continuando minha leitura, encontrei uma reportagem a respeito de um estudo sobre seguro contra imperícia, que poderia afetar os cidadãos americanos com relação às taxas cobradas pelos médicos e o direito de conseguir atendimento médico apropriado. Então, qual era a dimensão do problema? O artigo indicava que 1 em 5 médicos na Georgia haviam parado de realizar procedimentos de risco (como partos) graças ao contínuo aumento das taxas de seguro contra imperícia no estado. A situação foi descrita como uma "Epidemia Nacional" e uma "Crise da Saúde" por todo o país. O artigo incluiu poucos detalhes sobre o estudo, mas declarou que de 2.200 médicos entrevistados na Georgia, 2.800 – que eles disseram representar cerca de 18% da amostra – iriam parar de realizar procedimentos de alto risco. Mas espera um pouco! Isso está certo? De 2.200 médicos, 2.800 não realizarão procedimentos e esse número representa apenas 18%? Isso é impossível! Não dá para ter o numerador de uma fração maior do que seu denominador e ainda obter um resultado abaixo de 100%, certo? Este é um dos muitos exemplos de erros em estatística que são anunciados pela mídia. Então, qual é a porcentagem real? Você só pode estimar. O Capítulo 5 definirá precisamente as particularidades para se calcular as estatísticas, para que assim você possa saber o que procurar e perceber imediatamente quando alguma coisa estiver errada.

Invasão de terras

No mesmo jornal de domingo, havia um artigo sobre a dimensão do desenvolvimento urbano e da especulação imobiliária por todo o país. Dado o número de casas a ser construído em nosso pedaço de chão, essa é uma importante questão a ser debatida. As estatísticas mostravam o número de hectares de áreas rurais que estão sendo transformadas em áreas urbanas a cada ano e também transformava essa quantidade em acres em milhas quadradas. Mais a seguir, a fim de ilustrar a quantidade de terra que está se perdendo, o tamanho da área foi comparado a quantidades de campos de futebol americano. Nesse exemplo, em particular, estudiosos disseram que a região central de Ohio está perdendo 60.703,095 hectares por ano, ou 234 milhas quadradas, o equivalente a 115.385 campos de futebol americano (incluindo as zonas finais). Como se chegou a esses números? Será que eles são precisos? Comparar a campos de futebol o tamanho da área de terra perdida realmente ajuda na visualização dessa área?

Avaliação do desempenho escolar

O tópico seguinte do artigo era sobre a avaliação do desempenho escolar, em especial se as atividades extracurriculares estavam realmente ajudando na melhora do desempenho dos alunos. O artigo declarava que 81,3% dos alunos que frequentaram as atividades extras passaram nos exames de proficiência em redação, enquanto somente 71,7% dos que não participaram das atividades passaram no mesmo teste. No entanto, essa é, realmente, uma diferença que justifique o gasto de 386.000 dólares por ano? E o que está sendo feito durante essas atividades extras para melhorar o desempenho escolar? Será que os alunos estão apenas se preparando para o teste de proficiência ou realmente estão aprendendo mais sobre redação em geral? E aqui vai a grande pergunta: Esses que participaram das atividades extras foram mais motivados a melhorar suas notas do que os outros alunos? Ninguém sabe. Estudos como esses sempre aparecem e a única maneira de saber no que se pode confiar é entender quais perguntas devem ser feitas e estar apto a avaliar a qualidade do estudo apresentado. Tudo isso faz parte da estatística! A boa notícia é que, com algumas perguntas esclarecedoras, você poderá rapidamente avaliar um estudo estatístico e seus resultados. O Capítulo 17 o ajudará a fazer isso.

DICA

Tentando a sorte

Você já sonhou em ganhar na loteria, onde as chances são de uma em 89 milhões?

É isso mesmo! Para que você possa visualizar melhor esse número, imagine 89 milhões de bilhetes de loteria empilhados, sendo que um desses é o seu. Suponha que eu dissesse que você tem uma chance para tirar o seu bilhete da pilha — você acha que conseguiria? Essa é a mesma chance que você tem de ganhar na loteria. Mas, com algumas informações valiosas, você poderá aumentar sua chance de acerto (e eu vou querer uma parte de seu prêmio caso funcione para você). Para mais informações a esse respeito e sobre outras dicas para jogos de azar, veja o Capítulo 7.

LEMBRE-SE

Estudar levantamentos de todos os tamanhos e feitios

Enquetes e pesquisas de opinião são, provavelmente, os maiores veículos utilizados pela mídia hoje para chamar sua atenção. Parece que todos querem fazer uma enquete, incluindo gerentes de mercados, companhias de seguro, canais de TV, comunidades e, até mesmo, estudantes de segundo grau. Aqui estão alguns exemplos de resultados de pesquisas de jornal que fazem parte do nosso cotidiano.

Com o envelhecimento da mão-de-obra americana, as empresas estão se preparando para suas futuras lideranças (mas como eles sabem que a mão-de-obra está envelhecendo, e, se realmente está, o quanto está envelhecendo?). Uma pesquisa recente mostrou que quase 67% dos gerentes de recursos humanos entrevistados disseram que o planejamento para a sucessão havia se tornado mais importante nos últimos cinco anos do que havia sido no passado. Agora, se você estiver pensando em deixar seu emprego para se candidatar a um cargo de diretoria, espere um pouco. A pesquisa também mostrou que 88% dos 210 entrevistados disseram que, geralmente, preenchem os cargos seniores com candidatos internos (mas quantos gerentes não responderam? E 210 entrevistados realmente representam gente o suficiente para render uma notícia de primeira página no caderno de negócios?). Acredite ou não, quando você começa a procurar, encontra nos jornais numerosos exemplos de pesquisas feitas com muito menos participantes do que essa.

Algumas pesquisas baseiam-se em assuntos menos importantes. Por exemplo, qual objeto os americanos acreditam ser mais importante nos dias de hoje, a escova de dente, as panificadoras caseiras, computadores, carros ou telefones celulares? Em uma pesquisa feita com 1.042 adultos e 400 adolescentes (qual foi o critério utilizado para a adoção desses números?), 42% dos adultos e 34% dos adolescentes disseram que a escova de dente era mais importante do que carros, computadores ou celulares. Mas isso é mesmo uma grande notícia? Desde quando algo tão importante para a higiene diária de alguém deveria ser equiparada com telefones celulares e panificadoras caseiras? (o carro ficou em segundo lugar. Mas, é necessário se fazer uma pesquisa para isso?). Para mais informações sobre pesquisas de opinião, veja o Capítulo 16.

Estudando os esportes

O caderno de esportes é, provavelmente, o caderno mais cheio de números de todo o jornal. Além dos pontos do último jogo, as porcentagens de vitória/derrota de cada time e a posição relativa de cada time, a estatística especializada no mundo dos esportes é tão pesada que é necessário ter um bom preparo físico para encará-la. Por exemplo, as estatísticas para o basquete são divididas por equipe, por tempo e, até mesmo, por jogador. E você também precisa ser um expert em basquete para interpretá-las, pois tudo é abreviado (e sem legenda para ajudar os leigos):

- MIN: Minutos Jogados
- FG: Cesta (sigla do termo em inglês Field Goals)
- FT: Lances livres (sigla do termo em inglês Free throws)

- RB: Rebotes
- A: Assistência
- PF: Falta pessoal (sigla do termo em inglês Personal fouls)
- TO: Inversão da posse de bola (sigla do termo em inglês Turnover)
- B: Bloqueios
- S: Roubo de bola (sigla do termo em inglês Steals)
- TP: Total de pontos

Quem precisa saber disso? Talvez as mães dos jogadores? A estatística é algo que os torcedores nunca terão o bastante e algo sobre o qual os jogadores não suportam mais ouvir falar.

Bancando o detetive

No caderno de negócios do jornal, você encontra estatística sobre a bolsa de valores. Semana passada não foi uma boa semana, a bolsa caiu 455 pontos; caiu pouco ou muito? Precisamos calcular uma porcentagem para realmente entender isso. Ainda no mesmo caderno, você encontra reportagens sobre os CDs mais vendidos em todo país (a propósito, como eles sabem que são os mais vendidos?). Você também encontra reportagens sobre as taxas de juros para empréstimos: parcelas fixas em 30 anos, parcelas fixas em 15 anos, parcelas ajustáveis em 1 ano, empréstimos para carros novos, carros usados, para casa própria e para empréstimos feitos pelas avós (bom, na verdade não, mas se sua avó soubesse ler essas estatísticas, ela poderia considerar a possibilidade de aumentar essa taxa de juros insignificante que cobra de você!). Por fim, você também encontra milhares de anúncios dos tão amados cartões de crédito – anúncios mostrando as taxas de juros, a anuidade e o prazo para o pagamento de suas contas. Mas, como comparar todas as informações sobre investimentos, empréstimos e cartões de crédito a fim de tomar a melhor decisão? Quais estatísticas são as mais importantes? A verdadeira pergunta é: os números mostrados no jornal contam-lhe tudo, ou será necessária mais investigação para chegar à verdade? O Capítulo 3 o ajudará a começar a desmembrar esses números e tomar as decisões corretas a partir deles.

Visitando o caderno de turismo

Nem no caderno de turismo você consegue escapar do coral de números. Nessa seção, descobri que a pergunta mais frequente feita ao Transportation Security Administration Center (que recebe em média 2.000 telefonemas, 2.500 e-mails e 200 cartas por semana – você gostaria de ser a pessoa que fez essa conta?) foi: "Posso levar isso no avião?"; sendo que "isso" se referia a qualquer coisa, desde um animal até um balde de pipoca tamanho gigante (eu não recomendaria o balde de pipoca, pois você tem que guardá-lo horizontalmente no compartimento de bagagem acima dos assentos e, como as bagagens balançam durante o voo, o balde provavelmente se abrirá e, quando você for pegá-lo, você e seus companheiros de poltrona levarão um banho de pipoca. Sim, eu já vi isso acontecer uma vez).

Isso nos leva a uma interessante questão estatística: quantas pessoas por dia são necessárias para atender essas chamadas? O primeiro passo é estimar o número de chamadas e, se você errar, perderá muito dinheiro (caso tenha superestimado) ou ganhará muitas reclamações (caso tenha subestimado).

Conversando sobre sexo (e estatística) com a Dra. Ruth

Em uma página de destaque do jornal de domingo, pude ler sobre a última pesquisa realizada pela Dra. Ruth sobre a vida sexual das pessoas. Ela dizia que o sexo não parava aos 60 anos nem mesmo aos 70. Bom saber, mas como ela determinou isso? E com que frequência as pessoas estão tendo relações sexuais com essa idade? Ela não disse (talvez seja melhor omitir alguns números, não é mesmo?). Entretanto, a Dra. Ruth recomenda que as pessoas nessa idade desconsiderem as pesquisas que mencionam quantas vezes por semana, mês ou ano um casal tem relações sexuais. Segundo ela, essas pesquisas apenas mostram as pessoas se gabando. Ela pode estar certa com relação a isso. Pense, se alguém realiza uma pesquisa ligando para a casa das pessoas e pedindo alguns minutos da sua atenção para discutir sobre sua vida sexual, quem estaria disposto a responder a pesquisa? E o que responderiam para a pergunta: "Quantas vezes por semana você tem relações sexuais?" Será que responderiam a verdade, ou exagerariam um pouco? Pesquisas de autorrelato realmente são uma fonte real de preconceitos e podem levar a estatísticas enganosas. Portanto, pegue leve com a Dra. Ruth (que, por sinal, é a autora de Sex for Dummies, 2nd Edition, publicada pela Wiley Publishing, Inc). Como você a ajudaria a descobrir mais sobre esse assunto tão pessoal? Às vezes, as pesquisas são mais difíceis do que parecem. O Capítulo 2 tem mais exemplos de como a estatística pode errar e como você deve ficar atento a elas.

Mudança do tempo

As notícias sobre o tempo também fornecem um monte de estatística, com suas previsões de temperatura máxima e mínima para o dia seguinte (por que dizem 16°C e não 15°C?) e relatos sobre o fator UV do dia, o índice de poluição e a qualidade e quantidade de água (como eles conseguem esses números, coletando amostras? Quantas amostras são coletadas e onde elas são coletadas?) Atualmente já se pode fazer uma previsão metereológica para os próximos 3 dias, uma semana ou, até mesmo, um mês ou um ano! Mas essas previsões são realmente precisas? Dado o número de vezes que chove quando os metereologistas dizem que vai fazer sol, pode-se dizer que ainda há muito a melhorar!

Loucuras em Las Vegas

Quando vemos como os números são usados (e abusados) no dia-a-dia, não podemos ignorar o mundo das apostas, um negócio que movimenta bilhões de dólares por ano e que inclui os apostadores casuais, os profissionais e os compulsivos. Em que tipo de coisa você consegue apostar? Em qualquer coisa que possa ter dois resultados. As loucuras que alguém pode fazer em Las Vegas não têm limite (sem duplo sentido).

Aqui vão alguns exemplos de alguns tópicos quentes relacionados aos Super Bowl em que se pode apostar em um sports book (local de aposta) em Las Vegas:

- Qual time cometerá mais pênaltis?
- Qual time marcará pontos por último no primeiro tempo?
- Tentarão fazer uma conversão de 2 pontos?
- O que tentarão primeiro, marcar um ponto ou fazer um punt?
- O total de jardas para os dois times será maior ou menor do que 675?

Os dois times farão um Field Goal de 33 jardas ou mais? Hum. Por que não apostar no número de quilos de guacamole consumido pelos telespectadores do Super Bowl versus o número de lâminas de grama do campo? Façam suas apostas.

LEMBRE-SE

A probabilidade e os programas de computador realmente desempenham um importante papel para a previsão do tempo atualmente, embora sejam particularmente úteis no que se refere aos grandes eventos como furacões, terremotos e erupções vulcânicas. É claro que os computadores são apenas tão inteligentes quanto às pessoas que os programam e, portanto, os cientistas ainda têm muito a fazer antes que se possa prever um tornado, antes mesmo que ele se forme (não seria maravilhoso?). Para mais informações sobre computação e estatística, veja o Capítulo 6.

De olho nas estrelas

Indo para o caderno de artes, é possível ver vários anúncios de filmes. Cada anúncio contém uma citação de algum crítico de cinema, como: "A maior aventura de todos os tempos"; "Absolutamente hilário"; ou "Um dos dez melhores filmes do ano!". Você presta atenção nos críticos? Como você decide qual filme assistir? Especialistas dizem que enquanto a popularidade de um filme pode ser afetada pelos comentários dos críticos (bons ou ruins) na estreia de um filme, no final, no entanto, o boca a boca é mais importante para determinar o sucesso de um filme.

Estudos também mostram que quanto mais dramático for o filme, mais pipocas são vendidas. Sim, o show business também fica atento a quantas vezes você mastiga dentro do cinema. Como eles coletam todas essas informações e como elas influenciam o tipo de filme que é feito? Isso também é parte da estatística: projetar e conduzir estudos para ajudar a identificar um tipo de público, a fim de descobrir o que ele realmente gosta, e utilizar a informação para a criação de um produto. Então, da próxima vez que alguém segurando uma prancheta lhe perguntar se você tem um minuto, dê sua opinião.

Ouvindo os astros

Horóscopos: você os lê, mas você acredita no que eles dizem? Deveria acreditar? As pessoas podem prever o que irá acontecer com mais frequência do que por coincidência? Os estatísticos têm uma maneira de saber, usando o que eles chamam de teste de hipótese (veja o Capítulo 14). Até agora eles não encontraram alguém que possa ler mentes, mas as pessoas podem continuar tentando!

Usando a Estatística no Trabalho

Vamos dar um tempo com o jornal de domingo, que você lê no conforto de sua casa, e vamos para seu ambiente de trabalho. Se você trabalha para um escritório de contabilidade, é claro que os números fazem parte de sua vida diária. Mas, e as enfermeiras, fotógrafos de estúdio, gerentes de lojas, jornalistas, trabalhadores de escritórios ou, até mesmo, os trabalhadores da construção civil? Os números são parte de sua vida, caso você tenha algum desses trabalhos? Pode apostar. Esta seção lhe dará alguns exemplos de como a estatística está em todos os ambientes de trabalho.

LEMBRE-SE

Não é preciso ir muito longe para ver os rastros que a estatística deixa e os caminhos que ela traça dentro e fora de sua vida pessoal e profissional. O segredo é ser capaz de determinar o que tudo isso significa, no que você pode acreditar e ser capaz de tomar decisões sensatas, baseando-se na verdade por trás desses números para que você possa lidar e, até mesmo, acostumar,-se com a estatística da vida diária.

Entregando bebês – e informações

Sue trabalha como enfermeira durante o turno da noite na unidade de partos em um hospital universitário. Ela deve cuidar de várias pacientes em uma dada noite e fazer o máximo para acomodar a todos. Sua chefe lhe havia dito que todas as vezes em que ela assumir o turno, deveria se identificar para a paciente, escrever seu nome no quadro branco no quarto da paciente e perguntar-lhe se ela tem alguma pergunta. Por que ela faz isso? Pois, alguns dias depois que cada mãe deixa o hospital e volta para casa, ela recebe um telefonema perguntando sobre a qualidade do serviço, o que ficou faltando, o que o hospital pode fazer para melhorar a qualidade do serviço e do atendimento e o que os funcionários do hospital podem fazer para garantir que ele seja mais escolhido do que os outros hospitais da cidade. Um serviço de qualidade é importante e, para as novas mamães no hospital, com enfermeiras entrando e saindo a cada oito horas, saber o nome de suas enfermeiras é importante, pois isso faz que com que elas tenham suas perguntas respondidas prontamente. As comissões de Sue dependem de sua habilidade de satisfazer as necessidades das novas mamães.

Posando para a foto

Carol recentemente começou a trabalhar como fotógrafa para o estúdio fotográfico de uma loja de departamentos; um de seus pontos fortes é trabalhar com bebês. Baseando-se no número de fotos compradas pelos clientes ao longo dos anos, essa loja percebeu que as pessoas compram mais as fotos com poses do que as naturais. Consequentemente, os gerentes da loja irão encorajar os fotógrafos a tirarem mais fotos com poses.

Uma mãe entra na loja com seu bebê e faz um pedido especial: "gostaria que as fotos parecessem naturais". O que Carol diria: "Desculpe-me, não posso fazer isso, minha comissão depende da minha habilidade de tirar fotos com poses". Bom, pode ter certeza que essa mãe irá responder à pesquisa sobre qualidade do atendimento depois da sessão de fotos – e não apenas para retirar o cupom de desconto de $2,00 para a próxima foto (se é que ela voltará para a próxima foto).

Acabando em pizza

Terry é gerente de uma pizzaria local que vende pizza por fatias. Ele é o responsável por determinar quantos funcionários são necessários em um dado horário, quantas pizzas fazer antecipadamente para atender a demanda e quanto queijo é necessário comprar e ralar, tudo com o gasto mínimo de recursos e ingredientes. É meia-noite de uma sexta-feira e o local está vazio. Terry está com cinco funcionários e tem 5 massas de pizza para fazer 40 fatias. Será que ele deveria mandar dois de seus funcionários para casa? Deveria colocar as massas no forno ou esperar? Terry sabe o que, provavelmente, irá acontecer, pois, há semanas, o dono da pizzaria tem rastreado o movimento de seu estabelecimento e sabe que toda sexta-feira à noite o movimento diminui entre as dez e meia-noite, mas, depois desse horário, o movimento aumenta e os clientes não vão embora antes das duas e meia da manhã, horário em que a pizzaria fecha. Portanto, Terry fica com os funcionários, coloca as massas de pizza no forno começando com intervalos de 30 minutos a partir da meia-noite e acaba recompensado com uma noite lucrativa, clientes satisfeitos e um chefe feliz. Para mais informações sobre como fazer boas estimativas usando a estatística, veja o Capítulo 11.

Trabalhando no escritório de estatística

Vejamos o exemplo de DJ, uma assistente administrativa em uma empresa de informática. Como a estatística chega até seu ambiente de trabalho? Fácil. Todo escritório está repleto de pessoas que querem respostas a suas perguntas, e elas querem alguém que "destrinche os números" para "dizer-lhes o que eles realmente significam" para "descobrir se alguém tem dados mais precisos sobre um assunto" ou para simplesmente dizer, "Esses números fazem algum sentido?". Essas pessoas precisam saber tudo: desde os números de clientes satisfeitos até as alterações no inventário durante o ano; da porcentagem do tempo que os funcionários gastaram lendo seus e-mails até o valor gasto com suprimentos nos últimos três anos. Todos os ambientes de trabalho estão repletos de estatística, e o valor de mercado de DJ como assistente administrativa poderia aumentar se ela se tornasse a pessoa de confiança do chefe. Todo escritório precisa de um residente de estatística – por que não você?

Capítulo 2

As Estatísticas Erram

Neste Capítulo

- Examinando a dimensão dos abusos da estatística
- Analisando as proibições mais comuns da estatística
- Sentindo o impacto de um erro estatístico

A explosão de números pode lhe deixar desnorteado e confuso (mas, com a ajuda deste livro, você será capaz de entender muitas das estatísticas que encontrar em sua vida cotidiana!). O propósito deste capítulo é despertar outro sentimento: o ceticismo! Mas não um ceticismo como o de: "Não posso mais acreditar em nada", e sim um ceticismo como: "Hum, gostaria de saber de onde tiraram esses números". "Esses números serão verdadeiros?" "Preciso descobrir mais sobre esse estudo antes de acreditar em seus resultados". A mídia apresenta muitos exemplos de estatística errada e, depois que você descobrir como identificar tais problemas, ficará mais confiante com relação à estatística e pronto para desarmar qualquer explosão de números!

Assumindo o Controle: Tantos Números e Tão Pouco Tempo

As estatísticas acabam aparecendo na TV ou em seu jornal como o resultado de um processo. Primeiro, os pesquisadores que estudam um assunto geram um resultado, esse grupo é composto por entrevistadores, doutores, pesquisadores de marketing, pesquisadores do governo e outros cientistas. Eles são considerados as fontes originais da informação estatística. Depois de conseguirem seus resultados, esses pesquisadores querem contá-los às pessoas, então, geralmente, lançam um comunicado de imprensa ou um artigo científico. Aí entram os jornalistas, que são considerados a fonte de mídia da informação. Os jornalistas caçam os comunicados de impressa mais interessantes e vasculham as revistas científicas, basicamente em busca da próxima manchete. Quando eles terminam suas histórias, as estatísticas são enviadas ao público. Isso pode acontecer por meio de vários tipos de mídia: TV, jornais, revistas, sites na internet, newsletters e assim por diante. Agora, a informação está pronta para ser recebida pelo terceiro grupo, os consumidores de informação (você!). Você e outros consumidores de informação são as pessoas que terão que encarar a tarefa de ouvir e ler as informações, analisando-as e tomando decisões a partir delas. E, como você já deve ter adivinhado, em qualquer parte do processo de fazer a pesquisa, comunicar os resultados ou consumir a informação pode haver erros, intencionais ou não.

Detectando Erros, Exageros e Mentirinhas Leves

As estatísticas podem errar por muitas razões diferentes. Primeiramente, pode realmente haver um simples e honesto engano. Isso pode acontecer com qualquer um, não é mesmo?

Outras vezes, esse erro pode ser um pouco mais do que um simples e honesto engano. No calor do momento, por causa dos fortes sentimentos a favor de uma causa e pela falta de suporte dos números a esse ponto de vista, a estatística pode ser ajustada, ou, na maioria das vezes, exagerada, tanto em termos de valores quanto na maneira como é apresentada e discutida.

Por fim, você também pode encontrar situações em que os números haviam sido completamente fabricados e não poderiam ser repetidos por ninguém, pois os resultados nunca existiram. Esse é o pior caso e realmente acontece no mundo real.

Esta seção dá dicas para lhe ajudar a identificar erros, exageros e mentiras junto com alguns exemplos de cada tipo de erro que você, como um consumidor de informação, pode encontrar.

Conferindo a matemática

A primeira coisa a ser feita quando você se deparar com uma estatística ou com o resultado de um estudo estatístico é se perguntar: "Este número está correto?". Não o aceite como correto logo de cara! Você ficará surpreso com o número de ocorrência de erros aritméticos quando a estatística é coletada, resumida, relatada ou interpretada. Tenha em mente que outro tipo possível de erro é o *erro da omissão* – informação que não é citada e que poderia fazer uma grande diferença em termo de entendimento da verdade por trás dos números. Isso faz com que seja difícil direcionar a questão da exatidão, pois falta-lhe informação para seguir em frente.

DICA

Para identificar erros de aritmética ou de omissão nas estatísticas:

- Verifique se a soma está correta. Ou seja, as porcentagens do gráfico pizza realmente somam 100%? A soma do número de pessoas em cada categoria resulta no número de pessoas entrevistadas?
- Verifique duas vezes até mesmo os cálculos mais básicos.
- Sempre procure por um total, para que assim você possa colocar os resultados em uma perspectiva adequada. Ignore os resultados baseados em amostras de tamanho muito pequeno.

Contas que não batem

Muitas estatísticas dividem o resultado em grupos, mostrando a porcentagem de pessoas de cada grupo que foram entrevistadas com relação à determinada questão ou fator demográfico (tais como idade, sexo e outros). Essa é uma forma efetiva de relatar a estatística, desde que todas as porcentagens somem 100%.

Por exemplo, o USA Today relatou o resultado de uma pesquisa de opinião feita pela Tupperware com relação ao hábito de aquecer as sobras de alimento no forno microondas. A notícia dizia que 28% das pessoas entrevistadas disseram que aqueciam diariamente as sobras de alimentos no forno micro-ondas, 43% disseram que usavam de duas a quatro vezes por semana o forno microondas para aquecer as sobras de alimentos e 15% disseram fazer isso apenas uma vez por semana. Considerando que todos entrevistados deveriam se enquadrar em uma dessas categorias, a soma das porcentagens deveria ser de 100% ou o mais próximo possível. Se verificarmos rapidamente, a soma de 28% + 43%+ 15% = 86%. Então, o que aconteceu com os 14%? Quem foi deixado de fora? Onde foram colocados? Ninguém sabe. A estatística simplesmente não bate!

Quatro em cada cinco – verdade?

Outro item que você pode verificar rapidamente é se o número total de entrevistados foi informado. Para citar um exemplo simples, você deve se lembrar daquele comercial do Trident que dizia que “Quatro em cada cinco dentistas recomendam Trident para seus pacientes”. Esse comercial é um pouco antigo, mas, recentemente, ele foi relembrado por uma série de comerciais divertidos em que se perguntava o que teria acontecido com o quinto dentista e, depois, mostrava alguns incidentes que poderiam ter ocorrido, impedindo que ele ou ela apertasse o botão do “sim”. Mas, aqui vai a verdadeira pergunta: Quantos dentistas realmente foram entrevistados? Você não sabe, pois a pesquisa não informou. Você nem consegue verificar a boa qualidade da impressão, pois, no caso desse tipo de comercial, ela não é necessária.

Por que o fato de saber o número total de entrevistados faz diferença? Pois, a confiabilidade de uma estatística se deve, em parte, à quantidade de informação que entrou da estatística (desde que seja informação boa e correta). Quando a propaganda disse: “quatro em cada cinco dentistas”, isso realmente pode significar que apenas cinco dentistas foram entrevistados. Agora, talvez 5.000 dentistas tenham sido entrevistados e, nesse caso, 4.000 recomendaram o produto. O fato é que você não conseguirá saber quantos dentistas realmente recomendaram o produto, a menos que investigue mais para descobrir. Na maioria dos casos, sobra para você, o consumidor, o trabalho de descobrir essa informação. A não ser que você saiba o número total de pessoas que foram entrevistadas, você não conseguirá ter uma perspectiva da confiabilidade dessa informação.

Revelando estatísticas enganosas

Mesmo quando você revela um erro em uma estatística, não é possível determinar se este tenha sido apenas um simples e honesto engano ou se alguém estava planejando distorcer a verdade. Mas, de longe, o abuso mais comum da estatística vem em forma de um sutil, mas eficaz, exagero da verdade. Mesmo quando as somas batem, as próprias estatísticas podem ser enganosas; elas podem ser injustas, distorcer a verdade ou exagerar os fatos. É mais difícil de identificar as estatísticas enganosas do que os erros de matemática, mas elas podem causar um grande impacto na sociedade e, infelizmente, ocorrem a todo o momento.

Crimes que não compensam

Quando identificamos uma estatística enganosa, queremos questionar o tipo de estatística usada. Foi justa? Foi apropriada? Tem algum sentido prático? Se você estiver apenas preocupado em saber se as somas batem ou se os cálculos estão corretos, você pode estar perdendo um erro maior, o fato de que a estatística esteja medindo a característica errada.

As estatísticas que se referem à criminalidade são um grande exemplo de como a estatística é utilizada para mostrar os dois lados de uma história, em que apenas um é o verdadeiro. A criminalidade sempre é discutida em debates políticos, com um dos candidatos (geralmente o candidato à reeleição), argumentando que a criminalidade diminuiu durante seu mandato, enquanto o desafiante argumenta dizendo que a criminalidade aumentou (dando ao desafiante algo para criticar o candidato a reeleição). Como dois políticos podem argumentar o aumento e a diminuição da criminalidade ao mesmo tempo? Considerando que a matemática esteja correta, como isso é possível? Bom, dependendo da maneira como a criminalidade é medida é possível obter os dois resultados. A tabela 2-1 mostra o número de crimes nos Estados Unidos registrado pelo FBI de 1987 a 1997.

Tabela 2-1 Número de crimes nos Estados Unidos (1987 -1997)

Ano	*Número de Crimes*
1987	13.508.700
1988	13.923.100
1989	14.251.400

Ano	*Número de Crimes*
1990	14.475.600
1991	14.872.900
1992	14.438.200
1993	14.144.800
1994	13.989.500
1995	13.862.700
1996	13.493.900
1997	13. 175.100

A criminalidade está aumentando ou diminuindo? Parece que, em geral, está diminuindo, mas você poderia olhar esses dados de maneira diferente e apresentar esses números de modo que pareçam diferentes. A grande questão é: esses dados realmente falam a verdade?

Por exemplo, compare os números de 1987 e de 1993. Em 1987 um número estimado de 13.508.700 crimes ocorreu nos Estados Unidos e, em 1993, o número total de crimes era de 14.144.800. Parece que o índice de criminalidade aumentou durante esses seis anos. Imagine que você se candidatasse à presidência; você poderia construir sua plataforma em cima desse aparente aumento da criminalidade. E se você avançar até 1996, o número total de crimes estimado naquele ano foi de 13.493.900, pouco menos do que o número total de crimes em 1987. Então, pergunto: muito foi feito para diminuir a criminalidade durante o período de nove anos desde 1987 até 1996? Além disso, esses números não contam a história toda. O número total de crimes cometido em um dado ano pode realmente ser considerado como a estatística mais apropriada para se medir a dimensão da criminalidade nos Estados Unidos?

Outra parte importante da informação foi deixada de fora da história (e acreditem em mim, isso acontece com mais frequência do que você imagina!). Algo além do número de crimes também aumentou nos Estados Unidos no período entre 1987 a 1996: a população americana. O número total da população de um país desempenha um importante papel na estatística de criminalidade, pois quando o número de pessoas vivendo em um país aumenta, também se espera que aumente o número de criminosos e vítimas em potencial. Portanto, para colocar a criminalidade dentro de uma perspectiva, é necessário que se leve em consideração o número total de pessoas além do número de crimes. Como se faz isso? O FBI registra um índice de criminalidade que nada mais é do que uma taxa de criminalidade. Uma taxa é uma razão, ou seja, o número de pessoas ou eventos que lhe interessa, dividido pelo número total de todo o grupo.

A verdade sobre razões, taxas e porcentagens

A estatística pode ser expressa por meio de uma variedade de unidades diferentes, e essa variedade pode ser bem confusa.

- Uma razão é uma fração que divide duas quantidades. Por exemplo, "a razão de meninas por meninos é de 3 para 2", significa que para cada três meninas você encontra dois meninos. Mas isso não significa que apenas 3 meninas e 2 meninos fazem parte do grupo; razões se expressam em termos menores (simplificados ao menor termo possível). Portanto, você poderia ter 300 meninas para 200 meninos que a razão continuaria sendo 3 para 2.
- Uma taxa é uma razão que reflete uma quantidade por certa unidade. Por exemplo, seu carro faz 60 milhas por hora ou a taxa de assaltos em um bairro é de 3 assaltos por 1000 casas.
- Uma porcentagem é um número entre 0 e 100 que reflete a proporção de um todo. Por exemplo, uma camiseta está com 10% de desconto ou 35% da população é a favor de uma semana com quatro dias úteis. Para converter um porcento para um decimal, divida por 100 ou mova duas casas decimais para a esquerda. Para se lembrar disso mais facilmente, apenas se lembre que 100% é igual a 1 ou 1,00 e, para transformar 100 em 1, você precisa dividir por 100 ou mover duas casas decimais para a esquerda (é só fazer o contrário para transformar um número decimal em porcentagem).

As porcentagens podem ser utilizadas para se determinar quanto um valor aumentou ou diminuiu, em termos relativos. Suponha que os crimes em uma cidade aumentaram de 50 para 60, enquanto o número de crimes em outra cidade aumentou de 500 para 510. As duas cidades tiveram um aumento de 10% na criminalidade, mas, para a primeira cidade, essa diferença é muito maior. Para encontrar o aumento da porcentagem, faça a quantidade de "depois" menos a quantidade de "antes" e divida o resultado pela quantidade de "antes". Para a primeira cidade, isso significa que a criminalidade aumentou de (60-50) ÷ 50 = 10 ÷ 50 = 0,20 ou 20%. Para a segunda cidade, isso significa que a criminalidade aumentou apenas 2%, pois (510 - 500) ÷ 500 = 10 ÷ 500 = 0,02 ou 2%. Para encontrar a diminuição da porcentagem, faça o mesmo procedimento. Você encontrará um resultado negativo, o que indica a diminuição.

A tabela 2-2 mostra a população estimada dos Estados Unidos no período entre 1987 a 1997, juntamente com o número estimado de crimes e as taxas estimadas de criminalidade (crime a cada 100.000 habitantes).

Tabela 2-2 Número de crimes, Tamanho Estimado da população e Taxas de Criminalidade nos Estados Unidos (1987-1997)

Ano	*Número de crimes*	*Tamanho estimado da População*	*Taxa de Criminalidade (a cada 100.000 habitantes)*
1987	13.508.700	243.400.000	5.550,0
1988	13.923.100	245.807.000	5.664,2

Ano	*Número de crimes*	*Tamanho estimado da População*	*Taxa de Criminalidade (a cada 100.000 habitantes)*
1989	14.251.400	248.239.000	5.741,0
1990	14.475.600	248.710.000	5.820,3
1991	14.872.900	252.177.000	5.897,8
1992	14. 438.200	255.082.000	5.660,2
1993	14.144.800	257.908.000	5.484,4
1994	13.989.500	260.341.000	5.373,5
1995	13.862.700	262.755.000	5.275,9
1996	13.493.900	265.284.000	5.086,6
1997	13.175.100	267.637.000	4.922,7

Se compararmos novamente 1987 com 1993, veremos que o número de crimes saltou de 13.508.700 em 1987 para 14.144.800 em 1993 (note que isso representa um aumento de 4,7%, pois 14.144.800 menos 13.508.700 é igual a 636.100 que dividido pelo valor inicial de 13.508.700 é igual a 0,047 ou 4,7%). Portanto, olhando desse ponto, pode-se afirmar que a criminalidade subiu 4,7% no período de 1987 a 1993. Mas esses 4,7% representam um aumento do número total de crimes e não do número de crimes por habitante, ou do número de crimes para cada 100.000 habitantes. Para descobrir como o número de crimes por habitante mudou ao longo dos anos, é necessário que se calcule e compare as taxas de crimes para 1987 e 1993. Veja como: (5.484,4 – 5.550,0) ÷ 5.550,0 = - 65,6 ÷ 5.550,0 = - 0,012 = - 1,2%. Na verdade, os crimes por habitantes (a taxa de criminalidade) caíram 1,2%.

Dependendo de como você apresenta os números, os resultados podem ser fabricados para mostrar tendência opostas: que a criminalidade aumentou ou diminuiu entre os anos de 1987 e 1993. Mas agora que você sabe a diferença entre número de crimes e a taxa de criminalidade, você sabe que algumas estatísticas não podem simplesmente ser mostradas como o número total de eventos, mas, ao invés disso, devem ser mostradas como taxas (ou seja, o número de eventos dividido pelo número total do grupo).

LEMBRE-SE

Questione o tipo de estatística que foi utilizado antes de tentar achar um sentido para os resultados. Essa é mesmo uma medida justa e apropriada? Esse é o modo mais preciso de retratar a realidade que se esconde por trás dos dados ou há uma maneira melhor?

As escalas podem lhe dar uma mega chance!

Os gráficos são ótimas formas de se mostrar de maneira clara e rápida o ponto em que você quer chegar, desde que sejam feitos de maneira correta e honesta.

Infelizmente, muitas vezes, os gráficos que acompanham a estatística do dia-a-dia não são confeccionados de maneira correta e/ou honesta, e você precisa estar atento para identificar esses problemas. Um dos elementos mais importantes que deve ser observado é a escala em que o gráfico está. A escala de um gráfico é a quantidade utilizada para representar cada marca indicadora no eixo de um gráfico. Em que escala as marcas indicadoras aumentam? Isso pode fazer uma grande diferença em relação ao modo de ver um gráfico.

Por exemplo, a loteria do estado de Kansas frequentemente mostra os últimos resultados da Loteria Pick 3. Uma das estatísticas demonstrada é o número de vezes que cada número (de 0 a 9) é sorteado entre os três números vencedores. A tabela 2-3 mostra o número de vezes que cada número foi sorteado em 15 de março de 1997 (durante 1.613 jogos da loteria Pick 3, para um total de 4.839 números sorteados). Dependendo de como você escolhe enxergar esses resultados, mais uma vez você poderá fazer com que a estatística conte histórias muito diferentes.

Tabela 2-3 Número de Vezes que cada Número foi Sorteado (Loteria Pick 3 de Kansas em 15/03/97)

Número Sorteado	*Número de Vezes do Sorteio*
0	485
1	468
2	513
3	491
4	484
5	480
6	487
7	482
8	475
9	474

A maneira como as loterias geralmente mostram resultados como os da tabela 2-3 é ilustrada na Figura 2-1. Note que no gráfico de barras a seguir, parece que o número 1 não foi sorteado tantas vezes (apenas 468) quanto o número 2 (513). A diferença do tamanho das duas barras parece ser muito maior, exagerando a diferença entre o número de vezes em que esses dois números foram sorteados. No entanto, se pusermos esses números em perspectiva, verificaremos que a diferença real está em 513-468= 45 em um total de 4.839 números sorteados. Em se tratando do número total de cada número sorteado, a diferença entre o número de vezes que o número 1 foi sorteado em relação ao número 2 é de 45 ÷ 4.839 = 0,009, ou seja, apenas nove décimos de um por cento.

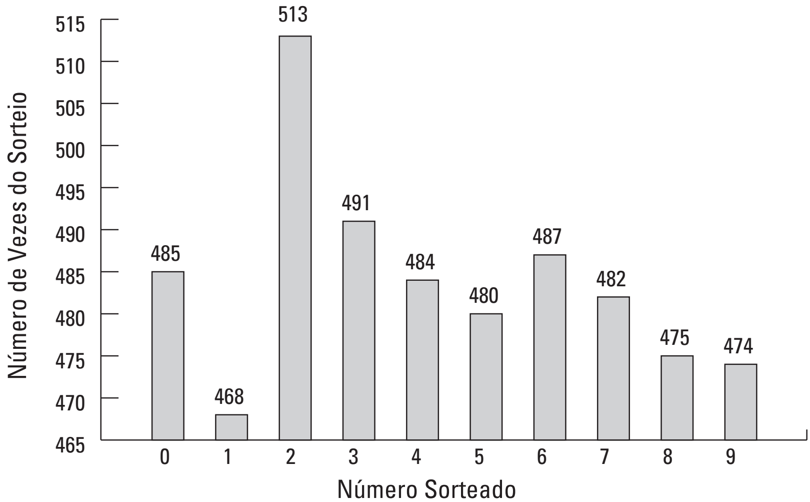


Figura 2-1: Gráfico de barras mostrando o número de vezes que cada número foi sorteado.

O que faz com que esse gráfico exagere as diferenças? Duas questões entram em cena agora, ambas com relação à aparência do gráfico. Primeiro, note que o eixo vertical mostra o número de vezes (frequência) que cada número foi sorteado e que sua escala aumenta de 5 em cinco. Assim, uma diferença de 5 em um total de 4.839 números sorteados aparece como se realmente significasse alguma coisa. Esse é um truque comum para exagerar os resultados – distorcer a escala para que as diferenças pareçam maiores do que realmente são. Em segundo lugar, o gráfico não começa a contar do zero, mas, sim, a partir de 465, portanto, ele realmente está apenas mostrando a parte superior de cada barra, onde as diferenças se encontram. Esse efeito também exagera o resultado.

A tabela 2-4 mostra um resumo mais real para cada um dos números sorteados na Loteria Pick 3, ao mostrar a porcentagem de vezes que cada número foi sorteado.

Tabela 2-4 Porcentagem de Vezes que Cada Número Foi Sorteado

Número Sorteado	*Número de Vezes do Sorteio*	*Porcentagem de Vezes do Sorteio*
0	485	10,0% = 485 ÷ 4.839
1	468	9,7% = 468 ÷ 4.839
2	513	10,6% = 513 ÷ 4.839
3	491	10,1% = 491 ÷ 4.839
4	484	10,0% = 484 ÷ 4.839
5	480	9,9% = 480 ÷ 4.839
6	487	10,0% = 487 ÷ 4.839
7	482	10,0% = 482 ÷ 4.839
8	475	9,8% = 475 ÷ 4.839
9	474	9,8% = 474 ÷ 4.839

A Figura 2-2 é um gráfico de barras que mostra a porcentagem de vezes que cada número foi sorteado, ao invés de mostrar o número de vezes que cada número foi sorteado. Note que o gráfico a seguir também utiliza uma escala mais realista do que a utilizada na Figura 2-1 e que começa do zero, essas duas alterações fazem com que as diferenças pareçam o que realmente são – não muito diferentes. Chato, né?

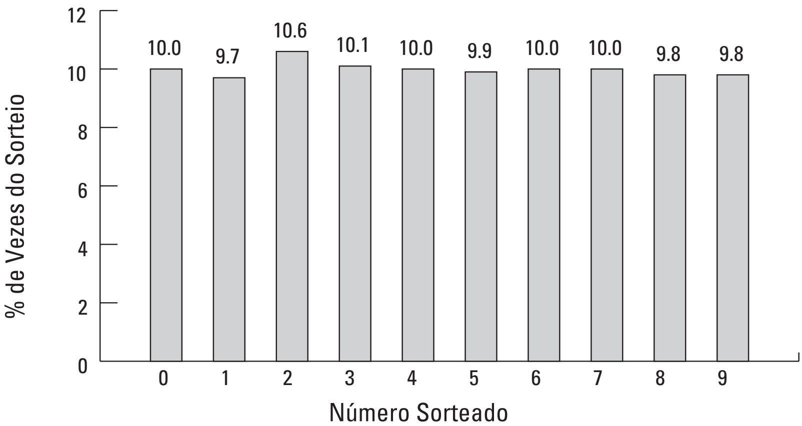


Figura 2-2: Gráfico de barras mostrando a porcentagem de vezes que cada número foi sorteado

Agora, por que a loteria faria isso? Talvez ela queira que você acredite que esteja conseguindo informações privilegiadas e que jogue no número 1 por acreditar que ele pode “sair” para você apesar do suposto fato de que ele não seja tão sorteado quanto os outros números (o que, por sinal, não é verdade; veja o Capítulo 7 para saber mais sobre o assunto). Ou, talvez, ela queira que você jogue no número 2, o número que mais foi sorteado e que com certeza sairá para você (mais uma vez, sem chances). Seja como for, a turma da loteria quer que você pense que exista algo “mágico” por trás dos números, e, assim, você não pode culpá-los; esse é o trabalho deles.

Gráficos enganosos aparecem o tempo todo na mídia! As escalas podem ser aumentadas (fazendo com que as marcas indicadoras representem pequenas quantidades de uma maneira maior) e/ou os gráficos podem se iniciar a partir de números diferentes do zero, para fazer com que as diferenças pareçam maiores do que realmente são. As escalas também podem ser comprimidas (fazendo com que as marcas indicadoras representem grandes quantidades de maneira muito menor) para dar a impressão de que “não houve alterações”. Esses são exemplos de representações falsas da verdade (veja o Capítulo 4 para obter mais informações sobre como identificar gráficos enganosos).

Observar a escala de um gráfico realmente pode lhe ajudar a manter os resultados em perspectiva.

Verificado suas fontes

Verifique a fonte de informação. Os melhores resultados geralmente são publicados por um periódico científico reconhecido por especialistas da área. Por exemplo, na área médica, entre os periódicos de maior reputação, nos quais médicos respeitados publicam os resultados de suas pesquisas e leem sobre as últimas descobertas, temos o Journal of the American Medical Association (JAMA), o New England Journal of Medicine, The Lancet e o British Medical Journal.

Ao examinar os resultados de qualquer estudo, leve em consideração a fonte e questione todos os estudos que foram conduzidos, e não apenas os estudos cujos resultados foram publicados em periódicos acadêmicos ou apareceram em propagandas. Um conflito de interesses por parte dos pesquisadores pode levar à informação incorreta.

Avaliando o tamanho amostral

O tamanho amostral não é tudo, mas, realmente, tem um papel importante em se tratando de pesquisas de opinião e estudos. Se o estudo for planejado e executado corretamente e se os participantes forem selecionados aleatoriamente (ou seja, de maneira não tendenciosa, veja os Capítulos 16 e 17 para mais detalhes sobre planejamento e execução de estudos).

Você deve estar pensando que todos os estudos fundamentam-se em grandes números de participantes. Isso realmente acontece com a maioria, mas nem sempre é verdade para outros tipos de pesquisa, tais como os estudos que envolvem experimentos cuidadosamente controlados. Experimentos consomem muito tempo, às vezes podem levar meses ou anos para realizá-los dentro de uma variedade de situações. Os estudos experimentais também podem ser caros. Alguns experimentos consistem em avaliar produtos, tais como chips de computador e equipamentos militares que custam milhares e, até mesmo, milhões de dólares. Caso o experimento consista na destruição do produto durante o processo de teste, o custo de cada experimento pode ser ainda mais alto. Em virtude do alto custo de alguns tipos de pesquisa, alguns estudos fundamentam-se em pequenos números de participantes ou produtos. No entanto, quanto menos o número de participantes ou de produtos testados, menos informação se tem no geral. Portanto, estudos com números pequenos de participantes (ou produtos) normalmente são menos precisos do que estudos semelhantes realizados com um tamanho amostral maior.

A maioria dos pesquisadores tenta incluir o maior tamanho amostral possível, equilibrando o custo deste com a necessidade de precisão. Às vezes, no entanto, as pessoas são um pouco negligentes e não se esforçam para obter uma amostra de tamanho suficiente. Outras vezes, esses pesquisadores realmente não entendem as ramificações de se ter uma amostra pequena. E alguns esperam que você não entenda a importância do tamanho amostral, mas agora você já sabe.

Os piores exemplos de tamanhos amostrais inadequados que já vi são os mostrados pelas propagandas da TV, em que o tamanho amostral consiste em apenas um. Geralmente, esses comerciais apresentam algo que se parece com um experimento para tentar persuadir os telespectadores a acreditar que um produto é superior ao outro. Você provavelmente já assistiu anúncios de TV que comparam duas marcas de papel toalha, em que um pedaço de cada marca de papel toalha é utilizado para absorver a mesma quantidade de suco. Esses exemplos podem parecer bobos, mas qualquer um pode cair facilmente na armadilha de tirar conclusões baseadas em um tamanho amostral de um (você já falou para alguém não comprar um determinado produto porque você não teve sorte com ele?). Lembre-se de que uma anedota (uma história) é, na verdade, um estudo com o tamanho amostral igual a um.

Verifique o tamanho amostral para garantir que os pesquisadores obtiveram informação o suficiente para fundamentar seus resultados. A margem de erro (veja o Capítulo 10) também lhe dá uma ideia do tamanho amostral, uma vez que, na maioria das vezes, uma pequena margem de erro significa que o estudo utilizou uma grande amostragem.

Qualidade ou Quantidade?

As manchetes são o arroz com feijão da mídia, mas também podem ser enganosas. Na maioria das vezes, as manchetes são mais grandiosas do que a "verdadeira" informação, em especial quando as histórias envolvem estatísticas e os estudos que geraram as estatísticas. Na realidade, você frequentemente encontrará verdadeiras lacunas entre a manchete e as "letrinhas miúdas" em tais histórias.

Um estudo conduzido há poucos anos avaliava 1.265 consultas gravadas em vídeo com 59 médicos de pronto-socorro e com 6 cirurgiões no

Colorado e em Oregon. Este estudo demonstrou que os médicos que não foram processados por erro médico haviam gasto em média 18 minutos com cada paciente, enquanto os médicos que foram processados tinham gasto 16 minutos em cada consulta.

Opa, mas dois minutos são realmente tão importantes assim? O estudo foi noticiado pela mídia sob a manchete: "Tempo de consulta com pacientes evita processos por imperícia médica"; dando a entender que se você é um médico que foi processado, tudo o que precisa fazer é passar mais tempo com seus pacientes para se livrar do problema.

O que realmente está acontecendo? Eu tenho mesmo que acreditar que tudo que um médico que tenha sido processado precisa fazer para estar a salvo é demorar dois minutos a mais em cada consulta? Pense a respeito de outras possibilidades que possam estar envolvidas aqui. O caso pode ser que os médicos que não foram processados, simplesmente sejam médicos melhores, que perguntam mais, ouvem mais e dizem mais a seus pacientes sobre o que esperar e, por isso, acabam gastando mais tempo em suas consultas; se for isso mesmo, o que o médico faz durante esse tempo conta muito mais do que o tempo que ele realmente gasta com cada paciente. Mas e essa outra possibilidade: talvez os médicos que foram processados estavam realizando procedimentos de maior risco ou talvez eles fossem algum tipo de especialista. Infelizmente, o artigo não dá essa informação. Uma terceira possibilidade é a de que os médicos que não foram processados tenham menos pacientes e, portanto, tenham mais tempo para dedicar a cada consulta e conseguir monitorar melhor seus pacientes. De qualquer forma, a notícia escrita em letra miúda, nesse caso, não se mostra adequada à manchete em letras garrafais e, quando você ler ou ouvir histórias como essa, procure por lacunas semelhantes entre o que está escrito na manchete e o que realmente foi descoberto pelo estudo.

Muito além do alcance

Você deve estar se perguntando como os candidatos a cargos políticos descobrem as opiniões dos eleitores. Eles realizam enquetes e pesquisas de opinião. Muitas dessas pesquisas são conduzidas por grupos independentes, tal como o Instituto Gallup; outras são feitas por representantes dos próprios candidatos, e os métodos podem se diferenciar em muito de um candidato para outro ou de uma pesquisa para outra.

Na corrida para a eleição presidencial norte americana de 1992, Ross Perot causou furor no cenário político. O seu grupo, United We Stand America, ganhou força e, no final das contas, Ross Perot e seus partidários causaram impacto nos resultados da eleição. Com frequência durante debates e discursos, Perot apontava estatísticas e tirava conclusões a respeito do que pensavam os cidadãos americanos sobre certos assuntos. Mas, o senhor Perot estava sendo sempre certo com relação ao que os americanos pensavam ou ele simplesmente estava certo sobre o que pensavam seus partidários? Um dos veículos que Ross Perot utilizou para saber das opiniões dos americanos foi a publicação de um questionário em 21 de março de 1992 juntamente com o Guia de TV, pedindo às pessoas para que o preenchessem e enviassem-no para o endereço fornecido. Depois ele compilava os resultados da enquete e os utilizava como parte de sua plataforma de campanha. A partir desses resultados, ele concluía que 80% dos americanos concordavam com ele em determinados assuntos (observe, no entanto, que ele recebeu apenas 18,91% dos votos em 1992).

Parte do problema com as declarações do senhor Perot está na maneira como a pesquisa foi conduzida. Para responder ao questionário, você tinha que comprar o Guia de TV, ou seja, você teria que realmente estar muito afim de preencher o questionário e ainda tinha que pagar o selo para enviá-lo por correio. Quem, provavelmente, faria isso? Aqueles que tivessem opiniões muito fortes. Além disso, a maneira como estavam escritas as perguntas encorajava às pessoas que concordavam com ele a responder e enviar o questionário; os que não concordavam com ele, no entanto, provavelmente ignoravam a pesquisa.

DICA

Se você conseguir perceber como o pesquisador quer que você responda um questionário, baseando-se na maneira como as perguntas foram escritas, você saberá que está diante de uma pergunta capciosa (veja o Capítulo 16 para mais informações sobre como identificar esse e outros problemas com pesquisas).

Aqui estão algumas perguntas feitas pelo senhor Perot em seu questionário. Eu as parafraseei, mas a intenção original está intacta (não entenda isso como uma crítica a Ross Perot; muitos candidatos e seus partidários fazem o mesmo tipo de coisa).

- Você acha que o line-item veto[1] deveria poder ser utilizado pelo presidente para acabar com gastos desnecessários?

- Você acha que o Congresso deveria excluir a si mesmo da legislação que faz para nós?

- Você acha que os principais novos programas deveriam ser apresentados em detalhes primeiramente à população?

As pessoas que sabiam da pesquisa e que escolhiam participar dela provavelmente seriam as que mais se identificavam com Ross Perot. Esse é um exemplo em que as conclusões do um estudo foram além de seu escopo, pois os resultados não representavam a opinião de "todos os americanos", tal como alguns eleitores foram levados a acreditar. Como seria possível conseguir a opinião de todos os americanos? Seria necessário realizar uma pesquisa muito bem planejada e implantada, que se fundamentasse em uma amostra de pessoas escolhida aleatoriamente (veja o Capítulo 16 para mais informações sobre como conduzir uma pesquisa).

FIQUE ATENTO

Ao examinar as conclusões de qualquer estudo, examine bem os grupos que foram realmente estudados (ou o grupo que realmente participou) e o maior grupo de pessoas (ou ratos de laboratório, ou moscas, dependendo do estudo) que o grupo estudado deveria representar. Depois, verifique as conclusões tiradas. Veja se elas se encaixam. Se não, tenha certeza de que você entendeu quais as reais conclusões e, antes de tomar qualquer decisão, seja realista com relação às afirmações feitas.

Procurando mentiras nos lugares certos

Você viu exemplos de enganos honestos que levam a problemas e de como a distorção, o aumento e o exagero da verdade também podem causar danos. Eventualmente, você também pode se deparar com situações em que as estatísticas são, simplesmente, inventadas, fabricadas ou forjadas. Isso não acontece muito, graças a revisores, comitês de supervisão e as leis e regulamentações governamentais.

Mas, uma vez ou outra, você escuta falar de alguém que falsificou dados ou "forjou números". Provavelmente, a mentira mais comumente cometida envolvendo estatística e dados ocorre quando são jogados fora os dados que não atendem às hipóteses formuladas por alguém, não se enquadram com padrões ou parecem ser "outliers". Nos casos de erros claros (por exemplo, a idade de uma pessoa foi grafada como sendo de 200 anos) faz sentido tentar arrumar o dado, removendo o dado incorreto ou tentando corrigir o erro. Mas não se deve, simplesmente, jogar uma parte do dado fora somente porque ele não atende seu ponto de vista. A eliminação de dados (exceto em caso de erro documentado) é eticamente condenável; ainda assim, acontece.

Ainda com relação à falta de dados de experimentos, uma frase comumente utilizada é "Entre aqueles que completaram o estudo..." Mas e os que não completaram o estudo, especialmente em se tratando de um estudo médico? Eles morreram? Não suportaram os efeitos colaterais das drogas e abandonaram o experimento? Sentiram-se pressionados a dar certas respostas para corroborar as hipóteses dos cientistas? Desistiram, pois se sentiram frustrados com a duração do experimento e pela sensação de não estarem melhorando?

Nem todo mundo responde a pesquisas, e até mesmo, as pessoas que tentam participar de algumas, às vezes, descobrem que não têm tempo nem interesse de responder a cada uma das muitas pesquisas com as quais são bombardeados. Atualmente, a sociedade americana está louca por pesquisas e, dificilmente, passa-se um mês sem que você seja convidado a responder a uma pesquisa feita por telefone, pela Internet ou por correspondência sobre tópicos que variam de preferência por produtos até sua opinião com relação à nova lei que regula o latido de cães na vizinhança. Os resultados dessas pesquisas são concluídos a partir das pessoas que realmente responderam, e as opiniões daqueles que optaram por não responder pode ser muito diferente das opiniões daqueles que optaram por responder. No entanto, se os pesquisadores darão um jeito ou não de contar isso a você, isso é outra história.

Por exemplo, alguém pode lhe dizer que distribuiu 5.000 questionários, recebeu 1.000 de volta e baseou os resultados nessas respostas. Você pode então pensar: "Nossa, 1.000 respostas. É muito dado; essa deve ser uma pesquisa muito precisa". Errado! O problema é que 4.000 pessoas selecionadas para participar da pesquisa optaram por não responder, e você não tem a menor ideia do que elas teriam falado se tivessem respondido. Você não tem garantia de que as opiniões dessas 4.000 pessoas estejam representadas pelas opiniões das que responderam. Na realidade, o contrário poderia ser verdadeiro.

DICA

O que poderia ser considerada uma alta taxa de respostas? (ou seja, o número de pessoas que responderam ao questionário dividido pelo número de questionários distribuídos). Alguns estatísticos estabeleceriam nada menos do que 70%, mas, como diria o Dr. Phil da TV, os estatísticos precisam "cair na real". Raramente uma pesquisa alcança uma taxa de respostas tão elevada. Mas, no geral, quanto menor for a taxa de respostas, menor é a credibilidade dos resultados e mais os resultados favoreceram as opiniões dos que responderam (e lembre-se que os respondentes tendem a ter opiniões muito fortes sobre os assuntos pelos quais optam a responder).

LEMBRE-SE

Para identificar dados falsos ou ausentes, procure informações a respeito do estudo, incluindo o número de pessoas selecionadas para participar, o número de pessoas que completaram o estudo e o que aconteceu com todos os participantes, não apenas com os que obtiveram um resultado positivo.

Sentindo o Impacto das Estatísticas Enganosas

Como as estatísticas enganosas afetam sua vida? Elas podem lhe afetar em grandes ou pequenas proporções, dependendo do tipo de estatística que atravessa seu caminho e do que você escolhe fazer com a informação que lhe é dada. A maneira mais significativa com que a estatística (boa ou má) lhe afeta é em seu processo de tomada de decisão diário.

Pense nos exemplos discutidos até agora e como eles afetariam sua tomada de decisão. Você provavelmente não vai ficar acordado até tarde da noite se perguntando se os 14% daqueles entrevistados realmente aquecem as sobras de alimento no forno micro-ondas. Mas talvez você passe por situações envolvendo estatística que possam lhe afetar de modo bastante significativo e, por isso, você precisa estar preparado e apto para separar tudo. Aqui vão alguns exemplos:

- Alguém poderá tentar lhe dizer que quatro em cada cinco pessoas entrevistadas concordam com o aumento dos impostos e, portanto, você também deveria concordar! Você se sentiria pressionado ou tentaria primeiro encontrar mais informações? (Será que você é um "Maria-vai-com-as-outras?").

- Um candidato envia-lhe um panfleto mostrando informações sobre a campanha baseado em estatísticas. Você acreditaria no que ele está dizendo?

- Se algum dia você for chamado para ser jurado, existe a possibilidade de que você se depare com um advogado utilizando estatística como parte de seu argumento. Você terá que decifrar toda essas informações e determinar se a evidência está convencendo "além de uma dúvida razoável". Em outras palavras, qual é a probabilidade do réu ser culpado? (Para mais sobre como interpretar probabilidades, veja os Capítulos 7 e 8).

- O noticiário do rádio diz que telefones celulares causam tumores no cérebro. Sua esposa ou esposo usa o celular o tempo todo. Será que você deve ser preocupar?
- E aquelas campanhas publicitárias de remédios que nunca acabam? Imagine como os médicos sentem-se pressionados por seus pacientes que acabam totalmente convencidos pelos anúncios de que necessitam daquele medicamente imediatamente. Ser informado é uma coisa, outra coisa é sentir-se informado por causa de um anúncio feito pelo fabricante de um produto.
- Caso você tenha algum problema médico, ou conheça alguém que tenha, provavelmente está sempre em busca de novos tratamentos e terapias que possam ser úteis. No entanto, o mundo da medicina está repleto de estatísticas que podem ser muito confusas.

Na vida, você pode se deparar com tudo: desde um inocente engano aritmético até exageros e distorções da verdade, falsificação e mineração de dados, além de relatórios que, por conveniência, omitem informações ou comunicam apenas a parte dos resultados que convenha ao pesquisador. Ao mesmo tempo em que é preciso enfatizar que nem todas as estatísticas são enganosas e que nem todo mundo age de má fé, também enfatizo que é preciso estar vigilante. Separe a boa informação da informação suspeita e ruim, de modo que você possa se manter afastado de estatísticas erradas.

Estatísticas para a vida cotidiana

Você toma decisões todos os dias baseado nas estatísticas e nos estudos estatísticos que ouviu ou viu, sem nem mesmo se dar conta disso. Eis alguns exemplos dos tipos de decisão que você toma ao longo de um dia:

- “Será que devo usar botas hoje? O que a previsão do tempo disse ontem à noite? Ah é, 30% de chances de nevar”.
- “Qual a quantidade de água que se deve tomar diariamente? Eu pensava que o ideal era 2 litros por dia, mas agora ouvi dizer que tomar muita água pode ser prejudicial para mim!”
- “Será que devo comprar aquelas vitaminas que minha amiga me recomendou? Mary diz que as vitaminas funcionam para ela, mas geralmente fazem mal ao meu estômago” (afinal, qual é a melhor hora para tomá-las?).
- “Minha cabeça está começando a doer; acho que devo tomar uma aspirina. Talvez eu devesse ficar um pouco mais no sol, ouvi dizer que isso ajuda a evitar a enxaqueca.”
- “Ah, espero que meu cachorro não tenha comido meus tapetes de novo. Ouvi dizer que dar Prozac aos cães os ajuda a lidar melhor com a ansiedade da separação do dono. Mas um cão usando Prozac? Como foi possível determinar a dosagem correta? E o que vou dizer aos meus amigos?”
- “O que vou comer no almoço hoje? Passar no drive-thru de novo? Eu ouvi alguma coisa sobre ‘colesterol ruim’. Mas eu suponho que todos fast-food são a mesma coisa — péssimos para a saúde, certo?”

Estatísticas para a vida cotidiana (cont.)

- "Gostaria de saber quando o chefe vai começar a tomar medidas mais enérgicas com os funcionários que checam seus e-mails pessoais. Ouvi falar sobre um estudo que mostrou que as pessoas passam em média duas horas por dia checando e enviando e-mails pessoais na hora do expediente. Eu é que não passo tanto tempo fazendo isso!"
- Vira e mexe tem alguém costurando o trânsito e falando no celular! Queria saber quando esses celulares serão realmente banidos! Tenho certeza de que eles são a causa de um grande número de acidentes!"

Nem todos os exemplos envolvem números, ainda assim, todos eles envolvem um assunto chamado estatística. A estatística realmente se refere ao processo de tomada de decisões, de teste de teorias, de comparação entre grupos ou tratamentos e de questionamento. Os cálculos numéricos ficam nos bastidores, deixando impressões e conclusões duradouras que, no final das contas, ficam embutidas em suas decisões diárias.

Capítulo 3

Os Segredos do Ofício

Neste Capítulo

- Entendendo a estatística como um processo e não apenas números.
- Familiarizando-se com alguns termos do jargão estatístico.

Na atual explosão de números, a palavra do momento é "dados", como em: "Você tem dados suficiente para corroborar sua afirmação?"; "Que dados você tem sobre o assunto?", "Os dados corroboraram a hipótese original de que ..."; "Dados estatísticos mostram que..." e "Os dados dão suporte a..." No entanto, o campo da estatística não se resume apenas a dados. A estatística é todo o processo envolvido na coleta de evidências para responder às perguntas feitas pelo mundo, nos casos em que tais evidências venham em forma de dados numéricos.

Neste capítulo, você verá em primeira mão como a estatística funciona como um processo e onde os números entram em cena. Você também ficará por dentro dos termos mais utilizados do jargão estatístico e entenderá como todas essas definições e conceitos se encaixam como partes desse processo. Assim, da próxima vez em que você ouvir alguém dizendo, "Esta pesquisa tem uma margem de erro de 3 pontos percentuais tanto para mais quanto para menos", você terá uma ideia do que isso significa.

Estatística: Mais do que Apenas Números

A maioria dos estatísticos não quer que a estatística seja vista como "apenas números". Enquanto o resto do mundo a vê dessa forma, os estatísticos não se veem como devoradores de números; na maioria das vezes, eles se veem como guardiões do método científico (é claro que os estatísticos dependem de que especialistas de outras áreas formulem perguntas interessantes, pois não é só de estatística que vive o homem). O método científico (formulação de perguntas, realização de estudos, coleta de evidências, análise de evidências e conclusões) é algo que você já deve ter visto antes, mas, ainda assim, deve estar se perguntando o que esse método tem a ver com estatística.

Todas as pesquisas se iniciam a partir de uma pergunta como:

- É possível beber muita água?
- Qual o custo de vida em San Francisco?
- Quem vencerá a próxima eleição presidencial?
- As plantas consideradas medicinais realmente auxiliam a manter uma boa saúde?
- Meu programa de TV favorito continuará a ser exibido no ano que vem?

Nenhuma das perguntas anteriores refere-se diretamente a números. Ainda que, para que se possa encontrar uma resposta a cada uma delas, seja necessário que se faça uso de dados e de processos estatísticos.

Suponha que um pesquisador queira determinar quem ganhará a próxima eleição presidencial dos Estados Unidos. Para responder a essa pergunta com o máximo de precisão, o pesquisador terá que seguir vários passos:

1. **Determinar o grupo de pessoas a ser estudado.**

   Nesse caso, o pesquisador utilizaria os eleitores registrados que estivessem planejando votar na próxima eleição.

2. **Coletar dados.**

   Esse passo constitui um desafio, pois você não pode, simplesmente, sair às ruas perguntando a todo cidadão americano se ele ou ela planeja votar na próxima eleição e, se fosse possível, em quem eles votariam. Além disso, suponha que alguém lhe responda: "Sim, eu planejo votar". Será que aquela pessoa realmente irá votar no dia da eleição? E será que aquela pessoa irá lhe contar para quem ela realmente pretende votar? E se aquela pessoa mudar de ideia e votar em outro candidato?

3. **Organizar, resumir e analisar os dados**

   Depois que o pesquisador saiu e conseguiu os dados de que ele precisava, pegando-os organizados, resumidos e analisados, isso o ajudará a responder sua pergunta. É o que muitas pessoas reconhecem como sendo a estatística propriamente dita.

4. **Reunir todos os resumos de dados, gráficos e tabelas, análises e chegar a uma conclusão a partir deles para tentar responder à pergunta inicial.**

   É claro que o pesquisador não será capaz de obter 100% de precisão para sua resposta, pois não foi possível entrevistar toda a população americana. Mas ele pode sim obter uma resposta que se aproxime bastante dos 100%. De fato, com uma amostra de 2.500 pessoas que tenham sido selecionadas aleatoriamente e de modo não tendencioso (para que cada membro da população tenha a mesma chance de ser selecionado), o pesquisador pode chegar a resultados precisos, dentro de uma margem de erro de 2,5% para mais ou para menos (isto é, se todos os passos do processo de pesquisa tiverem sido executados corretamente).

Ao tirar conclusões, o pesquisador tem que estar consciente de que todo estudo é limitado e que – por sempre haver uma chance de erro – os resultados podem estar errados. Um valor numérico pode ser relatado para informar aos outros a confiabilidade que o pesquisador deposita nos resultados e o grau de precisão que se espera de tais resultados (veja o Capítulo 10 para mais informações a respeito da margem de erro).

Depois que a pesquisa é feita e a pergunta respondida, normalmente os resultados levam a mais perguntas e a mais pesquisa. Por exemplo, se os homens pareceram estar mais a favor do candidato Y e as mulheres mais a favor de seu oponente, as próximas perguntas poderiam ser: "Quem vai mais às urnas no dia da eleição – homens ou mulheres – e quais fatores vão determinar a maioria de homens ou mulheres na votação?"

A estatística realmente se constitui em usar o método científico para responder as perguntas feitas pelo mundo. Os métodos estatísticos estão envolvidos em todos os passos de um bom estudo, desde o planejamento da pesquisa, na coleta de dados, na organização e no resumo da informação para se fazer uma análise até a chegada a uma conclusão, na discussão das limitações e, finalmente, no planejamento de novos estudos para responder as novas perguntas que surgiram. A estatística é mais do que apenas números, ela é um processo!

Entendendo Alguns Termos Básicos do Jargão Estatístico

Toda atividade tem suas ferramentas básicas e com a estatística não é diferente. Se você pensar em um processo estatístico com sendo uma série de estágios em que se parte de uma pergunta para uma resposta, você deve imaginar que, em cada estágio, será encontrado um conjunto de ferramentas e termos (ou jargão estatístico). Agora, se você começou a se arrepiar só em imaginar, não se preocupe. Ninguém está pedindo que você se torne um expert em estatística e atire-se em um treinamento intensivo, e ninguém está lhe pedindo para tornar-se um nerd em estatística e passar a usar o jargão o tempo todo. Você também não precisa carregar uma calculadora e um protetor de bolso para canetas no bolso esquerdo da camisa, como os estatísticos geralmente fazem.

No entanto, como o mundo está cada vez mais consciente dos números, os termos estatísticos são jogados cada vez mais na mídia e no ambiente de trabalho. Assim, saber o que esses termos realmente significam pode lhe ser muito útil. Além disso, se você está lendo esse livro, é porque quer conhecer mais sobre como calcular estatísticas básicas. O primeiro passo é entender mais alguns dos termos básicos do jargão estatístico. Portanto, nesta seção, você vai experimentar um pouquinho do jargão estatístico; eu o enviarei aos capítulos mais apropriados para obter mais detalhes.

População

Para praticamente qualquer pergunta que você queira investigar sobre o mundo, é necessário centrar sua atenção em um determinado grupo de indivíduos (por exemplo, um determinado grupo de pessoas, cidades, animais, tipos de rochas, pontuação em exames, e assim por diante). Por exemplo:

- O que a população americana acha sobre a política externa do presidente?
- Qual a porcentagem de lavouras destruídas por veados em Wisconsin no ano passado?
- Qual o prognóstico para as pacientes com câncer de mama que estão tomando medicamentos experimentais?
- Qual a porcentagem de que todos os tubos de pasta de dente sejam enchidos de acordo com suas especificações?

Em cada um desses exemplos, uma questão foi colocada. E, em cada caso, você pode identificar um grupo específico de indivíduos que está sendo estudado: a população americana, todas as lavouras em Wisconsin, todas as pacientes com câncer de mama e todos os tubos de pasta de dente que estão sendo enchidos, respectivamente. O grupo de indivíduos que você deseja estudar para obter a resposta à sua pergunta científica é chamado de população. Entretanto, as populações podem ser algo difícil de identificar. Em um bom estudo, os pesquisadores definem suas populações de maneira muito clara, enquanto em um estudo ruim, a população mal é definida.

A pergunta sobre se os bebês dormem melhor com música é um bom exemplo de como pode ser difícil determinar uma população. Como poderíamos definir um bebê exatamente? Como menos de três meses de idade? Com menos de um ano? E você quer estudar apenas os bebês dos Estados Unidos ou quer estudar os bebês do mundo todo? Os resultados podem ser diferentes para bebês de diferentes idades e para os bebês americanos versus os europeus versus os africanos e assim por diante.

Muitas vezes, os pesquisadores querem estudar e tirar conclusões a partir de populações muito abrangentes, mas no final, para economizar recursos financeiros, tempo ou, simplesmente, porque não têm outra saída, acabam estudando uma população bem restrita. Isso pode ocasionar grandes problemas para se chegar a conclusões. Por exemplo, suponha que um professor de faculdade queira saber como os comerciais de TV persuadem os consumidores a comprarem determinados produtos. Esse estudo se fundamentará em um grupo de alunos desse professor, que participaram a fim de garantir cinco pontos extras (você sabe como é!). Essa pode ser uma amostra adequada, mas os resultados do estudo desse professor não podem ser generalizados para quaisquer populações além de seus alunos, pois nenhuma outra população foi representada no estudo.

Amostra

Quando você quer saber se sopa ficou boa, o que você faz? Mexe a panela, retira um pouco com uma colher e prova. Depois tira uma conclusão sobre todo o conteúdo da panela sem, na verdade, ter provado tudo. Se a sua amostra for retirada de maneira justa (por exemplo, sem que você tenha escolhido todos os ingredientes mais saborosos) é possível ter uma ideia de com a sopa está sem ter que comer tudo. Isso é o que se faz na estatística. Alguns pesquisadores querem descobrir algo sobre uma população, mas eles não têm tempo nem dinheiro suficiente para estudar cada indivíduo dessa população. Então, o que eles fazem? Selecionam um pequeno número de indivíduos da população, estudam esses indivíduos e utilizam a informação para chegar a conclusões sobre toda a população. Isso se chama amostra.

Parece simples e fácil, certo? Infelizmente não é. Perceba que eu disse selecionam uma amostra. Isso parece ser um processo simples, mas, na verdade, não é. O modo como uma amostra é selecionada pode significar a diferença entre resultados que sejam corretos e justos de resultados sujos. Como exemplo, suponha que você queira colher uma amostra da opinião de adolescentes sobre o tempo que eles passam na Internet. Se você enviar a pesquisa por e-mail, seus resultados não representarão a opinião de todos os adolescentes, que é a sua população pretendida. Os resultados representarão apenas aqueles adolescentes que têm acesso à Internet. Esse tipo de incompatibilidade estatística é comum? Pode apostar que sim.

Uma das maiores culpadas pela má representação estatística causada por amostras ruins são as enquetes feitas pela Internet. É possível deparar-se com milhares de enquetes na Internet onde, para participar, é preciso acessar um site e dar sua opinião. Mesmo que 50.000 pessoas nos Estados Unidos participem de uma enquete na Internet, esse número não representa toda a população americana, ele apenas representa as pessoas que têm acesso à Internet, que estavam em determinado site e que estavam bastante interessados em participar da enquete (o que, normalmente, significa que essas pessoas possuem opiniões muito fortes com relação ao tópico abordado).

Da próxima vez que você for bombardeado com os resultados de um estudo, descubra informações sobre amostra de participantes e se pergunte se essa amostra realmente representa a população pretendida. Tenha cautela com relação a quaisquer conclusões tiradas a partir de uma população mais abrangente do que o que está realmente sendo estudado (mais no Capítulo 16).

Aleatoriedade

Uma amostra aleatória é boa; ela dá a mesma chance a todos os membros de uma população de ser selecionado, além de utilizar alguns mecanismos de casualidade para escolhê-los. O que isso realmente significa é que as pessoas não selecionam a si mesmas para participar e ninguém é favorecido em detrimento de outra pessoa no processo de seleção.

Um exemplo de como especialistas fazem isso é a maneira que o Instituto Gallup realiza seu processo de amostra aleatória. Tudo começa com uma lista computadorizada de todas as centrais telefônicas nos Estados Unidos, junto com a estimativa do número de residências familiares pertencentes a essas centrais. O computador utiliza um procedimento chamado discagem digital aleatória (DDA) para criar aleatoriamente números de telefones a partir dessas centrais e, então, seleciona amostras de números de telefones. Portanto, o que realmente acontece é que o computador cria uma lista de todos os possíveis números de telefone de residências familiares nos Estados Unidos e, depois, seleciona um subgrupo de números dessa lista para os quais o Instituto Gallup telefona (observe que alguns desses números de telefones podem ainda nem terem sido designados para uma residência, criando questões logísticas com as quais o instituto deve lidar).

Outro exemplo de amostragem aleatória envolve o setor de manufatura e o conceito de controle de qualidade. A maioria dos fabricantes segue especificações rígidas para a fabricação de seus produtos e erros no processo podem lhe custar dinheiro, tempo e credibilidade. Muitas empresas tentam frear os problemas, antes que eles se tornem grandes demais, através da monitoração do processo e do uso da estatística para tomar decisões para saber se o processo está correndo da maneira esperada ou se necessita ser interrompido. Para saber mais sobre controle de qualidade e estatística, veja o Capítulo 19.

Exemplos de amostra não aleatória (ou seja, amostras ruins) incluem amostras de enquetes para as quais você liga para dar sua opinião. Essa não é uma amostra verdadeiramente aleatória, pois não dá toda a população uma oportunidade igual de participar da pesquisa (se você tiver que comprar um jornal ou assistir um programa na TV e, ainda, concordar em escrever ou telefonar, isso lhe dá uma grande pista de que o processo de amostragem não é aleatório). Para saber mais sobre amostragem e enquetes, veja o Capítulo 16.

Todas as vezes em que você olhar os resultados de um estudo que se baseia em uma amostra de indivíduos, leia as letras miúdas e procure pelo termo "amostra aleatória". Se você encontrá-lo, vá mais fundo e procure saber como a amostra realmente foi selecionada e utilize a definição mencionada para verificar se a amostra foi mesmo selecionada aleatoriamente.

Viés

Viés é uma palavra que você ouve a todo o momento e, provavelmente, sabe que ela se refere a algo ruim. Mas o que realmente constitui um viés? O viés é um favoritismo sistemático presente no processo de coleta de dados, causando resultados enganosos e assimétricos.

O viés pode ocorrer de muitas maneiras:

- **Da maneira com a amostra é selecionada:** por exemplo, se você quiser uma estimativa de quanto as pessoas em sua comunidade pretendem comprar no Natal este ano e pegar uma prancheta rumo ao shopping um dia depois do feriado de Ação de Graças para perguntar às pessoas sobres seus planos, você tem um viés em seu processo de amostragem. Sua amostra tende a privilegiar aqueles compradores que estavam se gladiando no meio da multidão naquele dia e naquele shopping em particular.

- **Da maneira com os dados são coletados:** as perguntas de pesquisas são a principal fonte de viés. Devido ao fato de que os pesquisadores estão sempre procurando determinados resultados, as perguntas elaboradas por eles podem com frequência refletir o resultado esperado. Por exemplo, a questão da arrecadação de impostos para ajudar as escolas locais é algo que todo eleitor encara uma vez ou outra. Uma enquete com a pergunta: "Você não acha que a ajuda às escolas locais seria um grande investimento em nosso futuro?" Realmente tem um pouco de viés. Por outro lado, a pergunta: "Você não está cansado de pagar mais para educar os filhos dos outros além dos seus?"; também apresenta viés. A maneira como as perguntas são escritas pode ter um impacto enorme sobre os resultados. Veja o capítulo 16 para mais informações sobre o planejamento de enquetes e pesquisas de opinião.

Ao examinar o resultado de pesquisas que lhe interessem ou que sejam importantes para você, descubra quais perguntas foram feitas e como exatamente elas foram formuladas antes de tirar suas conclusões sobre os resultados.

Dados

Os dados são a medida real com a qual você vai lidar ao longo de seu estudo. A maioria dos dados cai em dois tipos de grupos: dados numéricos ou dados categorizados (veja o Capítulo 5 para informação extra).

- Dados numéricos são os dados que têm importância como medida, tais como a altura, o peso, o QI ou a pressão arterial de uma pessoa; o número de estoque que uma pessoa possui, o número de dentes que o cachorro de alguém tem; ou qualquer outra coisa que possa ser contada (os estatísticos também se referem aos dados numéricos como dados quantitativos ou dados de medida).

- Dados categorizados representam as características, tais como o sexo de uma pessoa, a opinião, a raça e até mesmo o tipo do umbigo (para dentro ou para fora – nada mais é considerado sagrado?) Enquanto essas características podem ser representadas por número (exemplo, "1" indica masculino e "2" indica feminino), tais números não têm significado especifico algum. Você não poderia os somar, por exemplo (observe que os estatísticos também se referem a esse tipo de dado como dado qualitativo).

Nem todos os dados são criados iguais. Descobrir como os dados foram coletados é dar o primeiro passo para determinar como você pesa os resultados e que conclusões podem ser tiradas a partir deles.

Conjunto de dados

Um conjunto de dados é a coleção de todos os dados tirados a partir de uma amostra. Por exemplo, se você mediu o peso de cinco embalagens e os pesos foram 12 lbs (5,4kg), 15 lbs (6,8 kg), 22 lbs (9,9 kg), 68 lbs (30,8 kg) e 3 lbs (1,3 kg), esses números constituem seu conjunto de dados. Entretanto, a maioria dos conjuntos de dados é maior do que esse.

Estatística

A estatística é um número que resume os dados coletados a partir da amostra. Muitos tipos diferentes de estatística são utilizados para resumir os dados. Por exemplo, os dados podem ser resumidos como porcentagens (60% das famílias estudadas nos Estados Unidos possuem mais de dois carros), como médias (o preço médio de uma casa nesta amostra é de...), como uma mediana (o salário mediano para os 1.000 cientistas da computação desta amostra foi de...) ou como um percentil (o peso de seu bebê está no percentil 90°, baseando-se nos dados coletados de mais de 10.000 bebês...).

Nem todas as estatísticas são corretas ou honestas, é claro. Não é porque alguém lhe mostra uma estatística que você precisa aceitá-la como sendo científica ou legítima! Você já deve ter ouvido alguém dizer: "Os números não mentem, mas os mentirosos fabricam os números."

As estatísticas baseiam-se em dados de amostra, e não em dados populacionais. Quando se coletam dados de toda uma população, temos o chamado censo. Se você depois resume toda a informação do censo em um número, esse número é um parâmetro, não uma estatística. Na maioria das vezes, os pesquisadores tentam estimar os parâmetros, usando a estatística. O U.S. Census Bureau é uma agência que relata o número total da população dos Estados Unidos, portanto eles conduzem um censo. Mas devido a problemas logísticos relacionados a essa árdua tarefa (tais como o de entrar em contato com os moradores de rua), os números do censo, no final das contas, podem apenas ser chamados de estimativas e eles acabam sendo ajustados para cima a fim de contabilizar as pessoas que o censo não alcançou. O longo formulário do censo é preenchido por uma amostra aleatória de residências; O U.S Census Bureau utiliza essa informação para esboçar conclusões sobre toda a população (sem ter que pedir a todos que preencham aquele enorme formulário).

Média

A média é a estatística mais comum utilizada para medir o centro ou o meio de um conjunto de dados numéricos. A média é a soma de todos os números divido pelo número total de números. Veja o Capítulo 5 para mais informações sobre a média.

A média pode não ser uma representação honesta dos dados, pois é facilmente influenciada pelos valores discrepantes (valores muito grandes ou muito pequenos dentro do conjunto de dados).

Mediana

A mediana é outra maneira de medir o centro de um conjunto de dados numéricos (além da boa e velha companheira, a média). A mediana estatística é muito parecida com a mediana de uma rodovia. Em uma rodovia, a mediana determina o meio da pista, separando um número igual de faixas para os dois lados da mediana. Em um conjunto de dados numéricos, a mediana é o ponto em que existe um número igual de pontos de dados cujos valores ficam tanto acima quanto abaixo do valor da mediana. Assim, a mediana é verdadeiramente o meio do conjunto de dados. Veja o Capítulo 5, para saber mais sobre a mediana.

Da próxima vez em que você ouvir o relato de uma média, observe se a mediana também foi falada. Caso não tenha sido, pergunte por ela! A média e a mediana são duas representações diferentes de um conjunto de dados e, geralmente, podem contar duas versões diferentes sobre os mesmos dados.

Desvio padrão

Você já ouviu alguém dizer que foi encontrado um determinado resultado "com dois desvios padrões acima da média"? Cada vez mais, as pessoas querem relatar o quanto seus resultados são importantes e o número de desvios padrões acima ou abaixo da média é uma maneira de fazer isso. Mas, o que é um desvio padrão afinal?

O desvio padrão é um modo que os estatísticos usam para medir a variabilidade entre os números em um conjunto de dados. Assim como o termo sugere, um desvio padrão é um padrão (ou seja, algo típico) de desvio (ou distância) da média. Portanto, o desvio padrão, em termos bem simples, é a distância média da média. Veja o Capítulo 5, para cálculos e mais informações.

O desvio padrão também é utilizado para descrever onde a maioria dos dados deveria se encaixar, em um sentido relativo, comparado à média. Por exemplo, em muitos casos, cerca de 95% dos dados ficarão dentro de dois desvios padrões da média (esse resultado é chamado de regra empírica. Veja o Capítulo 8 para mais sobre o assunto).

A fórmula para os desvios padrões é a seguinte:

$s = \sqrt{\dfrac{\sum (x - \bar{x})^2}{n - 1}}$, onde

= o número de valores do conjunto de dados

$\bar{x}$ = a média de todos os valores

x = cada valor do conjunto de dados

Para instruções mais detalhadas sobre o cálculo do desvio padrão, veja o Capítulo 5.

O desvio padrão é uma estatística importante, mas, frequentemente, é omitida quando os resultados são relatados. Sem ele, você está recebendo apenas uma parte da história sobre os dados. Os estatísticos gostam de contar a história do homem que estava com um dos pés em um balde de água gelada e o outro em um balde de água fervendo. O homem dizia que, na média, ele estava se sentindo ótimo! Mas imagine a variabilidade da temperatura para cada um dos pés. Agora, colocando os pés no chão, o preço médio de uma casa, por exemplo, não lhe diz nada sobre a variedade de preços de casas com a qual você pode se deparar enquanto estiver procurando uma casa para comprar. A média dos salários pode não representar verdadeiramente o que realmente está se passando em sua empresa se os salários forem extremamente discrepantes.

Não se satisfaça em saber apenas a média – pergunte também sobre o desvio padrão. Sem o desvio padrão, não há como saber a discrepância entre os valores (se você estiver falando sobre salário inicial, isso pode ser muito importante!).

Percentil

Provavelmente você já ouvido falar sobre percentil antes. Se você já fez algum tipo de teste padronizado, você sabe que quando sua pontuação é mostrada, ela lhe é apresentada como uma medida da posição em que você se enquadrou, comparado com as outras pessoas que também fizeram o teste. Essa medida comparativa geralmente é apresentada em termos de um percentil. O percentil apresentado para uma dada pontuação é a porcentagem dos valores do conjunto de dados que se encontram abaixo de certa pontuação. Por exemplo, se sua pontuação está no percentil 90°, isso significa que 90% das pessoas que fizeram a prova junto com você fizeram menos pontos do que você (e 10% tiveram uma pontuação mais alta do que a sua). Para mais informações sobre percentil, veja o Capítulo 5.

O percentil é utilizado de várias formas, sempre visando à comparação e a determinação da posição relativa (ou seja, como o valor de um dado individual é comparado ao resto do grupo). Por exemplo, o peso de bebês, que geralmente é demonstrado por meio de um percentil. O percentil também é utilizado por empresas, a fim de determinar sua posição quando comparadas com outras empresas em termos de vendas, lucro, satisfação do cliente e etc.

Escore padrão

O escore padrão (standard score) é uma maneira habilidosa de expressar os resultados em perspectiva sem ter que fornecer muitos detalhes – algo que a mídia adora. O escore padrão representa o número de desvios padrões acima ou abaixo da média (sem se importar em saber quais são os reais valores do desvio padrão e da média).

Para exemplificar, suponha que Bob tenha feito 400 pontos em uma prova. Mas o que isso significa? Pode não significar muito, pois não é possível colocar essa pontuação dentre uma perspectiva. Mas se você soubesse que o escore padrão do Bob no teste é de +2, saberia que a pontuação de Bob é de dois desvios padrão acima da média (parabéns, Bob!). Agora, suponha que o escore padrão de Bill seja de -2. Nesse caso, isso não é bom (para o Bill), pois significa que sua pontuação é de dois desvios padrões abaixo da média.

A fórmula para o escore padrão é:

$$\text{Escore padrão} = \frac{\text{escore original} - \bar{x}}{s}$$

Onde:

$\bar{x}$ é a média de todos os escores

s é o desvio padrão de todos

Para saber detalhes sobre o cálculo e a interpretação dos escores padrão, veja o Capítulo 8.

Distribuição normal (curva em forma de sino)

Quando os dados numéricos são organizados, eles geralmente são ordenados do menor para o maior, divididos em grupos de tamanho razoável e, depois, são colocados em gráficos para que se examine sua forma, ou distribuição. O tipo mais comum de distribuição de dados é a chamada curva em forma de sino, na qual a maioria dos dados é centralizada ao redor da média, de modo que quanto mais longe da média você se mover, cada vez menos pontos de dados você vai encontrar em ambos os lados. A Figura 3-1 mostra a ilustração de uma curva em forma de sino; observe que a forma da curva lembra o contorno de um sino antigo.

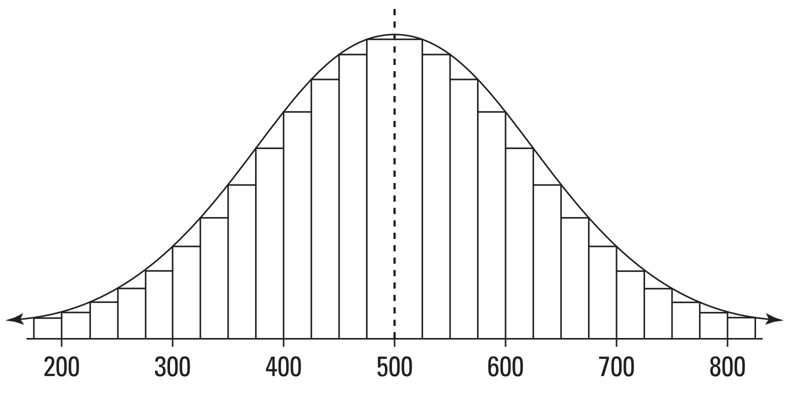


Figura 3.1: Curva em forma de sino

Os estatísticos têm outro nome para a curva em forma de sino para quando há muitos valores possíveis para os dados; eles a chamam de distribuição normal. Essa distribuição é utilizada para descrever os dados que seguem o padrão em forma de sino, incluindo a variedade de valores esperada e a posição ocupada por um escore individual em relação ao demais. Por exemplo, se os dados têm uma distribuição normal, espera-se que a maioria deles fique dentro dos dois desvios padrões da média, pois toda população distinta de dados tem uma média e uma desvio padrão diferente. Há um número infinito de diferentes distribuições normais, cada uma com sua própria média e com seu próprio desvio padrão para caracterizá-la. Veja o Capítulo 8 para muito mais informações sobre a distribuição normal.

PAPO TÉCNICO

A distribuição normal também é utilizada para ajudar a medir a precisão de muitas estatísticas, incluindo a média, por meio da utilização de um importante resultado em estatística, conhecido como teorema do limite central. Esse teorema possibilita medir o quanto sua média amostral irá variar, sem ter que pegar outra média amostral para fazer a comparação (felizmente!). Esse teorema basicamente diz que sua média amostral tem uma distribuição normal, independente da aparência da distribuição dos dados originais (desde que sua amostra seja grande o bastante). Veja o Capítulo 9 para saber mais sobre o teorema do limite central (conhecido entre os estatísticos como "a joia da coroa de toda a estatística". Você se importaria de dizer a eles para aproveitar mais a vida?).

FIQUE ATENTO

Se um conjunto de dados tem uma distribuição normal e você padroniza todos os dados para obter escores padrões, estes escores padrões são chamados de escore Z. Os escores Z possuem o que se convencionou chamar de distribuição normal padrão (ou distribuição Z). A distribuição normal padrão é uma distribuição normal especial com a média igual a 0 e o desvio padrão igual a 1. A distribuição normal padrão é útil para examinar os dados e determinar estatística como o percentil ou a porcentagem de dados entre dois valores. Assim, se os pesquisadores determinam que os dados têm uma distribuição normal, eles irão primeiramente padronizar os dados (convertendo cada ponto de dado em um escore Z) e, depois, utilizam a distribuição normal padrão para investigar e discutir os dados mais detalhadamente.

Experimentos

Um experimento é um estudo que impõe certo controle sobre os sujeitos do estudo e seu ambiente (por exemplo, restringindo sua alimentação, dando-lhes certas doses de medicamentos ou placebos ou lhes pedindo para que permaneça acordado por um determinado período). O objetivo da maioria dos experimentos é identificar uma relação de causa e efeito entre duas variáveis (como o consumo de álcool e problemas de visão). Eis algumas perguntas que um experimento tenta responder:

- Tomar zinco ajuda a reduzir a duração de um resfriado? Alguns estudos afirmam que sim.
- A forma e a posição de seu travesseiro afeta a qualidade de seu sono? O Emory Spine Center em Atlanta diz que "sim".
- Sapatos de salto alto afetam o conforto dos pés? Estudos realizados pelo UCLA dizem que é melhor um sapato com salto de 2,5 cm do que sapatos sem salto algum.

Nesta seção, você encontrará mais informações sobre como são (ou deveriam ser) conduzidos os estudos experimentais. E o Capítulo 17 é inteiramente dedicado a esse assunto. Mas, agora, apenas se concentre nos termos básicos relacionados aos experimentos.

Grupo experimental versus grupo de controle

Muitos experimentos tentam determinar se um tipo de tratamento (ou fator importante) tem algum tipo de efeito sobre um resultado. Por exemplo, o zinco ajuda a reduzir a duração de um resfriado? Os indivíduos que são convidados a participar dos experimentos são normalmente divididos em dois grupos, o grupo experimental e o grupo de controle. O grupo experimental é composto por indivíduos que receberão o tratamento que, supostamente, tenha um efeito sobre o resultado (neste caso, o zinco). O grupo de controle é composto pelos indivíduos que receberão o tratamento padrão conhecido, cujos resultados serão comparados ao do novo tratamento (como a vitamina C, no caso do estudo do zinco).

Placebo

O placebo é um tratamento falso, como um comprimido de açúcar. Geralmente é dado aos indivíduos do grupo de controle, para que, desta forma, eles não saibam se estão recebendo o tratamento (por exemplo, o zinco) ou se não estão recebendo nada. Os placebos são administrados ao grupo de controle devido a um fenômeno chamado efeito placebo, no qual os pacientes que receberam qualquer tipo de tratamento (ainda que seja um comprimido de açúcar) relataram algum tipo de resultado, seja positivo ("Sim, já estou me sentindo melhor") ou negativo ("Nossa, estou me sentindo um pouco tonto"), causado por um efeito psicológico. Sem o placebo, os pesquisadores não poderiam ter certeza de que os resultados foram obtidos a partir do tratamento ou a partir do efeito placebo.

Estudo cego e duplo-cego

Em um estudo cego, os indivíduos que fazem parte do estudo não sabem se estão no grupo experimental ou no grupo de controle. No exemplo do zinco, o placebo utilizado deveria se parecer com a pílula de zinco e os pacientes não devem saber qual tipo de pílula estão tomando. O estudo cego tenta eliminar qualquer tipo de parcialidade dos indivíduos estudados.

O estudo duplo-cego controla a parcialidade potencial por parte dos pacientes e também dos pesquisadores. Nem os pacientes, nem os pesquisadores que estão coletando os dados, sabem quais indivíduos receberam o tratamento e quais não o receberam. O estudo duplo-cego é melhor, pois, ainda que os pesquisadores se digam imparciais, sempre há o interesse por um determinado resultado – se não fosse assim, o estudo não estaria sendo realizado.

Pesquisas de opinião (enquetes)

Uma pesquisa de opinião (também chamada de enquete) é uma ferramenta de medição frequentemente utilizada para reunir a opinião da população juntamente com algumas informações demográficas relevantes. Devido ao fato de que muitos formadores de opinião, profissionais de marketing e outros querem "entender a alma do público americano" e descobrir o que um cidadão comum pensa e sente, muitas pessoas atualmente acreditam ser impossível escapar da avalanche de pedidos para que participem de enquetes e pesquisas de opinião. Realmente, você já deve ter recebido muitos pedidos para fazer parte de pesquisas e, talvez, até tenha se sentido entorpecido por elas. Simplesmente, jogue fora as pesquisas recebidas junto com sua correspondência ou diga "não" quando receber um convite pelo telefone.

Se feita de maneira adequada, uma pesquisa de opinião pode realmente ser informativa. Tais pesquisas são utilizadas para descobrir que programas de TV os americanos (e outros) gostam, como os consumidores se sentem em relação a compras pela Internet e se os Estados Unidos devem ter um sistema de defesa nuclear. As pesquisas também são utilizadas pelas empresas para medir o nível de satisfação de seus clientes, descobrir que produtos eles querem e determinar quem irá comprar seus produtos. As emissoras de TV utilizam as pesquisas para obter reações imediatas aos novos eventos e fatos e os cineastas as utilizam para determinar como terminar seus filmes.

Se eu tivesse que escolher uma palavra para descrever o estado geral das pesquisas na mídia atualmente, teria que usar a palavra quantidade ao invés de qualidade. Ou seja, você não encontra escassez de pesquisas ruins. Para determinar se uma pesquisa foi conduzida de maneira adequada, basta fazer algumas poucas perguntas; mais detalhes desse assunto no Capítulo 16.

Estimação

Uma das maiores utilidades da estatística é chutar um valor (o termo estatístico é estimação), como nos exemplos a seguir:

- Qual é a renda média de uma família americana?
- Qual a porcentagem de casas assistindo o Oscar esse ano?
- Qual a expectativa de vida média de um bebê nascido hoje?
- Qual a eficácia deste novo remédio?
- Qual a qualidade do ar de hoje comparado com o ar de dez anos atrás?

Todas essas perguntas necessitam de algum tipo de estimativa numérica para respondê-las, ainda que haja a necessidade de se conseguir uma estimativa correta e precisa. As seções a seguir estudam os principais elementos desse processo. Para mais informações sobre como fazer e interpretar uma estimativa, veja o Capítulo 11.

Margem de erro

Você provavelmente já ouviu alguém dizer: "Esta pesquisa tem uma margem de erro de 3 pontos percentuais para mais ou para menos". Mas o que isso significa? Todas as pesquisas baseiam-se em informações coletadas a partir de uma amostra de indivíduos, e não de toda a população. Certos erros podem ocorrer – não no sentido de erros de cálculos (embora, também possa haver alguns erros deste tipo), mas no sentindo de erro amostral, ou seja, um erro causado simplesmente porque os pesquisadores não irão questionar todo mundo. A margem de erro deve medir a quantidade máxima na qual se espera que os resultados da amostra se diferenciem dos resultados da população total. Devido ao fato de que os resultados da maioria das pesquisas são relatados em termos de porcentagem, a margem de erro também aparece, na maioria das vezes, como porcentagem.

Como interpretar uma margem de erro? Suponha que você saiba que 51% dos indivíduos de uma amostra disseram que pretendem votar em um determinado candidato na próxima eleição. Agora, para projetar esse resultado para o número total de eleitores, você teria que somar e subtrair a margem de erro e dar uma variação de possíveis resultados, a fim de garantir que você esteja cobrindo a lacuna entre sua amostra e a população total. Então, neste caso (supondo uma margem de erro de 3 pontos para mais ou para menos), você pode garantir, baseando-se nos resultados da amostra, que de 48% a 54% das pessoas irão votar para o candidato em questão. Neste caso, o candidato em questão poderia obter um pouco a menos ou um pouco a mais do que a maioria dos votos e tanto poderia perder como ganhar a eleição. Essa tem se tornado uma situação muito comum nos últimos anos, quando não é possível para a mídia noticiar os resultados no dia da eleição apenas se baseando nos resultados da pesquisa. Para mais informação sobre margem de erro, veja o Capítulo 10.

A margem de erro mede a precisão e não a presença de uma possível parcialidade. Resultados que pareçam numericamente científicos e precisos não representam nada se coletados de maneira tendenciosa.

Intervalo de Confiança

Quando você combina sua estimativa com a margem de erro, você tem um intervalo de confiança. Por exemplo, suponha que o tempo médio que você leva para chegar ao trabalho de carro é de 35 minutos, com uma margem de erro de 5 minutos para mais ou para menos. A sua estimativa é de que o tempo médio gasto até chegar ao trabalho fica em algum ponto entre 30 a 40 minutos. Esta estimativa é um intervalo de confiança, pois leva em consideração o fato de que os resultados da amostra irão variar e dá uma indicação da variação esperada. Para mais sobre os fundamentos do intervalo de confiança, veja o Capítulo 11.

Alguns intervalos de confiança são mais amplos do que outros (e quanto mais amplo, menor sua precisão). Vários fatores influenciam a amplitude de um intervalo de confiança, tais como o tamanho amostral, a variabilidade da população estudada e o quanto você espera obter de precisão (a maioria dos pesquisadores contenta-se com 95% de nível de confiança em seus resultados). Leia mais sobre os fatores que influenciam os intervalos de confiança no Capítulo 12.

Muitos tipos diferentes de intervalos de confiança são feitos nas pesquisas científicas, incluindo intervalos de confiança para médias, proporções, a diferença de duas médias ou proporções ou de diferenças emparelhadas. Para mais detalhes sobre os testes de hipótese mais comuns, veja o Capítulo 13.

Probabilidade versus Possibilidades

Uma probabilidade é a medida da possibilidade de um evento ocorrer. Em outras palavras, se a chance de chover amanhã é de 30%, há menos possibilidades de que chova amanhã do que de que chova, mas as chances de chuva é, ainda, é de 3 em 10 (dada essa possibilidade, você trará seu guarda-chuva amanhã?). Uma possibilidade de chuva de 30% também significa que durante muito e muitos dias com as mesmas condições climáticas como as previstas para amanhã, houve chuva em 30% das vezes.

Probabilidades são calculadas de muitas maneiras diferentes:

- A matemática é utilizada para fabricar os números (por exemplo, calculando suas chances de ganhar na loteria ou determinando a hierarquia das mãos no poker).
- Os dados são coletados e as probabilidades são estimadas com base no histórico dos dados (por exemplo, para a previsão do tempo).
- Modelos matemáticos e computacionais são usados na tentativa de prever o comportamento e a ocorrência de um fenômeno natural (por exemplo, furacões e terremotos).

As leis da probabilidade, com frequência, vão contra sua intuição e suas crenças a respeito do que pode acontecer (por isso existem os cassinos). Leia mais sobre probabilidade no Capítulo 6.

Possibilidades e probabilidades são ligeiramente diferentes. A melhor maneira de descrever essa diferença é estudando o seguinte exemplo: suponha que a probabilidade de que determinado cavalo vença uma corrida é de 1 em 10. Isso significa que sua probabilidade de vencer é 1 em 10 ou 1÷ 10 ou 0,10. A probabilidade reflete as chances de vencer. Agora, quais são as possibilidades de que o cavalo vença? São de 9 em 1. Isso porque, na verdade, a possibilidade é a razão entre as chances de perder e as chances de ganhar. Divida 9/10 por 1/10, cancelamos os 10s e temos 9/1, que na, linguagem das possibilidades, é lido como “9 para 1”. Leia mais sobre jogos de azar no Capítulo 7.

A lei das médias

Você provavelmente já ouviu alguém falar da lei das médias antes. Talvez tenha sido o comentarista de basebol local, lamentando-se que seu time, que havia desafiado as possibilidades ao vencer 50 jogos e perder apenas 12 nos primeiros três meses da temporada e que, agora, começa a perder, dada a lei das médias. Ou, talvez, em uma mesa de jogo ("A lei da média vai me pegar a qualquer momento – estou com muita sorte hoje!). O que é a lei das médias, exatamente, e, as pessoas estão usando esse termo de maneira adequada?

A lei das médias é uma regra de probabilidade. Ela diz que, com o tempo, os resultados irão fazer uma média do valor esperado, mas a curto prazo, ninguém sabe o que irá acontecer. Por exemplo, os cassinos ajustam todos seus jogos para que as chances de a casa vencer sejam ligeiramente maiores. Isso significa que, com o tempo, desde que as pessoas continuem jogando, na média, os cassinos certamente terão lucro. É claro que haverá ganhadores, é o que mantém as pessoas jogando e acreditando que elas poderão ser os próximos vencedores. Mas, a longo prazo, o número de perdedores sobrepõe-se ao de ganhadores (sem contar o fato de que muitas vezes as pessoas que ganham uma grande quantia, apostam seu prêmio e acabam ficando sem nada). Leia mais sobre a lei das médias no Capítulo 7.

Teste de Hipóteses

Teste de hipóteses é um termo com o qual provavelmente você ainda não se deparou quando lidando com números e estatísticas no seu cotidiano. Mas eu garanto que os testes de hipótese têm sido parte de sua vida pessoal e profissional, simplesmente devido ao importante papel que eles desempenham na indústria, na medicina, na agricultura, no governo e em muitas outras áreas. Todas as vezes em que você ouve alguém falando que seus resultados mostram "uma significativa diferença estatística", você está diante dos resultados de um teste de hipótese. O teste de hipótese é, basicamente, um procedimento estatístico em que os dados são coletados e medidos para comprovar uma alegação feita sobre uma população.

Por exemplo, se uma cadeia de pizzaria delivery alega entregar as pizzas dentro de 30 minutos a partir do pedido, você pode testar se essa alegação é verdadeira, coletando uma amostra aleatória do tempo de entrega durante um determinado período de tempo e observar o tempo médio de entrega para essa amostra.

Pelo fato de que sua decisão baseia-se em uma amostra, e não em uma população inteira, um teste de hipótese pode, às vezes, levar-lhe a tirar conclusões equivocadas. Entretanto, a estatística é tudo o que você tem e, se feita da maneira adequada, ela pode chegar o mais próximo possível da verdade sem realmente conhecer a verdade. Para mais sobre os fundamentos do teste de hipótese, veja o Capítulo 14.

Uma variedade de testes de hipótese é realizada em uma pesquisa científica, incluindo os testes-t, os testes-t emparelhados, testes de proporção ou médias para uma ou mais populações. Para mais detalhes sobres os testes de hipótese mais comuns, veja o Capítulo 15.

P-valor

Os testes de hipóteses são realizados para confirmar ou negar uma alegação feita sobre uma população. A alegação que está sendo testada é, em essência, denominada hipótese nula. A evidência em teste são seus dados e a estatística que os acompanha. Todos os testes de hipótese usam um ρ-valor para pesar a força da evidência (o que os dados estão lhe dizendo sobre a população). O ρ-valor é um número entre 0 e 1 que reflete a força dos dados que estão sendo usados para avaliar a hipótese nula. Um ρ-valor grande indica uma evidência fraca contra a hipótese nula. Por exemplo, se uma rede de pizzaria alega entregar pizzas em menos de 30 minutos (essa é a hipótese nula) e a sua amostra aleatória de 100 entregas tem uma média de 40 minutos para o tempo de entrega (mais do que 2 desvios padrões acima do que o tempo médio de entrega deveria ser) o ρ-valor para esse teste seria pequeno, e você diria que tem uma forte evidência contra a alegação feita pela rede de pizzaria.

Estatisticamente significativo

Sempre que se coletam dados para a realização de um teste de hipótese, os pesquisadores normalmente estão atrás de um resultado significativo. Geralmente, isso significa que o pesquisador encontrou algo fora do comum (pesquisa que, simplesmente, confirma o que já se sabe não dá manchete, infelizmente). Um resultado estatisticamente significativo é aquele que deveria ter tido uma probabilidade muito pequena de acontecer por acaso. O ρ-valor reflete essa probabilidade.

Por exemplo, se uma pesquisa descobre que um medicamento mostra-se mais eficaz no tratamento de câncer de mama do que o tratamento atual, os pesquisadores dizem que a nova droga mostra uma melhora estatisticamente significativa na taxa de sobrevivência das pacientes com câncer de mama (ou em algum outro aspecto que mostre esse efeito). Isso significa que baseados em seus dados, a diferença dos resultados entre as pacientes que utilizaram o novo medicamento com aquelas que utilizaram o tratamento antigo é tão grande que seria difícil dizer que foi apenas uma coincidência.

Às vezes, uma amostra não representa a população (por casualidade), e isso resulta em uma conclusão equivocada. Por exemplo, um efeito positivo experimentado por uma amostra de pessoas que tomaram o novo medicamento pode ter sido apenas um feliz acaso (presuma, para o momento, que você sabe que os dados não foram fabricados, forjados ou exagerados). A beleza da pesquisa médica está no fato de que assim que alguém lança uma comunicação de impressão dizendo que ele ou ela encontrou algo significativo, começa a corrida para tentar replicar os resultados e, se os resultados não puderem ser repetidos, isso, provavelmente, significa que os resultados originais estavam errados por alguma razão. Infelizmente, um comunicado de imprensa anunciando uma “nova descoberta” tende a render muito na mídia, por outro lado, os estudos de acompanhamento que o refutam dificilmente aparecem na primeira página.

LEMBRE-SE

Um resultado estatisticamente significativo não deve levar ninguém a tirar conclusões rápidas. Na ciência, o que conta não é apenas um único estudo notável, mas um grupo de evidências construído com o tempo juntamente com uma variedade de estudos de acompanhamento bem projetados. Considere qualquer nova descoberta que você ouvir como sendo um grão de areia e espere até que o trabalho de acompanhamento tenha sido realizado antes de usar a informação de um único estudo para tomar decisões importantes na sua vida.

Correlação e Causalidade

Entre todos os equívocos das questões estatísticas, o mais problemático é o abuso dos conceitos de correlação e causalidade.

Correlação significa que duas variáveis numéricas possuem algum tipo de relação linear. Por exemplo, o número de vezes que os grilos crilam por segundo está relacionado à temperatura; quando está frio, eles crilam com menos frequência e quando está calor, eles crilam mais vezes (isso realmente acontece!). Outro exemplo de correlação tem a ver com a alocação de policiais. O número de crimes (per capita) tem sido sempre relacionado ao número de policiais em uma dada área. Quando mais policiais patrulham uma área, a criminalidade tende a ser mais baixa e, quando menos policiais estão presentes, a criminalidade tende a ser mais alta. Entretanto, aparentemente, eventos sem relação alguma mostram-se correlacionados. Um desses exemplos é o consumo de sorvete (casquinha por pessoa) e o número de assassinatos em certas áreas. Agora, talvez um efetivo maior de policiais detenha a criminalidade, mas será que se as pessoas diminuírem seu consumo de sorvete, a criminalidade será detida? Qual a diferença? A diferença é que com a correlação, se pode descobrir a existência de uma ligação ou relação entre duas variáveis, x e y. Com a causalidade, alguém aparece e diz: "uma mudança em x irá causar uma mudança em y". Muitas vezes na pesquisa, na mídia ou no consumo público de estatística, isso é feito quando não deveria. Mas quando isso pode ser feito? Quando um experimento muito bem planejado é executado, eliminando quaisquer outros fatores que poderia estar relacionado com os resultados. Para mais sobre correlação e causalidade, veja o Capítulo 18.

Parte II

Fundamentos de Cálculos Numéricos

A 5ª Onda por Rich Tennant

"Prepare-se, acho que eles estão começando a dormir."

Nesta Parte...

Cálculos numéricos: alguém tem que fazer o trabalho sujo. Por que não você? Mesmo que você não seja uma pessoa de números e, definitivamente, calcular não é sua praia, a abordagem passo a passo nesta parte pode ser justamente o que você precisava para estimular sua confiança para fazer, e realmente entender, a estatística.

Nesta parte, você entrará em contato com os fundamentos dos cálculos numéricos, fazendo e interpretando gráficos e tabelas para fazer e entender médias, medianas, desvios padrões e mais. Você também irá desenvolver habilidades para avaliar a informação estatística de outra pessoa e desvendar a verdade por trás dos dados.

Capítulo 4

Interpretando Gráficos e Tabelas

Neste Capítulo

- Investigando como os dados com os quais você irá se deparar são exibidos
- Reconhecendo o objetivo escondido na imagem
- Interpretando e avaliando gráficos e tabelas
- Tomando cuidados com os gráficos enganosos

Uma vez alguém disse que uma imagem vale mais do que mil palavras. Em estatística, uma imagem pode valer mais do que milhares de pontos de dados – desde que essa imagem seja feita corretamente, claro. Com muita frequência, os dados nos são mostrados por meio de gráficos e tabelas. Essas figuras nos mostram tudo, desde os resultados da eleição, levando-se em consideração todas as possibilidades, até o desempenho da bolsa de valores nos últimos anos. A sociedade de hoje é a sociedade do fast-food e da fast-information; todos querem saber tudo, sem se importar com detalhes. O principal uso da estatística é o de condensar a informação e resumi-la, e mostrar os dados em forma de gráficos e tabelas é a maneira mais natural de se fazer isso. Mas, essas figuras realmente lhe mostram o que realmente está acontecendo com os dados? Isso depende da qualidade da imagem e seu propósito. As aparências podem enganar (às vezes propositadamente, às vezes sem querer) e nem todo gráfico e tabela que você vê estará correto. Este capítulo lhe ajudará a ganhar mais entendimento do uso de gráficos e tabelas pela mídia e pelo lugar onde você trabalha e lhe mostrará como ler e encontrar um sentido para essas imagens. Neste capítulo, eu também vou dar algumas dicas de como avaliar gráficos e tabelas e identificar os enganosos.

Unindo Gráfico com Estatística

O principal objetivo de gráficos e tabelas é o de estabelecer um ponto, deixá-lo claro e correto. Essas ferramentas são utilizadas, por exemplo, para dar um impacto a uma característica especifica dos dados, destacar mudanças ao longo do tempo, comparar e contrastar opiniões ou dados demográficos ou demonstrar ligações entre informações. Gráficos e tabelas quebram a história estatística que o autor quer transmitir a respeito de um conjunto de dados, para que, assim, o leitor possa rapidamente ter uma ideia do assunto e chegar a uma conclusão. Por esta razão, essas ferramentas são poderosas: usadas de maneira apropriada, elas podem ser informativas e eficazes; usadas de maneira inadequada, podem ser enganosas e destrutivas.

Gráficos e tabelas podem causar grandes ou pequenos impactos em sua vida – reagir de maneira crítica e entender o que eles querem, ou não, dizer, lhe ajudará a se tornar um consumidor de estatísticas atento. Você deve se familiarizar com as várias maneiras de como os dados são exibidos, pois você provavelmente irá se deparar com vários deles e, sendo assim, aprenderá a investigar a maneira como essas ferramentas são utilizadas pela mídia e pelo local onde você trabalha.

Pesquisadores e jornalistas utilizam diferentes maneiras de mostrar cada um dos dois tipos mais importantes de dados: dados categorizados, que representam qualidades ou características (tais como sexo ou partido político) e dados numéricos, que representam as quantidades (tais como altura e renda).

As ferramentas mais utilizadas para a exibição de dados categóricos são as seguintes:

- Gráfico Pizza (veja a seção “Digerindo as fatias de um Gráfico Pizza”)
- Gráfico de barras (veja a seção “Levantando as barras de um Gráfico de Barras”)
- Tabelas (veja a seção “Estatística de tabela”)
- Gráfico de linhas (veja a seção “Alinhado ao Gráfico de Linhas)

Para os dados numéricos, as tabelas são comumente utilizadas. Além disso, também deveria ser muito comum se utilizar histogramas para a exibição desse tipo de dados (mas não é), por isso eu os inclui na seção “Retratando os dados em um Histograma”.

Neste capítulo, apresento exemplos de cada uma das ferramentas para exibição dos dados, algumas considerações e dicas para avaliar cada uma delas de maneira mais crítica.

Digerindo as Fatias de um Gráfico Pizza

O gráfico pizza é um dos mais utilizados, pois é fácil de ser lido e consegue rapidamente estabelecer um objetivo. Você provavelmente já viu um – ele parece tão simples. O que pode ter de errado com um inocente gráfico pizza?

O gráfico pizza considera os dados categorizados e os divide em grupos, mostrando a porcentagem de indivíduos que se enquadram em cada grupo. Devido à forma circular do gráfico pizza, as fatias que representam cada grupo podem ser facilmente comparadas e contrastadas entre si. Devido ao fato de que cada indivíduo no grupo enquadra-se em apenas uma categoria, a soma de todas as fatias da pizza deveria ser 100% ou próximo a isso (estando sujeito a um erro de arredondamento).

Computando seus gastos pessoais

No que você gasta seu dinheiro? Quais são suas três maiores despesas? De acordo com o U.S. Bureau de Estatística do Trabalho, as três maiores despesas dos americanos em 1994 eram os gastos domésticos (32%), transporte (19%) e alimentação (14%). A Figura 4-1 mostra esses resultados em um gráfico pizza (observe que a categoria "outras" está um pouco grande neste gráfico. Mas neste caso, seria difícil determinar que outros itens deveriam ser incluídos em uma única fatia de um gráfico pizza, uma vez que existem muitas possibilidades de tipos de gastos para pessoas diferentes).

Como o governo americano conseguiu essa informação? Da chamada Consumer Expenditure Survey (Pesquisa dos Gastos dos Consumidores). Muitas agências federais ficam responsáveis por coletar dados (frequentemente por meio de pesquisas) e por espalhar os resultados por meio da imprensa escrita (o governo americano é uma ótima fonte sobre muitos aspectos da vida cotidiana nos Estados Unidos).

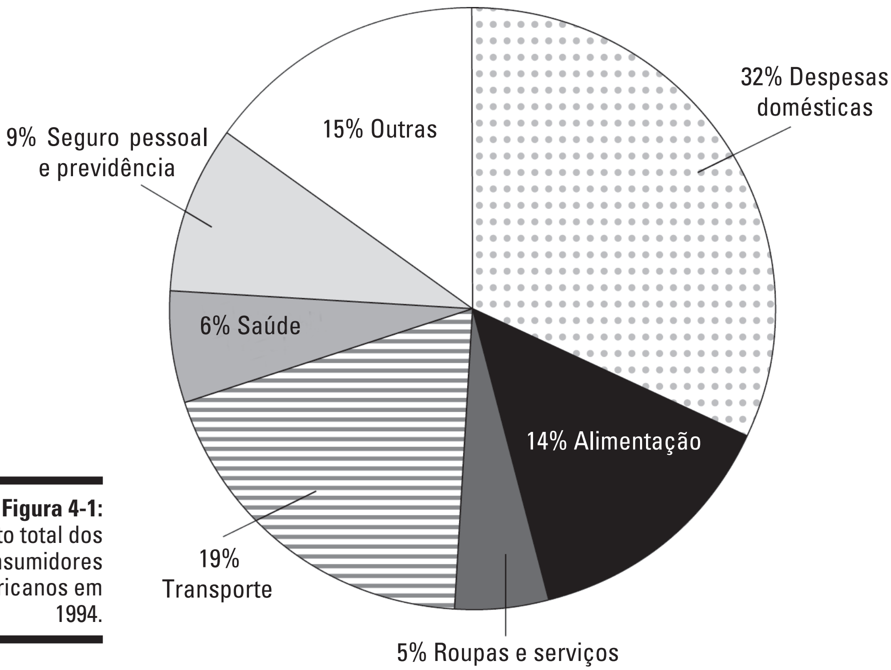


Figura 4-1: Gasto total dos consumidores americanos em 1994.

Analisando a loteria

As loterias estaduais americanas geram uma receita elevada e também devolvem uma grande porção do dinheiro recebido, sendo que parte da receita é dada como prêmio e parte é direcionada para os programas estaduais, tais como educação. De onde vem o dinheiro? A Figura 4-2 mostra um gráfico pizza exibindo os tipos de jogos e a porcentagem de receita gerada pela loteria do estado de Ohio.

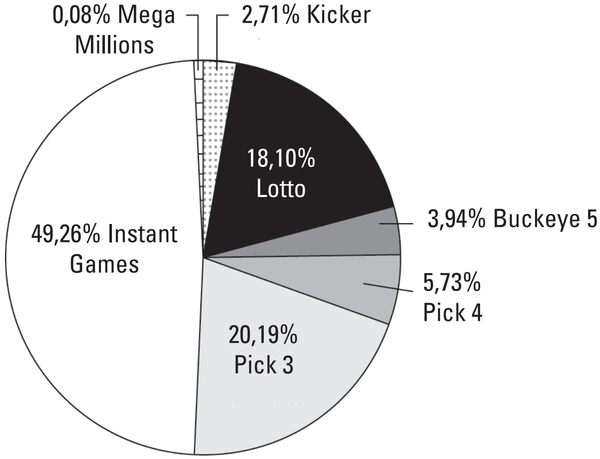


Figura 4-2: A divisão da receita gerada pela Loteria do estado de Ohio nos Estados Unidos (1993-2002).

Nesse gráfico, você pode observar que a maior parte da receita da loteria de Ohio (49,25%) vem dos Instant Games (a raspadinha). O restante da receita vem dos vários jogos do tipo loteria, em que os apostadores escolhem uma série de números e ganha aquele que acertar um determinada quantidade de números sorteados. Por que as raspadinhas contabilizam uma porção tão grande da receita de vendas da loteria? Uma das razões possíveis é que os prêmios pagos para esse jogo são frequentes, embora não sejam grandes. Além disso, a raspadinha dá o resultado imediatamente, enquanto que com os jogos de loteria você precisa esperar até que o sorteio ocorra para saber se é ou não o ganhador. Por outro lado, talvez as pessoas simplesmente se satisfaçam com o fato de raspar as figuras!

Note que esse gráfico pizza não fala quanto dinheiro foi ganho, somente fala qual a porcentagem de dinheiro foi ganho com cada tipo de jogo. Em outras palavras, para começar, você sabe que a pizza foi dividida, mas não sabe o tamanho dessa pizza. Isso é algo que você, como consumidor, deve saber. Cerca da metade do dinheiro (49,25%) vem das raspadinhas, essa receita representa um milhão de dólares, dois milhões, dez milhões ou mais? O gráfico pizza da Figura 4-2 não lhe traz essa informação, e você não consegue determiná-la sem saber o total da receita em dólares. No entanto, consegui encontrar essa informação em outro gráfico fornecido pela loteria de Ohio: a receita total gerada das vendas da loteria em 2002 foi registrada como sendo de “1.983,1 milhões de dólares” – popularmente conhecido como 1, 983 bilhões de dólares. Devido ao fato de que 49,25% da receita das vendas vêm de jogos instantâneos, essa porcentagem representa uma receita de $ 976.676.750 durante um período de dez anos. Haja raspadinha!

FIQUE ATENTO

Com frequência, os gráficos pizza mostram a divisão da porção ou da porcentagem do total que se enquadra em cada grupo ou categoria, mas quase nunca mostram o número total de cada grupo, em termos de unidades originais (número de dólares, número de pessoas e assim por diante). Essa abordagem gera a perda de informações e não necessariamente lhe apresenta toda a história por trás dos dados, além de deixar-lhe curioso para saber o total que está sendo dividido. Sempre é possível chegar a porcentagens partindo dos valores, mas é impossível chegar aos valores partindo das porcentagens e sem saber o valor total. Em se tratando de pesquisas, essa falta de informação pode ser um problema real; frequentemente, os gráficos pizza mostram a porcentagem de pessoas que deram uma determinada resposta, mas não dizem quantas pessoas responderam a pesquisa – uma fatia essencial da informação necessária para avaliar a precisão dos resultados (veja o Capítulo 10, para mais informações sobre precisão e margem de erro para pesquisas).

LEMBRE-SE

Sempre procure pelo número total de indivíduos para qualquer gráfico ou tabela. Se ele não estiver disponível, pergunte!

A loteria da Flórida utiliza um gráfico pizza para mostrar para onde seu dinheiro vai quando você compra um bilhete de loteria (ver a Figura 4-3). Você pode observar que metade da arrecadação da loteria da Flórida (50 centavos para cada dólar gasto) é destinada para os prêmios e 38 centavos de cada dólar são destinados para programas educacionais. Este gráfico não detalha a maneira como cada dólar arrecadado é gasto, mas você provavelmente também quer saber quantos dólares são gastos com os jogos da loteria da Flórida. As vendas de bilhetes da loteria da Flórida para 2001 totalizaram $2,360.6 milhões (ou 2.36 bilhões), ou seja, $147,70 per capita (ou seja, por pessoa), como mostrado na Tabela 4-1.

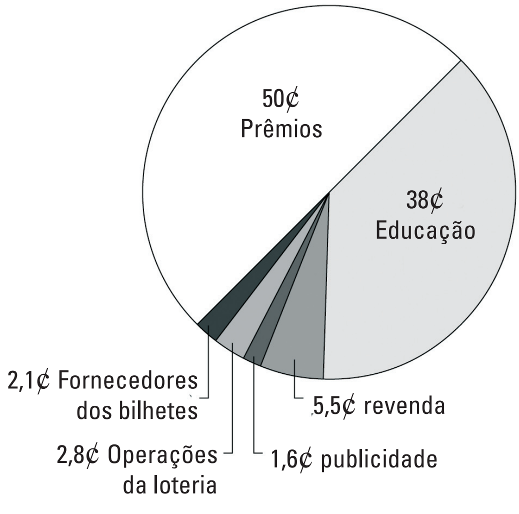


Figura 4-3: Gastos da loteria da Flórida (ano fiscal 2001-2002).

Tabela 4-1 A primeiras dez loterias (2001)

Classificação	Loteria	População (Milhões)	Venda de bilhetes (Milhões)	Renda Líquida (Milhões)	Lucro Líquido (Milhões)	Prêmio (% da receita)	Vendas per capita ($)
1	Nova Iorque	18.976	4,178	2,274	1,447	54.4%	220.16
2	Mass.	6.349	3,923	2,274	865	70.7%	617.85
3	Calofornia	33.872	2,896	1,492	1,048	51.5%	85.49
4	Texas	20.852	2,826	1,639	865	58.0%	135.50
5	Flórida	15.982	2,361	1,180	862	50.0%	147.70
6	Georgia	8.186	2,194	1,142	692	52.0%	267.98
7	Ohio	11.353	1,920	1,113	637	58.0%	169.11
8	Nova Jersey	8.414	1,807	991	695	54.8%	214.72
9	Pennsylvania	12.281	1,780	996	627	55.9%	144.93
10	Michigan	9.938	1,615	874	586	54.1%	162.49

É interessante notar que o web site da loteria de Michigan informa a quantidade em dólares que é anualmente destinada à educação, mas não indica a porcentagem da arrecadação total que é destinada a esse fim. Por exemplo, a quantia destinada à educação em 2001 pela loteria de Michigan foi de $587 milhões. Graças à tabela 4-1, sabe-se que a receita total das vendas da loteria de Michigan foi de $1.615 milhões ($1,6 bilhões) e, por isso, podemos calcular a porcentagem da receita que foi destinada aos programas educacionais naquele estado. Em Michigan, cerca de 36% ($587 milhões ÷ $1.615 milhões × 100%) do dinheiro arrecadado com a venda de bilhetes é destinado à educação.

O gráfico pizza é fácil de ser utilizado para comparar os tamanhos das fatias dentro de uma única pizza, mas também pode ser utilizado para comparar uma pizza inteira a outra. Por exemplo, a loteria de Nova Iorque informa seus gastos por meio de gráficos pizza (veja a Figura 4-4).

Prêmios	56%	$2, 664 Bilhões
Ajuda à Educação	33%	$1,58 Bilhões
Comissões de revendedores	6%	$284 Milhões
Gastos com funcionários	3%	$119 Milhões
Custos administrativos	2%	$107 Milhões

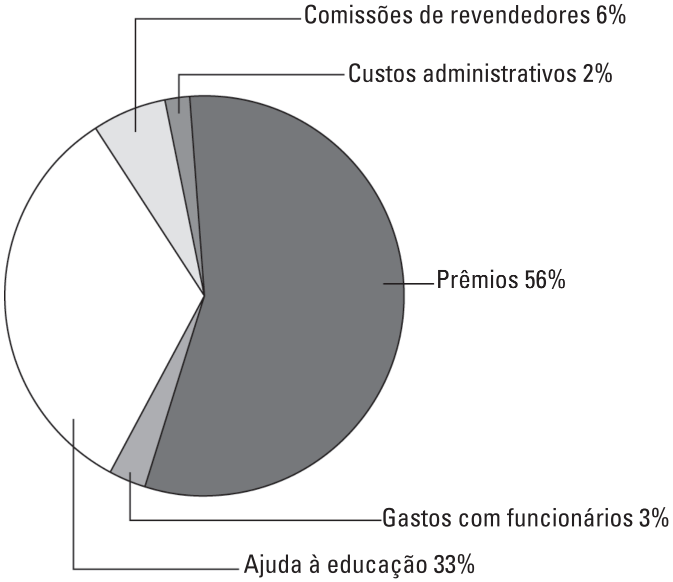


Figura 4-4: Gastos da loteria de Nova Iorque

Comparando as figuras 4-3 e 4-4, é possível notar que a loteria de Nova Iorque destina 56% do dinheiro arrecadado para o pagamento de prêmios (ligeiramente a mais do que na Flórida) e 33% para a educação (ligeiramente a menos do que na Flórida). Além disso, o gráfico pizza da loteria de Nova Iorque vem acompanhado de uma tabela mostrando a real quantia destinada em dólares, o que possibilita saber mais sobre a história toda por trás dos números (no entanto, a loteria de Nova Iorque faz você somar o total arrecadado por sua conta; É bem acima de $4,5 bilhões de dólares).

O estado de Nova Iorque também quer que você se dê conta de quanto dinheiro está sendo investido em educação, exibindo um gráfico pizza referente a todas as verbas recebidas pela educação (uma tática muito boa do ponto de vista político). A Figura 4-5 mostra que enquanto 4% das verbas para a educação no período de 2001-2002 vieram do governo federal, 5% vieram da loteria de Nova Iorque. Novamente, o gráfico vem acompanhado de uma tabela mostrando a real quantia em dólares (na realidade, a única quantidade necessária é a quantidade total em dólares, pois a partir dessa quantidade e das porcentagens exibidas no gráfico, você é capaz de calcular os números que estão na tabela. No entanto, não é nada mal receber tudo prontinho).

Receita local	45%	$15, 168 Bilhões
Ajuda de outros estados para a educação	42%	$14,148 Bilhões
Loteria de Nova Iorque	5%	$1,58 Bilhões
Ajuda Federal	4%	$1,48 Bilhões
Outros recursos	4%	$1,43 Bilhões

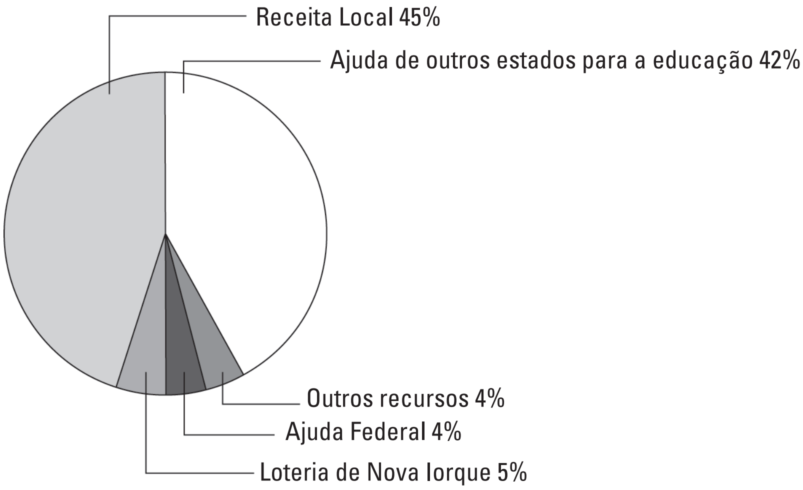


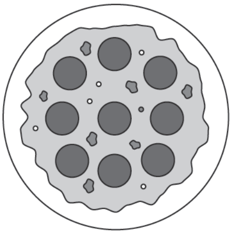
Figura 4-5: Verbas para a educação em Nova Iorque (2001-2002)

Fatiando seus impostos

O Internal Revenue Service (IRS) quer que você saiba onde o dinheiro de seus impostos está indo e se você lhes falar o quanto pagou de imposto o ano passado, eles lhe dirão como seu imposto foi fatiado. A Figura 4-6 mostra um exemplo dos resultados obtidos com o IRS, caso de ter pago $10.000 em impostos no ano passado.

Esse gráfico é criativo, mas um pouco diferente (será que eu serei auditada se disser "estranho"?). Primeiro, porque ele realmente se parece com uma pizza, mas o que você tem que se perguntar é o que esta pizza está fazendo aí, se não tem fatias para lhe mostrar para onde seu dinheiro está indo. As porcentagens aparecem na tabela que acompanha o gráfico, portanto, estão disponíveis. Este gráfico teria mais impacto visual se o IRS tivesse mostrado as "fatias" da pizza que correspondem às porcentagens da tabela. O interessante neste gráfico é que ele mostra tanto a quantidade em dinheiro quanto a porcentagem que está sendo gasta em cada área (a propósito, independentemente do total de dólares arrecadados em impostos, as porcentagens que mostram onde o dinheiro é investido não se alteram, o único valor a se alterar é a quantidade de dólares).

Gastos Federais



Seu dinheiro é gasto	Sua parcela	Porcentagens
Na defesa nacional	$1.700,00	17%
Sistema de saúde	$700,00	7%
Seguro saúde	$1.200,00	12%
Auxilio desemprego, pensão por invalidez e outros	$1.400,00	14%
Previdência social	$2.300,00	23%
Pagamentos de juros	$1.100,00	11%
Outras despesas	$1.600,00	16%
Total pago	**$10.000,00**	**100%**

Figura 4-6: Como seu imposto é fatiado (2202)

Ao analisar a Figura 4-6, você nota que a maior fatia de seu imposto vai para a Previdência Social (23%), e a segunda maior fatia vai para a defesa nacional (17%). Parece estranho, no entanto, que o IRS divide certas categorias em apenas um dígito (por exemplo, 7% para o sistema de saúde), mas a terceira maior fatia da pizza vai para "outras despesas" (16%).

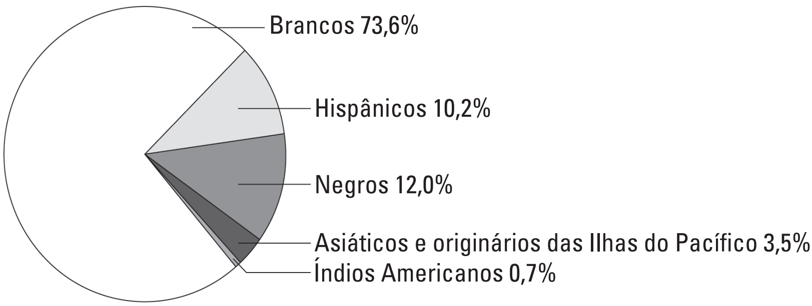
Um gráfico pizza não deve ter muitas fatias, pois um grande número de fatias distrai o leitor da questão que está sendo abordada. No entanto, agrupar todas as categorias restantes em uma chamada "outras" resulta em uma das maiores categoria de todo o gráfico e os leitores acabam se indagando sobre o que se inclui naquela fatia da pizza.

Talvez você agora esteja se perguntando o que são aquelas "outras despesas" no gráfico do IRS. Se você continuar sondando o web site do IRS, vai descobrir que a categoria "outras despesas" refere-se a "benefícios pagos a servidores públicos federais aposentados, subsídios a agricultores e outras atividades". Isso não fornece muita informação extra, mas talvez seja tudo que você queira realmente saber. Fazendo justiça ao IRS, tenho certeza que os detalhes estão esclarecidos em algum relatório governamental devidamente arquivado.

Prevendo as tendências populacionais

O U.S. Census Bureau fornece muitos gráficos em seus relatórios sobre a população americana. A Figura 4-7 mostra dois gráficos pizza, comparando a divisão racial nos Estados Unidos em 1995 (números reais) com a projeção da divisão racial para 2050, caso a tendência atual continue. Você pode observar que em 1995, cerca de 73,6% da população americana era branca, enquanto os negros constituíam o segundo maior grupo com 12%, seguindo pelos de origem hispânica, que compunham 10,2% da população (note que, embora os hispânicos possam ser brancos ou negros, aqui, eles são classificados como uma categoria separada, independente do histórico racial). O Census Bureau prevê que, no futuro, os Brancos serão uma fatia decrescente do total da população americana, enquanto a fatia de Hispânicos irá crescer mais rapidamente do que a de Negros não-Hispânicos. Essa tendência é muito bem mostrada com o uso de dois gráficos pizza, ao invés da simples utilização de tabelas mostrando porcentagens.

1995



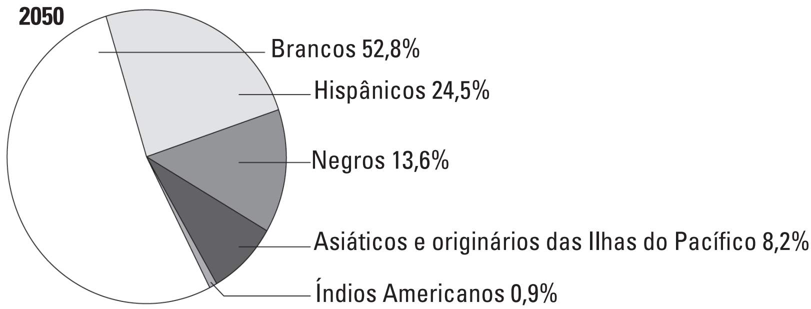


Figura 4-7: Tendência étnica para os Estados Unidos

Avaliando um gráfico pizza

DICA

Para provar a exatidão estatística de um gráfico pizza, experimente-o:

- Verifique se a soma das porcentagens é igual a 100% ou se se aproxima desse valor (quaisquer erros de arredondamento devem ser muito pequenos).
- Fique atento às fatias da pizza chamada de "outros/as" que sejam maiores do que várias outras fatias.
- Procure pela informação do número total de unidades, para que assim você possa determinar qual era o tamanho da pizza antes de ser dividida nas fatias que você está observando.

Levantando as barras de um Gráfico de Barras

O gráfico de barras talvez seja o gráfico mais comum utilizado pela mídia. Como o gráfico pizza, o gráfico de barras divide os dados categorizados em grupos, mostrando quantos estão em cada grupo. Um gráfico de barra, no entanto, representa esses grupos utilizando barras de diferentes comprimentos, ao invés de fatias de pizza de vários tamanhos. E, enquanto o gráfico pizza, com frequência, informa a quantidade em cada grupo por meio de porcentagem, o gráfico de barra utiliza ou o número de indivíduos em cada grupo ou a porcentagem do total em cada grupo. Em um gráfico de barras, o comprimento das barras indica o número ou a percentagem em cada grupo.

Rastreando os gastos com transporte

Quanto as pessoas gastam com transporte? Isso depende de quanto elas ganham. O Bureau of Transportation Statistics (Departamento de Estatísticas de Transporte) – você sabia que este departamento existe?-

realizou um estudo sobre o transporte nos Estados Unidos em 1994 e muitos de seus achados são apresentados em gráficos de barras como o mostrado na Figura 4-8.

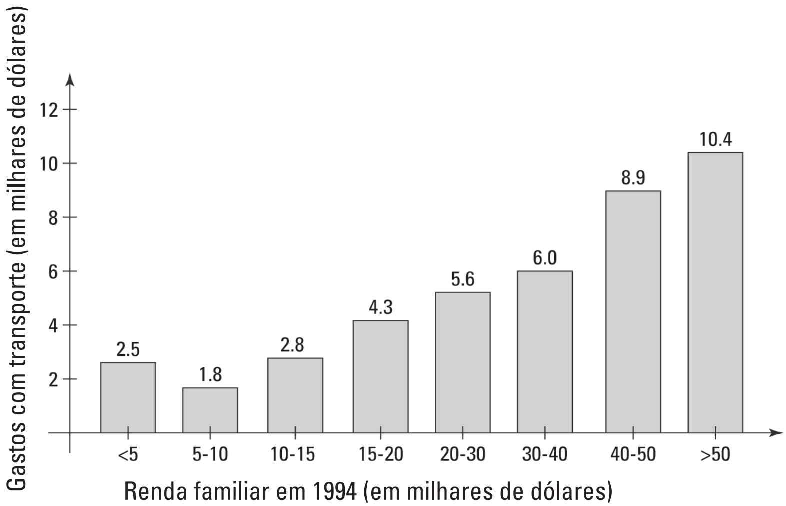


Figura 4-8: Gastos com transporte versus renda familiar em 1994.

Esse gráfico em particular mostra quanto dinheiro é gasto em transporte por pessoas com diferentes rendas familiares. Parece que, conforme a renda familiar aumenta, o gasto com transporte também aumenta. Isso provavelmente faz sentido, pois quanto mais dinheiro as pessoas têm, mais elas têm para gastar. Mas será que o gráfico seria diferente se olhássemos os gastos com transporte em termos de porcentagem da renda familiar e não em termo de quantidade total de dólares? O primeiro grupo ganha menos que $ 5.000 ao ano e tem que gastar $ 2.500 com transporte por ano (observe que no gráfico se lê "2.5", mas como as unidades estão em milhares de dólares, 2.5 se traduz em $2.500). Esses $2.500 representam 50% da renda anual daqueles que ganham $5.000 por ano; e representam uma porcentagem ainda maior para os que ganham menos do que $5.000 ao ano. As famílias cuja renda é de $30.000 - $40.000 por ano gastam $ 6.000 por ano com transporte, o que representa 15% a 20% de sua renda. Portanto, embora as pessoas que ganhem mais também gastem mais com transporte, elas não gastam mais com relação a porcentagem de sua renda total. Dependendo da maneira como você olha, o gráfico de barras pode contar duas versões diferentes da mesma história.

O gráfico de barras tem outra peculiaridade. As categorias para a renda familiar como mostrada não são equivalentes. Por exemplo, cada uma das primeiras quatro barras representa as rendas familiares com intervalos de $ 5.000, mas os três grupos seguintes aumentam o intervalo para $10.000 cada e o último grupo contém todas as famílias que ganham mais de $50.000 por ano, uma grande porcentagem das famílias, mesmo em 1994. Os gráficos de barras com diferentes variações de categorias, como o mostrado na Figura 4-8, dificulta a comparação entre os grupos.

Destacando as mães no mercado de trabalho

Os gráficos de barra são frequentemente utilizados para comparar dois grupos, ao dividir as categorias para cada grupo e mostrá-las como barras lado-lado. Um exemplo disso é mostrado na Figura 4-9, que responde à pergunta: "a porcentagem de mães no mercado de trabalho mudou ao longo dos anos?". A Figura 4-9 mostra que a porcentagem total de mães no mercado de trabalho subiu de 47% a 72% entre 1975 e 1998. Se levarmos a idade do filho em consideração, perceberemos que menos mães trabalham enquanto seus filhos são pequenos e ainda não estão na escola, mas a diferença entre 1975 a 1998 ainda é cerca de 25% em cada caso, como mostrado nas barras lado a lado.

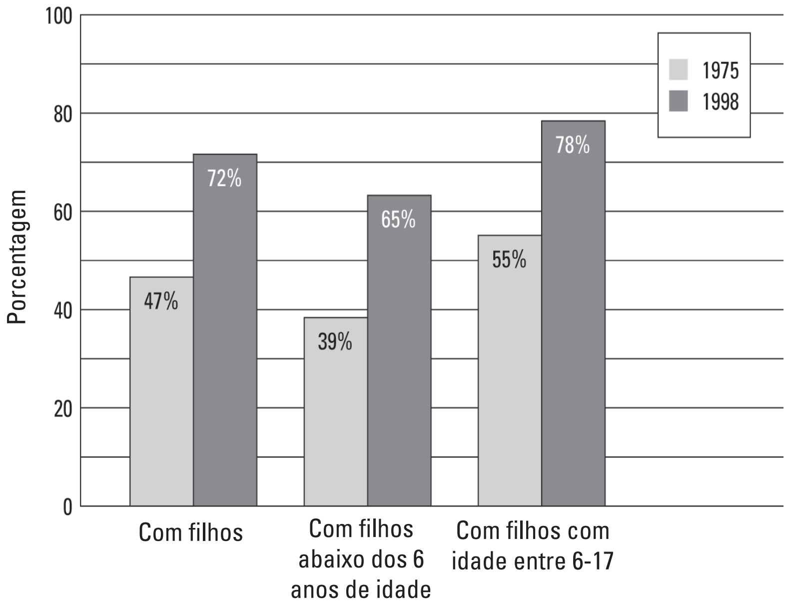


Figura 4-9: Porcentagem de mães no mercado de trabalho versus a idade dos filhos (1975 e 1998).

Jogando a loteria de Ohio

A loteria de Ohio mostra suas vendas e gastos para 2002 utilizando o gráfico de barra (veja a Figura 4-10). Esse gráfico de barra requer um trabalho extra nos bastidores para que ele seja compreensível. A primeira questão com relação a esse gráfico é que as barras não representam tipos similares de entidades. A primeira barra representa as vendas (uma forma de receita) e as outras barras representam os gastos. Este gráfico de barras poderia ser mais claro se a primeira barra não tivesse sido incluída; por exemplo, a venda total poderia ser listada como uma nota de rodapé. Também, os gastos poderiam ser representados em um gráfico pizza, como é feito por algumas outras loterias estaduais (como as Figuras 4-3 e 4-4). A questão seguinte é que a soma de todos os gastos ($2.013,2 milhões - em outras palavras, $ 2, 0132 bilhões) é maior do que as vendas ($1, 9831 bilhões), portanto alguma receita adicional não está sendo mostrada neste gráfico de barra (ou é isso, ou a loteria de Ohio está quase falindo!). Observando mais atentamente a informação fornecida pelo web site da loteria de Ohio, descobri que, além das vendas, uma receita adicional foi informada sendo de $ 124,1 milhões, advinda de juros e outras receitas". Isso contabiliza a receita total de 1,9831 bilhões + $ 124,1 milhão = $2,1072 bilhões em 2002, gerando um lucro de $ 2,1072 bilhões - $ 2,0132 bilhões = $,094 bilhões ou $ 94.000.000.

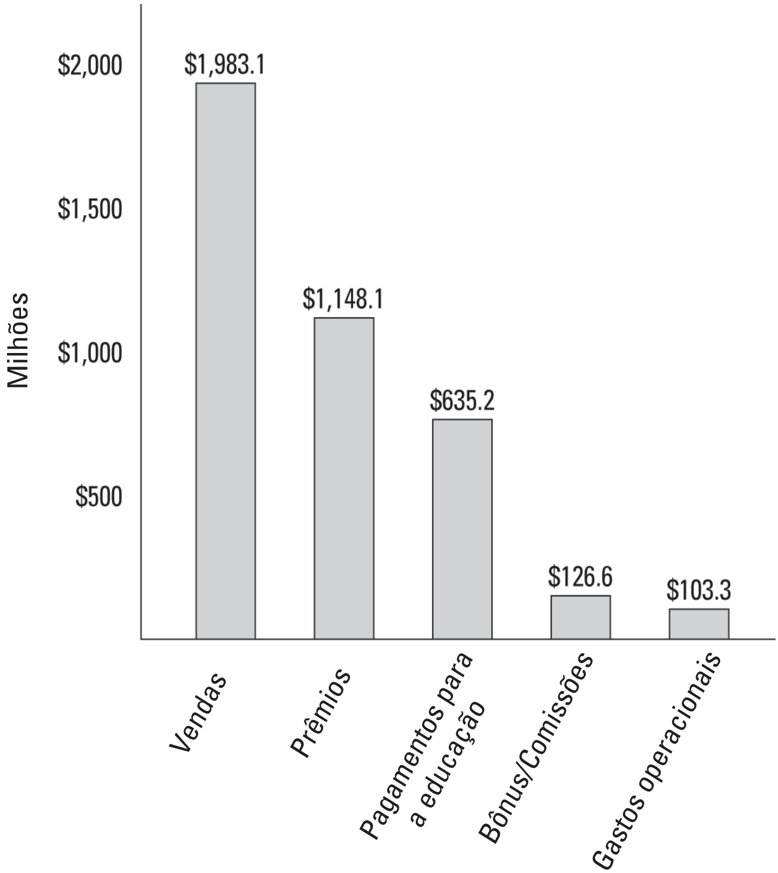


Figura 4-10: Vendas e Gastos da loteria de Ohio em 2002.

Observe que ao longo de todo o exemplo da loteria de Ohio, as unidades estavam em termo de milhões, portanto, você viu números engraçados como $ 1.983,1 milhões, quando na realidade deveriam ser escritos como $ 1,9831 bilhões. Por que a loteria de Ohio utiliza a unidade milhões? Talvez para não parecer que está ganhando tanto dinheiro como realmente está. Os números realmente parecem um pouco domados na Figura 4-10. Não pense que as loterias simplesmente deixam passar essas sutilezas: elas são mestres em sutilezas (pense nisso: a fim de descobrir os lucros, você mesmo tem que fazer os cálculos e, além disso, precisa chegar a dois lugares diferentes para conseguir os números necessários para tais cálculos. As loterias não facilitam!).

Não pense que a informação apresentada em gráficos e tabelas representa tudo o que você necessita saber; esteja preparado para ir mais a fundo, caso precise completar quaisquer informações omitidas (fim para o qual gráficos e tabelas são muito utilizados pela mídia!). Normalmente, você não leva muito tempo para encontrar o que estava procurando (ou, pelo menos, para descartar a informação que lhe está sendo apresentada, caso o que você encontre mostre parcialidade ou imprecisão).

Os gráficos de barra permitem uma grande licença poética para quem quer que os projete. Isso porque a pessoa que constrói o gráfico de barra determina a escala que ele ou ela deseja utilizar, o que significa que a informação pode ser apresentada de uma maneira enganosa. Ao utilizar uma escala menor (por exemplo, colocando meia polegada da altura da barra para representar 10 unidades versus 50 unidades), você pode distorcer a verdade, fazendo as diferenças parecerem mais drásticas, ou exagerar os valores. Ao utilizar uma escala maior (por exemplo, colocando meia polegada da altura da barra para representar 50 unidades versus 10 unidades) você pode minimizar as diferenças, fazendo os resultados parecerem menos drásticos do que eles realmente são e, até mesmo, fazer com que pequenas diferenças pareçam nem existir (veja o Capítulo 2 para exemplos dessa manipulação).

No entanto, observe que em um gráfico pizza a escala não pode ser alterada para super enfatizar (ou minimizar) os resultados. Independente de como você fatia o gráfico pizza, você sempre está fatiando um círculo e a proporção do gráfico total pertencente a quaisquer fatias não se alterará, mesmo que você faça a pizza maior ou menor.

Avaliando o gráfico de barras

Para avaliar as barras estatísticas nos gráficos de barra, verifique estes itens:

- Barras que dividem os valores das variáveis numéricas (tais como a renda financeira) deveriam ser de largura igual para facilitar a comparação.
- Fique atento para a escala do gráfico de barras (as unidade na qual a altura das barras está representada) e determine se esta é uma representação apropriada da informação.
- Não pense que a informação contida em um gráfico de barras é tudo que você precisa saber; esteja pronto para ir mais a fundo caso seja necessário.

Estatística de Tabela

A tabela é uma das formas utilizada para exibir dados, resumindo a informação obtida a partir de um conjunto de dados em um formato linha e coluna. Algumas tabelas são claras e fáceis de ler, outras, no entanto, deixam um pouco a desejar. Embora a intenção ao se usar um gráfico pizza ou um gráfico de barras seja a de enfatizar um ou dois ponto no máximo, uma tabela pode enfatizar vários pontos de uma só vez (o que pode ser bom ou ruim, dependendo do efeito que isso pode causar no leitor).

As informações estatísticas são compiladas pelos pesquisadores, não apenas para seus próprios relatórios, mas para que outras pessoas também possam utilizar esses dados para suas próprias pesquisas ou para responder suas próprias perguntas. As tabelas geralmente são utilizadas nessas situações.

Examinando as estatísticas de nascimento

O Departamento de Saúde Pública e Meio Ambiente do Colorado compila tabelas sobre as estatísticas de nascimentos para os residentes do estado. A tabela 4-2 mostra o número de nascidos vivos pelo sexo da criança e o status da pluralidade (nascimentos únicos versus nascimento de gêmeos, trigêmeos e assim por diante) para o período selecionado entre 1975-2000. Algumas questões que podem ser respondidas por meio dessa tabela são: Qual a taxa de natalidade de meninos comparada com a taxa de nascimento de meninas no Colorado? A taxa de nascimentos múltiplos está alterando? A partir dessa tabela, você pode observar que ao longo de um período de 25 anos, a porcentagem de nascimentos de meninas permaneceu constante, logo abaixo dos 49%, enquanto a porcentagem dos nascimentos de meninos permaneceu constante um pouco acima dos 51% (você pode estar se perguntado o porquê essas porcentagens não estão mais próximas de 50% cada uma. Essa é uma pergunta para os demógrafos – cientistas que estudam as tendências populacionais – e biólogos, não estatísticos). Também é possível notar que a taxa de nascimentos múltiplos (em oposição aos nascimentos únicos) parece ter se alterado ao longo dos anos. Parece que a porcentagem dos nascimentos múltiplos tem aumentado, mas em qual coluna você deve olhar: na de números de nascimentos múltiplos ou na de porcentagens de nascimentos múltiplos? Mas isso faz diferença? A resposta é sim!

Porcentagem versus total

Como tirar conclusões acerca das tendências sobre os nascimentos múltiplos ao longo do tempo utilizando as estatísticas apresentadas nessa tabela? Se você observar apenas os números dos nascimentos múltiplos em 1975 comparado aos de 2000, eles saltaram de 763 para 1.982. Alguns podem tentar dizer que isso representa um aumento de 160%, ou cerca de 1,6 vezes mais nascimentos múltiplos em 25 anos ([1.982 – 763] ÷ 763). Realmente houve mais nascimentos múltiplos em 2000 do que em 1975, mas, também, houve mais nascimentos únicos ao longo desse período. Em virtude disso, a única maneira precisa de comparar essas estatísticas é calcular a porcentagem de nascimentos únicos versus a porcentagem de nascimentos múltiplos e compará-las. Observando a Tabela 4-2 é possível notar que a porcentagem de nascimentos múltiplos em 1975 era de 1,9%, enquanto que em 2000 a porcentagem de nascimentos múltiplos era de 3,0%. Você pode concluir que a porcentagem de nascimentos múltiplos realmente aumentou ao longo do tempo, mesmo depois de levar em consideração o elevado número de nascimentos. Entretanto, esse aumento não é de 160%, mas se aproxima de 58%: ([3,0 – 1,9]) ÷ 1,9) × 100%.

Tome cuidado com as conclusões tiradas a partir de gráficos e tabelas que comparam o número de indivíduos em oposição à porcentagem de indivíduos. As porcentagens representam uma comparação relativa de quantidades (e com frequência levando em consideração um período de tempo); essa é, geralmente, uma maneira precisa de comparar quantidades, especialmente quando o número total de itens ou eventos também se altera ao longo do tempo. Ao observar a mudança das porcentagens, deve-se levar em consideração o fato de que o número total de itens ou eventos também se alterou (claro que se você estiver mesmo interessado em examinar como o número de cada item se alterou, você deve observar os números e não as porcentagens).

Tabela 4-2 Nascimentos Vivos em Colorado de Acordo com o Sexo e com o Status da Pluralidade

Ano	*Número total de nascimentos*	*Número total de nascimentos de meninas*	*% de Nascimentos de Meninas*	*Número total de nascimentos de meninos*	*% de Nascimentos de meninos*	*Número de nascimentos únicos*	*% de Nascimentos únicos*	*Número de nascimentos múltiplos*	*% de Nascimentos múltiplos*
1975	40.148	19.447	48,4	20.701	51,6	39.385	98,1	763	1,9
1980	49.716	24.282	48,8	25.434	51,2	48.771	98,1	945	1,9
1985	55.115	26.925	48,9	28.190	51,1	53.949	97,9	1.166	2,1
1990	53.491	26.097	48,8	27.394	51,2	52.245	97,7	1.246	2,3
1995	54.310	26.431	48,7	27.879	51,3	52.669	97,0	1.641	3,0
2000	65.429	31.953	48,8	33.476	51,2	63.447	97,0	1.982	3,0

A Tabela 4-3 mostra a divisão do número de nascimentos vivos no estado do Colorado em comparação com a idade da mãe, para o período selecionado de 1974 a 2000. A variável numérica idade é dividida em categorias que tem a mesma medida (5 anos) e não se sobrepõe. Isso propicia uma comparação justa e imparcial entre as faixas etárias. No entanto, a tabela apenas mostra o número de nascimentos em cada caso, por isso não é possível observar nenhuma tendência que possa estar se desenvolvendo ao longo do tempo, em termos da idade materna, apenas por meio desta tabela. Este problema poderia ser resolvido, colocando-se as porcentagens em parênteses junto com o número total de cada categoria, para que o leitor pudesse facilmente fazer a comparação. Outra maneira de incluir essa informação é inserindo um gráfico pizza para cada ano, onde se possa observar a porcentagem do total de nascimentos para as mulheres em cada uma das oito categorias de faixa etária que não se sobrepõe.

Tabela 4-3 Nascimentos Vivos em Colorado de Acordo com a Idade da Mãe

Ano	*Número Total de Nascimentos**	*10-14*	*15-19*	*20-24*	*25-29*	*30-34*	*35-39*	*40-44*	*45-49*
1975	40.148	88	6.627	14.553	12.565	4.885	1.211	222	16
1980	49.716	57	6.530	16.642	16.081	8.349	1.842	198	12
1985	55.115	90	5.634	16.242	18.065	11.231	3.464	370	13
1990	53.491	91	5.975	13.118	16.352	12.244	4.772	717	15
1995	54.310	134	6.462	12.935	14.286	13.186	6.184	1.071	38
2000	65.429	117	7.546	15.865	17.408	15.275	7.546	1.545	93

**Observe: talvez a soma dos nascimentos não resulte no número total de nascimentos, devido aos nascimentos ocorridos de mulheres com idade acima de 50 anos.*

Note que, como os totais são apresentados nessa tabela, você consegue encontrar as porcentagens sozinho, se você quiser (se a tabela apenas apresentasse as porcentagens, sem nenhum total, seria fácil compará-las, mas suas conclusões seriam limitadas, pois você não teria como saber os números totais). Apenas para poupar o seu tempo, calculei as porcentagens para o grupo das mães com idades entre 40-49. A Tabela 4-4 mostra esses cálculos. A partir desta tabela, você consegue ver que a tendência com relação à idade maternal parece estar emergindo. Mais e mais mulheres estão tendo filhos aos 40 anos, e a porcentagem continua em ascensão.

O rodapé da Tabela 4-3 (parafraseado da observação escrita originalmente pelo Departamento de Saúde Publica e Meio Ambiente do estado do Colorado) indica que todas as mães com idade acima de 50 anos não foram incluídas nesse conjunto de dados. Estudos recentes demonstram que existe uma porcentagem crescente (mas ainda pequena) de mulheres que estão tendo seus filhos com 50 anos. Sendo assim, esse conjunto de dados pode ser aumentado com o tempo para incluir essa faixa etária.

Tabela 4-4 Porcentagem de Nascimentos Vivos no Colorado com relação a Mulheres com idades de 40-49 anos

Ano	*Total de Nascimentos*	*Número de Nascimentos para Mulheres com idade entre 40-49 anos*	*% de Nascimentos para Mulheres com idade entre 40-49 anos*
1975	40.148	238	0,59%
1980	49.716	210	0,42%
1985	55.115	383	0,69%
1990	53.491	732	1,4%
1995	54.310	1.109	2,0%
2000	65.429	1.638	2,5%

Colocando as porcentagens em perspectiva

Não se engane pensando que só porque algumas porcentagens são pequenas, elas não são significativas e/ou sirvam para comparação. Todas as porcentagens na Tabela 4-4 são pequenas (iguais ou menores do que 2,5%), mas a porcentagem para o ano de 2000 (2,5%) é quatro vezes maior do que a porcentagem para o ano de 1975 (0,59%) e, em termos relativos, isso representa um aumento muito elevado. Da mesma forma, não presuma que o aumento elevado da porcentagem signifique o envolvimento de um grande número de pessoas. Suponha que alguém anuncie que a taxa de determinada doença quadruplicou ao longo dos últimos anos. Isso não significa que uma grande porcentagem de pessoas esteja infectada, mas apenas significa que a porcentagem está quatro vezes maior do que era. Para começar, a porcentagem de pessoas infectadas com a doença em questão podia ser extremamente pequena. Um aumento ainda é um aumento, mas, em algumas situações, relatar apenas as porcentagens pode causar enganos; a alta incidência da doença precisa ser colocada em perspectiva em termos do número total de pessoas infectadas.

LEMBRE-SE

A porcentagem é uma medida relativa. Por isso, também procure pelo número total, a fim de manter as reais quantidades dentro de uma perspectiva apropriada.

Fique de olho nas unidades

Às vezes, as tabelas podem ser um pouco confusas se você não as observar cuidadosamente. Por exemplo, o IRS anuncia "Resumo das Estatísticas de Impostos" em seu site na internet, e algumas dessas estatísticas (exatamente como foram informadas pelo IRS) são mostradas na Tabela 4-5.

Tabela 4-5 Estatística da Restituição de Imposto de Renda Individual

Número de Restituições (AF 2001)	129.783.221
Arrecadação Bruta (AF 2001 em milhões de dólares)	1.178.210
Desconto acima de 1% para RBA (AB 1999)	$293.415
Desconto acima de 10% para RBA (AB 1999)	$87.682
Desconto menor do que 10% para RBA (AB 1999)	$4.718
Mediana da Renda Bruta Ajustada (RBA, AF 2000)	$27.355
Porcentagem de descontos padrões reivindicados (AB 2000)	66,2%
Porcentagem de itens de descontos reivindicados (AB 2000)	32,9%
Porcentagem dos que pagaram pelo serviço de preenchimento de declaração (AB 2000)	53,4%
Porcentagem dos que declararam via internet (AB 2001)	38,3%
Número de reembolsos com RBA > $1 milhão (AB 2000)	241.068
Número de reembolsos individuais (AB 2000 em milhões)	93,0
Quantia de reembolsos individuais (AB 2000 em bilhões de dólares)	167,6

Duas características dessa tabela podem ser imediatamente notadas. Primeiro, o IRS está informando as estatísticas para anos diferentes na mesma tabela, por exemplo, AF 2001 (que significa ano fiscal de 2001 ou o período de 12 meses a partir de 1 de julho de 2000 até 30 de junho de 2001), AB 1999 (que significa ano base 1999) e AB 2000. Observe que o ano base e o ano fiscal se sobrepõem e que as restituições de imposto para o ano base 2000, por exemplo, referem-se a Abril de 2001 (que está no ano fiscal 2001). Que confuso, não? Também perceba que não é possível comparar a mediana da renda bruta ajustada (mediana RBA) nessa tabela com os descontos acima de 1% e 10% da RBA na mesma tabela, pois esses números representam anos diferentes (ano base 2000 2 anos base 1999, respectivamente).

Em segundo lugar, a maneira como o IRS informa as unidades de dólares pode levar a equívocos. Por exemplo, as arrecadações brutas para os reembolsos individuais para o ano fiscal de 2001 (informada em milhões de dólares) é escrita como 1.178.210. Isso significa que $ 1.178.210 milhões de dólares foram arrecadados por meio de reembolsos individuais. No entanto, essa quantidade de dólares não deveria ter sido representada dessa maneira. Para colocá-la na perspectiva correta, esse número deveria ser escrito da seguinte forma: $1.178.210.000.000, ou seja, $ 1, 178 trilhões. Não me admira o fato do IRS não ter informado essa receita dessa maneira, é número grande demais para se entender.

Quando você for analisar uma tabela, certifique-se de que tenha entendido as unidades que estão sendo expressas e fique atento às mudanças de unidades (tais como anos) ao longo da tabela.

A Tabela 4-5 pretende mostrar vários pontos; aqui vão alguns deles. No ano base de 2000, a proporção de pessoas que reivindicaram o desconto padrão em relação às pessoas que reivindicaram os itens de desconto é cerca de 2 para 1 (um gráfico pizza poderia mostrar isso muito bem). Cerca de metade das pessoas que entregaram a declaração de reembolso no ano base de 2000 utilizaram serviços pagos a porcentagem dos que entregaram o documento via internet no ano base de 2001 foi de cerca de 38% (uma porcentagem que o IRS provavelmente gostaria de ver aumentar a cada ano). A média de reembolso para o ano base de 2000 foi de $1.802,15 (o total de reembolso em bilhões de dólares, dividido pelo número total de reembolsos em milhões). (Agora, não é mais impactante informar o total de reembolso em dólares do que a média de reembolsos?). Observe também que o IRS não informou o número total de declarações entregues para o ano base de 2000 nem a porcentagem de contribuintes que receberam o reembolso referente a aquele ano base. Saber esta porcentagem seria mais útil do que saber o número total de reembolsos em um dado ano base. Afinal de contas, nem todos conseguiram o reembolso – muitas pessoas tiveram que pagar ou pagaram exatamente o valor integral.

As tabelas são projetadas para tornar determinados pontos mais proeminentes e outros menos notáveis. Às vezes, a informação deixada de lado – ou, até mesmo, omitida – é a mais reveladora!

Avaliando uma tabela

Para descobrir se a tabela é estatisticamente suficiente, siga os seguintes passos:

- Conheça a diferença entre as porcentagens e números totais e como essas duas estatísticas são utilizadas na interpretação dos resultados. As porcentagens são, com frequência, a estatística mais adequada para a comparação entre diferentes resultados.
- Certifique-se, com relação aos dados numéricos, se os grupos da tabela não se sobrepõem e que estejam divididos de maneira equilibrada para se chegar a uma comparação imparcial.
- Observe atentamente as unidades e como elas estão apresentadas na tabela.
- Observe o modo como a informação é apresentada. Frequentemente, as tabelas são projetadas para diminuir a importância de certos pontos, ao mesmo tempo em que enfatizam apenas os pontos que são convenientes para pesquisadores e jornalistas.

Alinhado-se ao Gráfico de Linhas

O gráfico de linhas é uma maneira de exibir os dados cujo maior objetivo é o de examinar as tendências ao longo do tempo. Normalmente, um gráfico de linhas apresenta algumas unidades de tempo no eixo horizontal (tais como anos, dias, meses e assim por diante) e alguns valores no eixo vertical (tais como a renda familiar média, a taxa de nascimento, vendas totais, porcentagem de pessoas a favor do presidente e outras). Em cada período, o valor é representado por um ponto, e esses pontos são ligados por um traço para formar um gráfico de linhas.

Analisando as tendências salariais

Em 1999, o U.S Bureau of Labor Statistics emitiu um relatório sobre as tendências trabalhistas nos Estados Unidos e as perspectivas para o futuro. Esse relatório inclui muitos gráficos de linhas, inclusive os dois mostrados na Figura 4-11 e na Figura 4-12. A Figura 4-11 mostra a tendência ao longo do tempo com relação à remuneração por hora de operários da indústria no período de 1947 a 1998 (devido à inflação, não faria sentido apenas mostrar a real remuneração por hora ao longo desse período. Você necessita saber a informação em termos de "remuneração real" – em outras palavras, algo que possa ser comparado ao longo do tempo. Aqui, o bureau mostra tudo em termos do dólar de 1998 para uma comparação imparcial ao longo do tempo). Você pode observar, a partir da Figura

4-11, que a remuneração de operários da indústria aumentou de 1947 até o início dos anos 70, declinou durante os anos 70 e, basicamente, permanece com a mesma variação até o final dos anos 90, quando um aumento começa a surgir.

A Figura 4-12 enfatiza que a lacuna salarial entre os trabalhadores com diferentes níveis de escolaridade tem se ampliado entre 1979 e 1997.

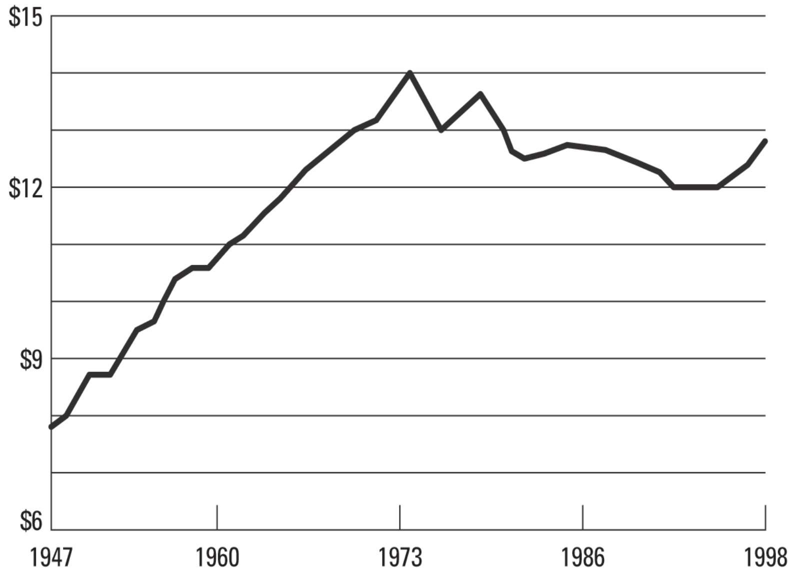


Figura 4-11: Remuneração média por hora para os operários da indústria no período de 1947 a 1998 (com o valor do dólar referente a 1998).

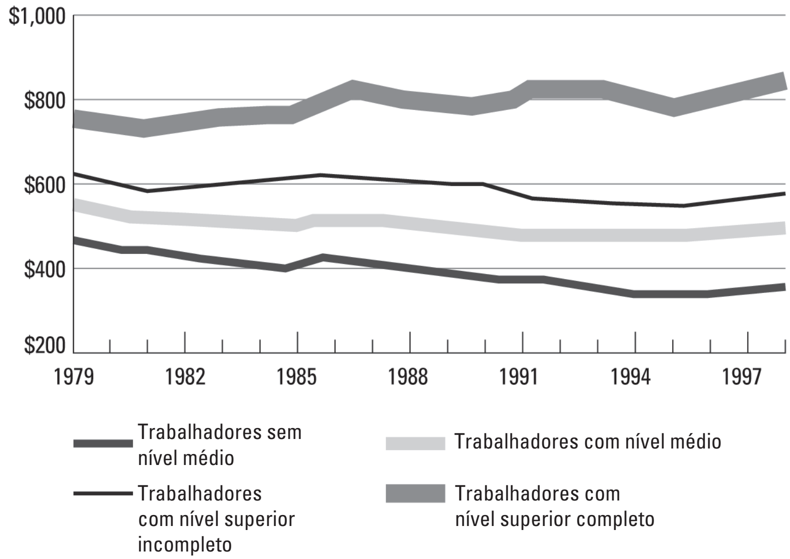


Figura 4-12: Remuneração semanal média com relação ao nível de escolaridade no período de 1979-1997 (com o valor do dólar referente a 1998).

As estatísticas apontam fatos – ou seja, elas informam o que está se passando, mas não explicam o porquê de tais eventos estarem ocorrendo. O relatório emitido pelo Bureau of Labor Statistics não apenas apresenta dados demonstrando as tendências salariais, mas, também, vai além da apresentação dos dados para discutir algumas das razões pelas quais a média salarial da classe operária estagnou desde o final dos anos 70 até a metade dos anos 90 e o porquê a lacuna salarial entre os trabalhadores com mais e menos nível de escolaridade está aumentando. Responder à pergunta: "por quê?"; é muito mais complicado do que responder à pergunta: "qual?". Embora o Bureau of Labor Statistics certamente tenha outras estatísticas que sustentem suas avaliações de porque as tendências estão causando tais efeitos, nem todo mundo que emite uma estatística tem esse suporte.

Muitos tentam usar um simples gráfico não apenas para mostrar o que está acontecendo, mas também para tentar explicar o porquê as coisas estão acontecendo dessa maneira. Sem terem dados suficientes, essas pessoas podem estar esboçando conclusões falsas. Se você achar que alguém está indo muito longe com suas conclusões, questione se tais conclusões têm fundamento.

Traçando os nascimentos múltiplos

Ao demonstrar as estatísticas de nascimentos para os habitantes do estado do Colorado, um gráfico de linhas pode ser utilizado para examinar a tendência da taxa dos nascimentos múltiplos com o passar do tempo. Tal gráfico se pareceria com o ilustrado na Figura 4-13. Você pode notar que a porcentagem de nascimentos múltiplos parece estar crescendo com o passar do tempo, especialmente se você simplesmente comparar o período compreendido entre 1975 a 2000.

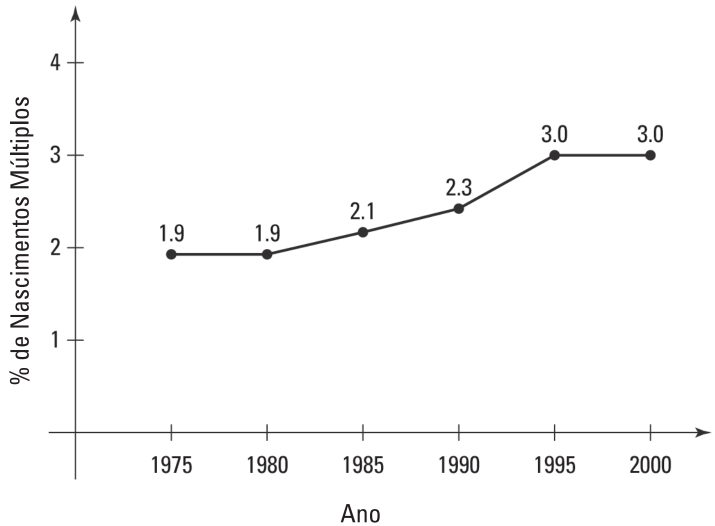


Figura 4-13: Porcentagem dos nascimentos múltiplos para os habitantes do estado do Colorado no período de 1975-2000.

Assim como acontece com o gráfico de barras, as diferenças representadas por meio de um gráfico de linhas podem ser manipuladas através da alteração da escala nos eixos verticais. Portanto, certifique-se de ter levado a escala em consideração ao analisar os resultados de um gráfico de linhas

Para o exemplo dos nascimentos múltiplos no estado do Colorado, o aumento da porcentagem dos nascimentos múltiplos ao longo dos anos pode parecer mais drástico se alterássemos as faixas de aumento no eixo vertical de 1% para 0,2%, fazendo com que o gráfico aumentasse verticalmente. Da mesma forma, essa tendência poderia parecer quase inexistente se alterássemos as faixas de incremento do eixo vertical de 1% para 5%. Os aumentos mais razoáveis ficam em torno de 0,2% e 5%. Eu escolhi 1% para a minha versão desse gráfico.

Outro fator a ser considerado com relação aos gráficos de linha é o eixo do tempo, ou a proximidade ou a distância em que estão os pontos de dados. Dados que estejam próximos no eixo do tempo, mas que possuam quantidades muito diferente, farão com que o gráfico pareça serrilhado. Isso é muito comum com os gráficos que representam dados voláteis, assim como os preços de ações ao longo de um dia, como você já deve ter visto em programas de televisão sobre economia. Outros gráficos de linha podem mostrar mudanças mais radicais a longo prazo, tais como a demonstrada na Figura 4-13, que apenas mostra a porcentagem dos nascimentos múltiplos de cinco em cinco anos e não anualmente. Mais uma vez, isso depende de como as pessoas que estão criando o gráfico quer que você veja a informação. É necessário que você realmente pense sobre como o gráfico foi estabelecido e, potencialmente, faça perguntas adicionais, a fim de esclarecer quaisquer pontos obscuros.

Você também pode se deparar com um problema em que um gráfico de linhas apresente uma informação de maneira desonesta como, por exemplo, exibindo o número de crimes ao longo de um período ao invés de mostrar a taxa de criminalidade (número de crimes per capita). Certifique-se de que você tenha compreendido quais estatísticas estão sendo apresentadas em tal gráfico e depois avalie sua imparcialidade e adequação (veja o Capítulo 2 para mais sobre o assunto).

Avaliando o gráfico de linhas

Para ver se um gráfico de linhas está estatisticamente alinhado:

- Examine a escala do eixo vertical (valor) assim como o eixo horizontal (linha do tempo), os resultados podem parecer mais ou menos drásticos do que realmente são com apenas uma alteração da escala.
- Leve em consideração as unidades utilizadas no gráfico e assegure-se de que elas sejam apropriadas para uma comparação ao longo de um período (por exemplo, os dólares foram ajustados de acordo com a inflação?).
- Tome cuidado com pessoas que tentam explicar os motivos de uma tendência sem utilizar estatísticas extras para sustentar suas afirmações. Um gráfico de linhas mostra apenas o que está acontecendo. O porquê de algo estár ocorrendo é uma outra história!

Retratando os dados em um Histograma

Os dados numéricos em sua forma crua e desorganizada são de difícil compreensão. Por exemplo, veja a Tabela 4-6, que mostra a estimativa populacional feita em 2000 para os 50 estados americanos (e o Distrito da Columbia), colocados pelo U.S. Census Bureau. Olhe fixamente para a tabela por cerca de 30 segundos. Depois de fazer isso, vá em frente e tente responder a estas perguntas rapidamente:

- Quais estados têm a maior e a menor população?
- Quantas pessoas habitam a maioria dos estados? Dê uma variação aproximada de valores
- Qual é a variabilidade existente entre as populações de cada estado? (Os estados são muito semelhantes ou muito diferentes uns dos outros com relação ao total populacional?)

Tabela 4-6 Estimativa Populacional por Estado (Censo de 2000)

Estado	*Censo Populacional de 2000*
Alabama	4.447.100
Alaska	626.932
Arizona	5.130.632
Arkansas	2.673.400
Califórnia	33.871.648
Colorado	4.301.261
Connecticut	3.405.565
Delaware	783.600
Distrito de Colunbia	572.059
Florida	15.982.378
Georgia	8.186.453
Hawaii	1.211.537
Idaho	1.293.953
Illinois	12.419.293
Indiana	6.080.485
Iowa	2.926.324
Kansas	2.688.418
Kentucky	4.041.769
Louisiana	4.468.976
Maine	1.274.923

Estado	*Censo Populacional de 2000*
Maryland	5.296.486
Massachusetts	6.349.097
Michigan	9.938.444
Minnesota	4.919.479
Mississippi	2.844.658
Missouri	5.595.211
Montana	902.195
Nebraska	1.711.263
Nevada	1.998.257
New Hampshire	1.235.786
New Jersey	8.414.350
New Mexico	1.819.046
New York	18.976.457
North Carolina	8.049.313
North Dakota	642.200
Ohio	11.353.140
Oklahoma	3.450.654
Oregon	3.421.399
Pennsylvania	12.281.054
Rhode Island	1.048.319
South Carolina	4.012.012
South Dakota	754.844
Tennessee	5.689.283
Texas	20.851.820
Utah	2.233.169
Vermont	608.827
Virginia	7.078.515
Washington	5.894.121
West Virginia	1.808.344
Wisconsin	5.363.675
Wyoming	493.782
Total EUA	**281.421.906**

Sem encontrar uma maneira de organizar esses dados, você terá dificuldade para responder aquelas perguntas. Embora a mídia favoreça o uso de tabelas para organizar os dados numéricos, os estatísticos preferem o histograma como o preferido para a exibição desse tipo de dados. “O que é um histograma?”, você deve estar se perguntando.

O histograma é, basicamente, um gráfico de barras que se aplica a dados numéricos. Como os dados são numéricos, as categorias são ordenadas da menor para a maior (ao contrário dos dados categorizados, tais como o sexo, que não possui uma ordem inerente). E, pelo fato de que você tem que assegurar que todos os números enquadram-se em apenas um grupo, as barras de um histograma se tocam, mas não se sobrepõem. Cada barra é marcada no eixo *x* (ou eixo horizontal) pelo valor que representa seu centro. Por exemplo, suponha que um histograma demonstrando a duração da vida útil de uma peça de um carro (em horas) tenha duas barras adjacentes marcadas com valores médios (centro) de 1.000 e 2.000 horas, e cada barra tem uma largura de 500 horas. Isso significa que a primeira barra representa as peças do carro que duram entre 1.500 a 2.500 horas (os números limites podem ficar em qualquer um dos lados, desde que você tenha dados consistentes para todos os valores limites).

A altura de cada barra de um histograma representa tanto o número de indivíduos em cada grupo (também conhecido com frequência de cada grupo) ou a porcentagem de indivíduos em cada grupo (também conhecida como frequência relativa de cada grupo). Por exemplo, se 50% das peças de um carro durar entre 500 e 1.500 horas, a primeira barra no exemplo anterior teria uma frequência relativa de 50% e sua altura seria o reflexo disso.

A Figura 4-14 ilustra um histograma dos dados populacionais estaduais e, por meio dele, você facilmente consegue responder às perguntas feitas no início desta seção. E, na minha opinião, os histogramas, na maioria das situações, fornecem um resumo organizacional de um conjunto de dados de maneira mais interessante que uma tabela.

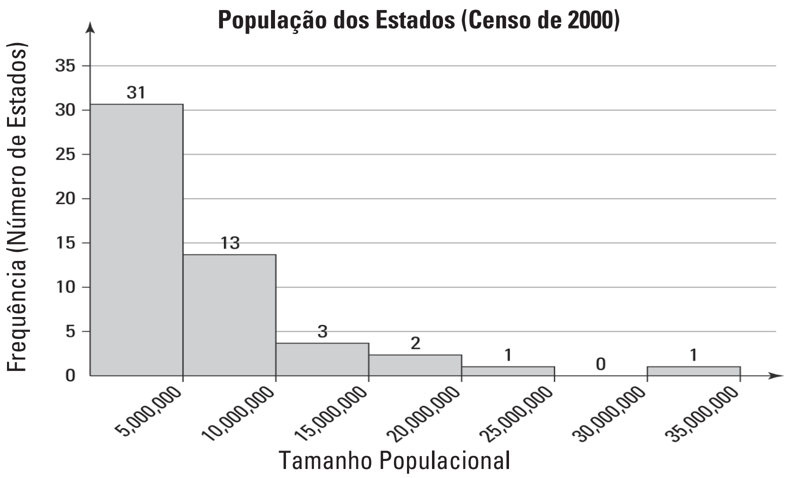


Figura 4-14: Tamanho populacional estadual (censo de 2000).

A maioria dos estados e o Distrito da Columbia (31 de 51 ou 60.8%) têm menos do que 5 milhões de habitantes. 25,5% dos estados têm um número de habitantes que varia de 5 a 10 milhões de habitantes. Isso significa que 86,3% dos estados têm menos do que 10 milhões de habitantes cada. Cada um dos sete estados restantes possui um número bem elevado de habitantes, fazendo com que o histograma pareça assimétrico e arrastando-se para a direita (chamado de distorcido para a direita). Com exceção de poucos estados, o tamanho populacional dos estados americanos não é tão variável quanto você deve ter pensado. O histograma não fala que estado é qual, é claro, mas uma rápida vasculhada nos dados originais pode revelar quais estados é o menor e qual é o maior. Os cinco estados mais populosos são California, Texas, New York, Florida e Illinois (seguido bem de perto pela Pennsylvania). O menor estado é Wyoming com cerca de 494.000 habitantes.

Se aparecer alguma dúvida enquanto você estiver examinando um gráfico ou uma tabela, tente ter acesso ao conjunto de dados originais. Os pesquisadores devem lhe fornecer esses dados, caso você os peça.

Analisando a idade materna

Em um exemplo sobre estatísticas de nascimento (ver Tabela 4-3), a idade das mães aparece para alguns anos compreendidos no período de 1975 a 2000. Para cada ano da tabela, a idade variável é dividida em grupos e nos é dado o número de mães em cada grupo. Devido ao fato de que nos são apresentados os números totais, podemos fazer um histograma da idade das mães mostrando, ou as frequências, ou as frequências relativas, dependendo do que for mais importante para o que você pretende fazer.

Suponha que você queira comparar as idades das mães em 1975 e 2000. Você pode fazer dois histogramas, um para cada ano e comparar os resultados. A Figura 4-15 mostra dois desses histogramas para 1975 (acima) e 2000 (abaixo). Observe que as frequências relativas (ou porcentagens) são exibidas no eixo vertical e as faixas etárias das mães são exibidas no eixo horizontal.

O histograma pode resumir muito bem as características dos dados numéricos. Uma das características que um histograma pode mostrar é a chamada forma dos dados (ou seja, o modo como os dados estão distribuídos entre os diferentes grupos). Estão distribuídos igualmente de um modo uniforme? São simétricos, ou seja, o lado esquerdo do histograma é o reflexo do lado direito? O histograma possui formato em U, com muitos dados nas extremidades e poucos no meio? O histograma dos dados possui uma forma sinuosa, ou seja, parecendo com uma colina no meio com extremidades tendendo em ambas as direções, conforme você se afasta do centro? Ou o histograma é distorcido, parecendo uma colina assimétrica com uma extremidade longa indo para a direita (indicando que os dados estão distorcidos para a direita) ou para a esquerda (indicando que os dados estão distorcidos para a esquerda)?

As idades das mães na Figura 4-15 para os anos de 1975 e 2000 se parecem com colinas, embora os dados de 1975 estejam ligeiramente distorcidos para a direita, indicando que, conforme a mulher envelhecia, menos filhos elas tinham, se compararmos com a situação em 2000. Outra maneira de dizer isso é que em 2000, uma maior proporção de mulheres mais velhas estava tendo filhos, quando comparado com 1975.

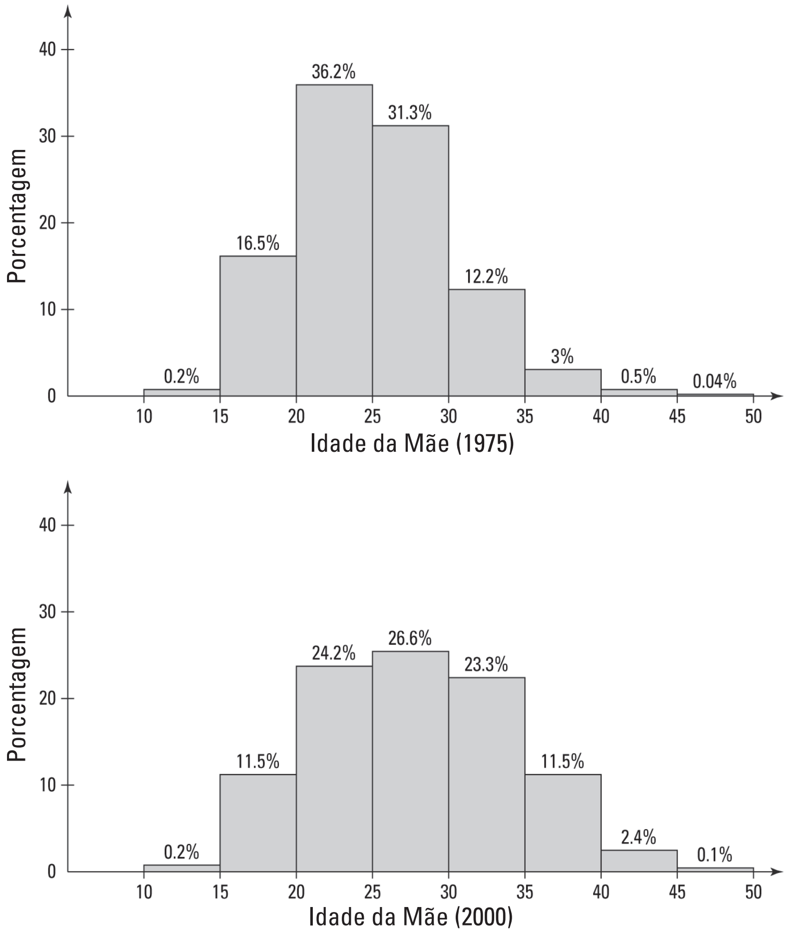


Figura 4-15: Nascimentos Vivos no Estado do Colorado segundo a idade da mãe em 1975 e em 2000.

Você também pode ter uma noção de quanta variabilidade existe nos dados ao observar um histograma. Se o histograma for razoavelmente plano, com barras de alturas próximas, você pode pensar que existe menos variabilidade, pois as alturas das barras são semelhantes. De fato, o contrário é verdadeiro. Isso porque você tem um número igual em cada barra, mas as barras propriamente ditas representam diferentes variações de valores, portanto o conjunto inteiro de dados está, na verdade, espalhado. Agora, se o histograma possui um agrupamento de dados no centro e com caudas dos lados, isso indica que mais dados estão nas barras do meio do que nas barras externas, portanto os dados estão na realidade mais agrupados. Ao compararmos as idades das mães em 1975 com as idades em 2000, você irá observar uma maior variabilidade em 2000 do que em 1975. Isso, mais uma vez, indica a mudança dos tempos: as mulheres de hoje estão esperando cada vez mais para terem filhos, quando comparadas com as mulheres de 1975 que, em sua maioria, tinham filhos até os 30 anos, e o tempo que elas estão esperando varia (o Capítulo 5 mostra as maneira para medir a variabilidade em um conjunto de dados).

A variabilidade em um histograma não deve ser confundida com a variabilidade em um gráfico de linhas (veja a seção "Alinhando-se ao gráfico de linhas). Se os valores se alteram com o passar do tempo, em um gráfico de linhas, eles são representados como altos e baixos, e muitas mudanças do alto para baixo (com o passar do tempo) indicam muita variabilidade. Portanto, uma linha estável em um gráfico de linhas não indica mudança nem variabilidade dos valores ao longo do tempo. Entretanto, quando a altura das barras de um histograma parecer estável (uniformes), elas mostram o contrário – os valores estão espalhados uniformemente ao longo dos grupos, indicando uma grande variabilidade dos dados em um ponto no tempo.

Um histograma também pode lhe dar uma ideia de onde está o centro dos dados. O centro de um conjunto de dados é medido de diferentes maneiras (veja o Capítulo 5 para uma discussão sobre essas medidas). Uma maneira de medir a olho o centro de um histograma é imaginar o histograma como se fosse alguém sentado em uma gangorra e o centro como sendo o ponto onde fica o eixo responsável pelo equilíbrio do peso de cada um dos lados. Volte à Figura 4-15, que mostra as idades das mães do Estado do Colorado em 1975 e 2000, e note que o ponto central parece estar próximo dos 25 anos para o histograma de 1975 e próximo de 27,5 anos para o histograma de 2000. Isso sugere que no ano de 2000, as mulheres do Estado do Colorado estavam tendo filhos a uma idade mais avançada, em média, do que em 1975.

Os histogramas não são encontrados na mídia com a frequência que deveriam ser. A razão disso não é clara, e as tabelas são muito mais utilizadas para mostrar as divisões dos dados numéricos. Porém, o histograma pode ser informativo, principalmente quando usado para comparar um grupo ou um período de tempo ao outro. De qualquer forma, se você quiser observar um dado de maneira gráfica, sempre será possível transformar os dados da tabela em gráficos.

Tenha cuidado com os histogramas que utilizam escalas incomuns para enganar os leitores. Assim como o que acontece com o gráfico de barras, você pode exagerar as diferenças ao utilizar um escala menor no eixo vertical de um histograma e também pode subestimar as diferenças ao utilizar uma escala maior.

Os leitores podem ser enganados por histogramas de maneiras que não são possíveis com um gráfico de barras. Lembre-se que um histograma lida com dados numéricos e não categoriais. Isso significa que você precisa determinar como quer que os dados numéricos sejam divididos em grupos, a fim de exibi-los nos eixo horizontal. O modo como esses grupos são determinados faz com que o gráfico tenha uma aparência muito diferente.

Engatinhado com um bebê

Qual a distância que um bebê de oito meses percorre engatinhando? A Figura 4-16 mostra dois histogramas que representam o mesmo conjunto de dados (distância que meu bebê engatinhou durante seis horas de teste). Em casa caso, as distancias foram arredondadas. Na porção superior da figura, os valores foram divididos em incrementos de 5 pés (1,5 metros) e os dados parecem estar distribuídos de maneira uniforme. Em outras palavras, o número de vezes que ele engatinhou cada distância (0-5 pés, 5-10, e 10-15 pés) foi, aproximadamente, o mesmo. Os dados não parecem muito interessantes. Mas, na parte de baixo da figura, dividi as distâncias em tamanhos menores de 1 pé (0,30 metros ou 30 cm) e o histograma ganhou uma aparência diferente e muito mais interessante.

Neste histograma, você consegue observar dois grupos distintos de distâncias, indicando que meu bebê tendia ou a engatinhar uma distância mais curta (cerca de 5 pés) ou uma distancia mais longa (cerca de 10 pés) para chegar aonde ele queria ir. Isso faz sentido, pois quando eu comecei a coleta dos dados, a caixa de brinquedos do meu bebê estava cerca de 5 pés do ponto de partida e a pilha de jornais (outro de seus brinquedos favoritos) estava a 10 pés de distância. O segundo histograma é uma representação muito melhor do dado: a distância que meu bebê engatinhou no cenário em questão.

Portanto, qual a distância percorrida por ele durante essas seis horas? Usando a parte inferior da Figura 4-16, é possível encontrar a distância total engatinhada, pois o valor de incremento das barras é de 1 pé (30 cm). Multiplica a altura de cada barra pela distância e depois some tudo. A distância total percorrida pelo meu bebê em um período de seis horas foi a estonteante marca de 398 pés (121,31 metros) – mais do que o comprimento de um campo de futebol americano!

Observe que eu poderia ter dividido as distâncias em tamanhos ainda menores, mas isso apenas faria com que o histograma parecesse confuso e cheio, sem lhe dar nenhuma informação adicional. Existe um meio termo em relação ao número de grupos utilizados e a variação de valores que eles representam. Cada histograma é ligeiramente diferente um do outro, mas algo entre 6 e 12 grupos é um bom número de barras para um histograma. Se o histograma tiver muito poucas barras, os dados não mostrarão nada; se o histograma tiver muitas barras, os dados ficarão muito desarticulados e os padrões se perderão.

LEMBRE-SE

Certifique-se de levar em consideração a escala do eixo horizontal e também do vertical quando for examinar os resultados apresentados no histograma. Os mesmos dados podem ser feitos para parecerem diferentes, dependendo de como eles estão agrupados (por exemplo, em poucos versus muitos grupos) e dependendo da escala do eixo vertical, que pode fazer com que as barras pareçam ser mais altas ou mais baixas do que o esperado.

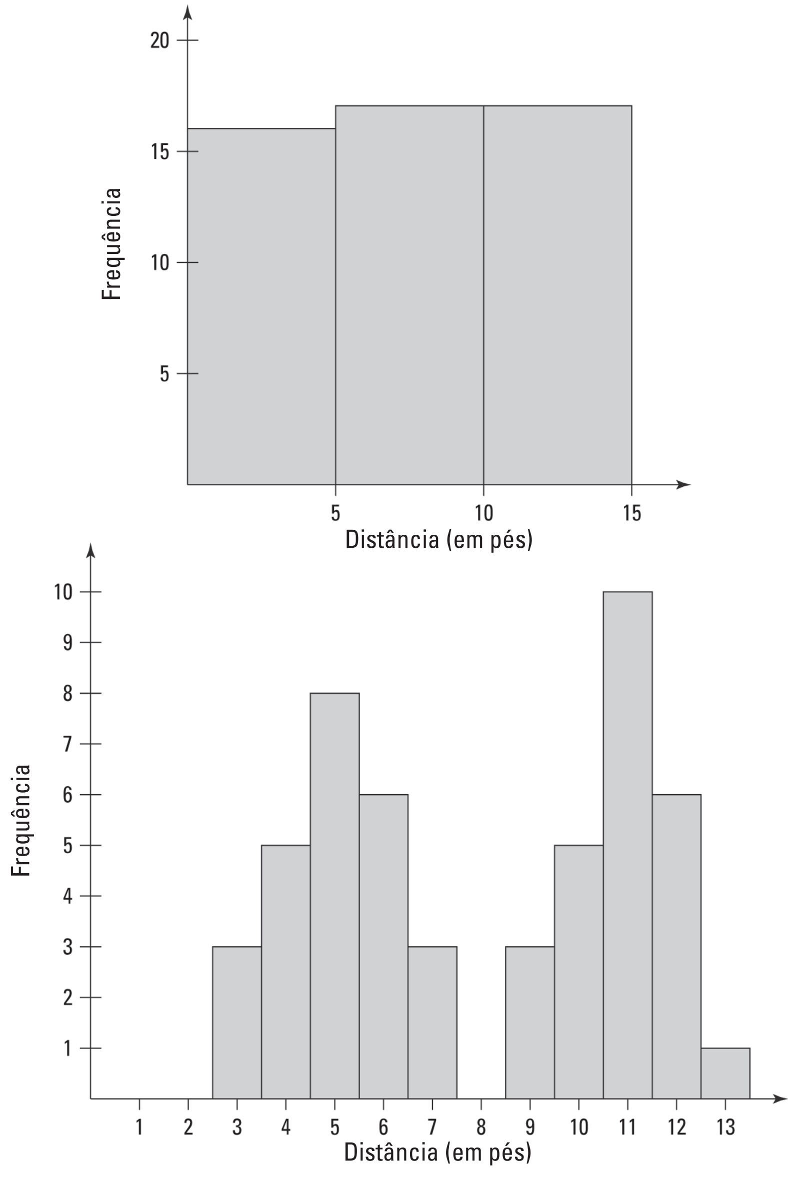


Figura 4-16: Distância engatinhada pelo bebê

Interpretando um histograma

Você pode usar um histograma para obter três características principais dos dados numéricos:

- O modo como os dados se distribuem (simétrico, distorcido para a direita, distorcido para a esquerda, formato de sino ou sinuoso e outros)
- A variabilidade encontrada nos dados
- Onde fica o centro dos dados (aproximadamente)

Avaliando um histograma

Para retratar a qualidade estatística de um histograma:

- Examine a escala utilizada para o eixo vertical (frequência ou frequência relativa) e tome cuidado com resultados que pareçam exagerados ou subestimados devido ao uso de escalas inadequadas.
- Verifique as unidades no eixo vertical para ver se o histograma relata as frequências (números) ou frequências relativas (porcentagem) e, então, leve-as em consideração ao avaliar a informação.
- Observe a escala utilizada para os grupos de variáveis numéricas (no eixo horizontal). Se a variação para cada grupo for muito pequena, os dados podem parecer mais suaves do que realmente são.

Capítulo 5

Médias, Medianas e Mais

Neste Capítulo

- Resumindo dados de maneira rápida
- Interpretando estatísticas comumente usadas
- Compreendendo o que as estatísticas dizem e o que não dizem

Uma estatística é um número que resume algumas características sobre um conjunto de dados. Das centenas de estatísticas existentes, algumas delas são utilizadas com tanta frequência, que comumente aparecem no ambiente de trabalho e em outras facetas da vida cotidiana. Neste capítulo, você irá descobrir quais estatísticas são utilizadas com mais frequência, como elas são utilizadas, o que elas significam e como elas são usadas indevidamente.

Todos os conjuntos de dados têm uma história e, se usadas de maneira apropriada, as estatísticas são a melhor maneira de contar essa história. As estatísticas usadas de maneira inadequada podem contar uma história diferente ou contar apenas um lado. Portanto, é muito importante saber como tomar boas decisões a partir das informações que lhe foram fornecidas. Neste capítulo, você verá algumas das mais comuns sínteses de dados com estatísticas e descobrirá o que essas estatísticas dizem e o que elas não dizem sobre os dados, que podem ser agrupados tanto como numéricos ou categoriais.

Resumindo Dados com Estatísticas

As estatísticas são usadas para sintetizar algumas das informações mais básicas dentro de um conjunto de dados. O resumo das informações tem vários propósitos diferentes. Imagine seu chefe vindo perguntar para você: "Qual é a nossa base de clientes atualmente e quem está comprando nossos produtos?". Como você gostaria de responder a esta questão – com um monte de números e estatísticas enormes e complicadas que com certeza lhe deixariam atordoado? Provavelmente não. Você vai querer números claros, limpos e concisos, que resuma a base de clientes, de maneira que seu chefe possa perceber o quanto você é brilhante e que, para, depois, ele possa pedir a você que colete ainda mais dados, para, assim, poder estudar um modo de incluir mais pessoas na base de clientes (isso é o que você ganha por mostrar eficiência).

Assim, as estatísticas são, geralmente, usadas para fornecer informações que sejam fáceis de compreender e que respondam às perguntas que nos são feitas (caso responder essas perguntas seja possível).

As estatísticas de síntese também têm outros propósitos. Depois que todos os dados tiverem sido coletados por meio de uma pesquisa ou de algum outro tipo de estudo, o próximo passo para os pesquisadores é tentar encontrar algum sentido nesses dados. Normalmente, a primeira coisa que os pesquisadores fazem é realizar algumas estatísticas básicas nos dados para ter uma ideia do que está acontecendo nos dados. Depois, os pesquisadores podem fazer mais análises para formular ou testar as hipóteses feitas sobre a população, estimar determinadas características a respeito da população, procurar relações entre os itens medidos e assim por diante.

Uma parte muito importante da pesquisa é o relato dos resultados, não apenas para seus companheiros, mas para a mídia e para o público em geral. Enquanto os companheiros de um pesquisador podem estar esperando ouvir todas aquelas complexas análises, feitas a partir do conjunto de dados, o público em geral, ao contrário, não está pronto para ouvir ou não está interessando em ouvir tudo isso. O que o público quer? Informação básica. Portanto, as estatísticas que conseguem deixar os assuntos claros e concisos são comumente utilizadas para transmitir tais informações à mídia e ao público.

Muitas vezes, as estatísticas são usadas para transmitir um rápido resumo de uma situação que é, na verdade, bastante complexa. Em tais situações, menos realmente não significa muito e, às vezes, a história real por detrás dos dados acaba se perdendo. Ao mesmo tempo que você precisa aceitar o fato de que, atualmente, apenas recebemos ecos das informações, certifique-se de que essa informação não esteja sendo ainda mais enfraquecida. Pense nas estatísticas que nos são passadas, no que elas realmente significam e na informação que está faltando. Este capítulo está focado nessas perguntas.

Sintetizando Dados Categorizados

Os dados categorizados coletam qualidades ou características a respeito de um indivíduo, como a cor dos olhos de uma pessoa, o sexo, o partido político ou a opinião sobre determinados assuntos (por meio do uso de categorias como concorda, discorda e não opina). Os dados categorizados têm a tendência natural de enquadrar-se em grupos ou categorias. “Partido Político”, por exemplo, nos Estados Unidos, normalmente apresenta quatro grupos: Democratas, Republicanos, Independentes e Outros. Os dados categorizados frequentemente são coletados através de uma pesquisa de opinião, mas também podem ser coletados por meio de experimentos. Por exemplo, em um teste experimental para um novo tratamento médico, os pesquisadores podem utilizar três categorias para avaliar o resultado do experimento: Os pacientes melhoraram, pioraram ou permaneceram iguais durante o tratamento?

Os dados categorizados são, frequentemente, resumidos por meio das porcentagens de indivíduos que se enquadram em cada uma das categorias definidas. Por exemplo, os entrevistadores podem relatar a porcentagem de Republicanos, Democratas, Independentes e Outros que participaram de uma pesquisa. Para calcular a porcentagem de indivíduos enquadrados em determinada categoria, encontre o número de indivíduos pertencentes à categoria desejada, divida esse número pelo total de pessoas que participaram do estudo e, depois, multiplique por 100%. Por exemplo, se uma enquete com 2.000 adolescentes tivesse incluído 1.200 meninas e 800 meninos, a porcentagem resultante seria $(1200 \div 2.000) \times 100\% = 60\%$ de meninas e $(800 \div 2000) \times 100\% = 40\%$ de meninos.

Você pode ainda dividir os dados categorizados criando algo conhecido como tabela cruzada (ou tabela de dupla entrada). As tabelas cruzadas são tabelas com linhas e colunas que resumem a informação a partir de duas variáveis categorizadas em uma, como sexo e partido político, para que você possa observar (ou facilmente calcular) a porcentagem de indivíduos em cada combinação de categorias. Por exemplo, se você tivesse dados sobre o sexo e o partido político dos respondentes, você seria capaz de ver a porcentagem de mulheres republicanas, homens republicanos, mulheres democratas, homens democratas, e assim por diante. Neste exemplo, o número total de combinações possíveis em sua tabela seria $2 \times 4 = 8$, ou seja, o número total das categorias do sexo vezes o total das categorias de afiliação partidária.

O governo americano calcula e sintetiza muitos dados categorizados utilizando as tabelas cruzadas. O U.S Census Bureau não apenas contabiliza a população como também coleta e sintetiza os dados a partir de um subconjunto de todos os cidadãos americanos (que tenham preenchido o longo formulário de pesquisa) sobre as diversas características demográficas, tais como sexo e idade. Dados típicos sobre a idade e o sexo, relatados pelo U.S Census Bureau para uma pesquisa realizada em 2001, são ilustrados na Tabela 5-1 (normalmente, a idade seria um dado numérico, mas em virtude do modo como o governo americano a relata, a idade é dividida em categorias, transformando-se em um dado categorial. Veja a seção a seguir para mais informações sobre dados numéricos).

Tabela 5-1 População Americana Dividida por Idade e Sexo (2001)

Idade	*Total*	*%*	*Masculino*	*%Masculino*	*Feminino*	*% Feminino*
Abaixo de 5 anos	19.369.341	6,80	9.905.282	7,08	9.464.059	6,53
5 a 9 anos	20.184.052	7,09	10.336.616	7,39	9.847.436	6,79
10 a 14 anos	20.881.442	7,33	10.696.244	7,65	10.185.198	7,03

Tabela 5-1 (continuação)

15 a 19 anos	20.267.154	7,12	10.423.173	7,46	9.843.981	6,79
20 a 24 anos	19.681.213	6,91	10.061.983	7,20	9.619.230	6,63
25 a 29 anos	18.926.104	6,65	9.592.895	6,86	9.333.209	6,44
30 a 34 anos	20.681.202	7,26	10.420.677	7,45	10.260.525	7,08
35 a 39 anos	22.243.146	7,81	11.104.822	7,94	11.138.324	7,68
40 a 44 anos	22.775.521	8,00	11.298.089	8,08	11.477.432	7,92
45 a 49 anos	20.768.983	7,29	10.224.864	7,31	10.544.119	7,27
50 a 54 anos	18.419.209	6,47	9.011.221	6,45	9.407.988	6,49
55 a 59 anos	14.190.116	4,98	6.865.439	4,91	7.324.677	5,05
60 a 64 anos	11.118.462	3,90	5.288.527	3,78	5.829.935	4,02
65 a 69 anos	9.523.702	3,35	4.409.658	3,15	5.123.044	3,53
70 a 74 anos	8.780.521	3,08	3.887.793	2,78	4.892.728	3,37
75 a 79 anos	7.424.947	2,61	3.057.402	2,19	4.367.545	3,01
80 a 84 anos	5.149.013	1,81	1.929.315	1,38	3.219.698	2,22
85 a 89 anos	2.887.943	1.01	926.654	0,66	1.961.289	1,35
90 a 94 anos	1.175.545	0,41	303.927	0,22	871.618	0,60
95 a 99 anos	291.844	0,10	58.667	0,04	233.177	0,16
100 ou mais	48.427	0,02	9.860	0,01	38.567	0,03
Total das idades	**284.796.887**	**100**	**139.813.108**	**100**	**144.983.779**	**100**

Você pode examinar muitas facetas diferentes da população ao observar e trabalhar com os diferentes números da Tabela 5-1. Se observarmos o sexo, notamos que o número de mulheres sobrepõe-se ligeiramente ao número de homens, pois a população em 2001 era 51% de mulheres (divida o número total de mulheres pelo número total da população e multiplique por 100%) e 49% de homens (divida o número total de homens pelo número total da população e multiplique por 100%). Você também pode observar a idade: a porcentagem de toda a população com 5 anos ou menos era de 6,8%; o maior grupo foi o das pessoas com 40-44 anos, que perfaziam 8% da população. A seguir, você também pode investigar uma possível relação entre sexo e idade ao comparar as várias partes da tabela. É possível, por exemplo, comparar a porcentagem de mulheres em comparação com a de homens com 80 ou mais anos. Pelo fato dos dados serem informados com valores de incremento de cinco em cinco anos, é necessário que você faça alguns cálculos para chegar à sua resposta. A porcentagem da população que é mulher e tem 80 anos ou mais é 2,22% + 1,35% + 0,6% + 0,16% + 0,03% = 4,36%; A porcentagem de homens com 80 anos ou mais é 1,38% + 0,66% + 0,22% + 0.04% + 0,01% = 2,31%. Isso nos mostra que a faixa etária acima de 80 anos contém quase duas vezes mais mulheres do que homens. Esses dados parecem confirmar a crença de que as mulheres tendem a viver mais tempo do que os homens.

FIQUE ATENTO

Se você tiver o número de indivíduos de cada grupo, sempre será possível calcular as porcentagens. Porém, se você tiver apenas as porcentagens sem o número total de cada grupo, nunca será possível saber o número original de indivíduos de cada grupo. Por exemplo, você pode ouvir que 80% das pessoas entrevistadas preferem mais os biscoitos Chessy cheese do que os biscoitos Crummy cheese. Mas quantas pessoas foram entrevistas? Podem ter sido apenas 10, pois 8 em 10 é 80%, assim como 800 em 1000 também é 80%. Essas duas frações (8 em 10 e 800 em 1000) têm diferentes significados para os estatísticos, pois, no primeiro caso, a estatística baseia-se em pouquíssimos dados e, no segundo caso, ela se baseia em muitos dados (leia o Capítulo 10 para saber mais sobre precisão dos dados e margem de erro).

PAPO TÉCNICO

Depois de fazer uma tabela cruzada para demonstrar a divisão de duas variáveis categoriais, você pode conduzir alguns testes estatísticos para determinar se há uma relação significativa entre as duas variáveis (veja o Capítulo 18 para mais informações sobre tais testes estatísticos).

Sintetizando Dados Numéricos

Com os dados numéricos, as características mensuráveis, tais como altura, peso, QI, idade ou renda financeira são representadas através de números. Pelo fato dos dados terem um significado numérico, você possui mais maneiras de sintetizá-los. Determinadas características de um conjunto de dados numéricos podem ser descritas usando as estatísticas, tais como a localização do centro, a maneira com os dados estão distribuídos e a localização de determinados marcos. Esses tipos de síntese aparecem com frequência na mídia, portanto saber o que essas estatísticas de síntese dizem e também o que elas não dizem lhe ajudará a entender melhor as pesquisas que lhe são apresentadas diariamente.

Chegando ao centro

O modo mais comum de sintetizar um conjunto de dados numéricos é descrever onde está o centro. Uma maneira de pensar o que o centro de um conjunto de dados significa é perguntar: "qual seria um valor normal?" ou, "onde está o meio dos dados?" Na verdade, o centro de um conjunto de dados pode ser medido de diferentes modos, e o método escolhido pode influenciar em muito as conclusões tiradas a partir dos dados.

Fazendo a média dos salários da NBA

Os jogadores da NBA ganham muito dinheiro, não é? Comparados a muita gente, eles com certeza ganham. Mas quanto eles ganham? Será mesmo que é tudo isso que você imagina? A resposta depende de como você escolhe sistetizar a informação. Você sempre ouve falar de jogadores como o Shaquille O'Neal, que ganhou $21,4 milhões nas temporadas de 2001-2002. Mas é esse o valor normal que um jogador da NBA ganha? Não, Shaquille O'Neal foi o jogador mais bem pago da NBA naquela temporada.

Portanto, quando um jogador comum da NBA ganha? Uma maneira de responder a essa pergunta é observar a média salarial. A média é, provavelmente, a estatística mais comumente utilizada de todos os tempos. Ela é uma maneira de se determinar onde está o "centro" dos dados.

Aqui está o que você precisa para encontrar a média para o conjunto de dados, denotado como $\overline{x}$.

1. **Some todos os números do conjunto de dados**
2. **Divida pelo número de números do conjunto de dados, *n*.**

Por exemplo, os dados sobre os salários dos jogadores para as temporadas de 2001-2002 são mostrados na Tabela 5-2 para os 13 jogadores do Los Angeles Lakers (excluindo os que foram liberados antes do final de cada temporada).

Tabela 5-2 Salários dos Jogadores do Los Angeles Lakers nas temporadas de 2001-2002

Jogador	*Salário ($)*
Shaquille O'Neal	$21.428.572
Kobe Bryant	$11.250.000
Robert Horry	$5.300.000
Rich Fox	$3.791.250
Lindsey Hunter	$3.425.760

Tabela 5-2 (continuação)

Derek Fisher	$3.000.000
Samaki Walker	$1.400.000
Mitch Richmond*	$1.000.000
Brian Shaw*	$963.415
Devean George	$834.250
Mark Madsen	$759.960
Jelani McCoy	$565.850
Stanislav Medvedenko	$465.850
Total	**$54.184.907**

**sem os ajustes da franquia paga pela NBA (salary cap)*

Somando todos os salários, o total da folha de pagamento desse time é de $54.184.907. Se dividirmos esse número pelo número total de jogadores (n = 13) teremos uma média salarial de $4.168.069,77. Uma ótima média salarial, não é mesmo? Mas observe que Shaquille O'Neal é o primeiro da lista e, naquela temporada, seu salário foi o mais alto de toda a liga. Se tirarmos a média salarial de todos os jogadores do Lakers sem contar o do Shaq, teremos uma média de $32.756.335 ÷ 12 = $2.729.694,58. Esse número ainda é alto, mas significativamente mais baixo do que a média salarial dos jogadores incluindo Shaquille O'Neal (é claro que os fãs vão dizer que isso apenas mostra o quanto ele é importante para o time e esse assunto é apenas a ponta do iceberg dos longos debates que os torcedores adoram ter sobre estatísticas).

Assim, para a temporada 2001-2002, a média salarial dos jogadores do Lakers foi, aproximadamente, $4,2 milhões. Mas será que a média sempre conta a história toda? Em alguns casos, a média pode ser um pouco enganosa, e esse é um desses casos. Isso porque todos os anos, alguns jogadores de elite (como Shaq) ganham muito mais do que qualquer outro (a propósito, assim como Shaq, eles, geralmente, são os mais altos). Esses são os chamados valores discrepantes (números em um conjunto de dados que são extremamente altos ou baixos quando comparados com o resto dos dados). Devido a maneira como a média é calculada, os valores discrepantes muito altos tendem a elevar a média (assim como fez o salário de Shaquille O'Neal no exemplo acima). Da mesma forma, valores discrepantes muito baixos tendem a reduzir a média.

Lembra quando você e o resto da classe se davam mal na prova enquanto uns nerds tiravam 10? Lembra como a professora não alterava a graduação da escala, a fim de mostrar o desempenho da maioria da classe? Sua professora estava provavelmente utilizando a média e, nesse caso, a média não representava realmente o verdadeiro meio das notas dos alunos.

Que outra forma que você teria, a não ser a média, para mostrar o salário de um jogador "comum" da NBA ou para mostrar a pontuação de um aluno comum da sua sala de aula? Outra estatística utilizada para medir o meio de um conjunto de dados é a mediana. A mediana ainda é não tão usada como deveria ser, embora, atualmente, seu uso esteja aumentando.

Reduzindo os salários à mediana

A mediana de um conjunto de dados é o valor que se encontra exatamente no meio. Aqui vão alguns passos para você encontrar a mediana de um conjunto de dados:

1. **Ordene os números do menor para o maior.**
2. **Se o conjunto de dados possui um número impar de números, escolha o número que estiver exatamente no meio.**

   Este é a mediana.
3. **Se o conjunto de dados possui um número par de números, pegue os dois números que estiverem exatamente no meio e faça a média deles para encontrar a mediana.**

Os salários da equipe Los Angeles Lakers durante a temporada de 2001-2002 (volte à tabela 5-2) já estão ordenados do menor para o maior (de baixo para cima). Já que a lista contém os nomes e os salários de 13 jogadores, o salário mediano é o sétimo de baixo para cima ou vice-versa, ou seja, é o salário do jogador Samaki Walker, que ganhou $1,4milhões naquela temporada jogando pelos Lakers. Esta é a mediana.

Este salário mediano para os Lakers está bem abaixo da média de $4,2 milhões para o time. Porém, pelo fato de que a média salarial dos Lakers incluía valores discrepantes (como o salário de O'Neal), o salário mediano representa melhor o meio dos salários do time (observe que apenas 3 jogadores ganharam mais do que a média salarial de $4,2milhões, enquanto 6 jogadores ganharam mais do que o salário mediano de $1,4 milhões). Diferente do que acontece com a média, a mediana não é influenciada pelos salários dos jogadores que estavam na extremidade mais alta (a propósito, o salário mais baixo do Lakers para temporada 2001-2002 foi de $ 465.850 – muito dinheiro para os padrões da maioria das pessoas, mas uma merreca quando, comparado ao que você imaginava ser o salário de um jogador da NBA!).

O Governo americano usa com frequência a mediana para representar o meio referente a seus dados. Por exemplo, o U.S. Census Bureau registrou que em 2001 a renda familiar mediana era de $ 42.228, caindo 2,2% com relação ao ano 2000, quando a renda familiar mediana era de $43.162.

Interpretando o centro: Comparando médias e medianas

Agora, suponha que você faz parte de um time da NBA tentando negociar salários. Se você representar os proprietários, vai querer mostrar quanto cada um está ganhando e quanto você está gastando, portanto você deve levar em consideração os jogadores estrelas do time e mostrar a média. Mas, se você estiver do lado dos jogadores, vai querer mostrar a mediana, pois ela representa melhor o que os jogadores do meio ganham. Cinquenta por cento dos jogadores ganham acima da mediana e cinquenta por cento ganham abaixo dela. É por isso que ela é chamada de mediana – representa o ponto que está exatamente no meio.

PAPO TÉCNICO

O histograma é um tipo de gráfico que organiza e exibe os dados numéricos de uma forma ilustrada, mostrando os grupos de dados e o número ou porcentagens dos dados que se encontram em cada grupo (leia o Capítulo 4 para mais informações sobre histogramas e outros modos de exibição de dados). Caso os dados apresentem valores discrepantes na extremidade superior, o histograma ficará distorcido para a direita e a média será maior do que a mediana (veja a parte superior do histograma da Figura 5-1 para um exemplo de dados distorcidos para a direita). Se os dados apresentarem valores discrepantes na extremidade inferior, o histograma dos dados ficará distorcido para a esquerda e a média será menor do que a mediana (o histograma do meio na Figura 5-1 mostra um exemplo de histograma com dados distorcidos para a esquerda). Se os dados são simétricos (tendo, aproximadamente, a mesma forma dos dois lados), a média e a mediana terão, aproximadamente, o mesmo valor (a parte inferior da Figura 5-1 mostra um exemplo de dados simétricos retratados em um histograma).

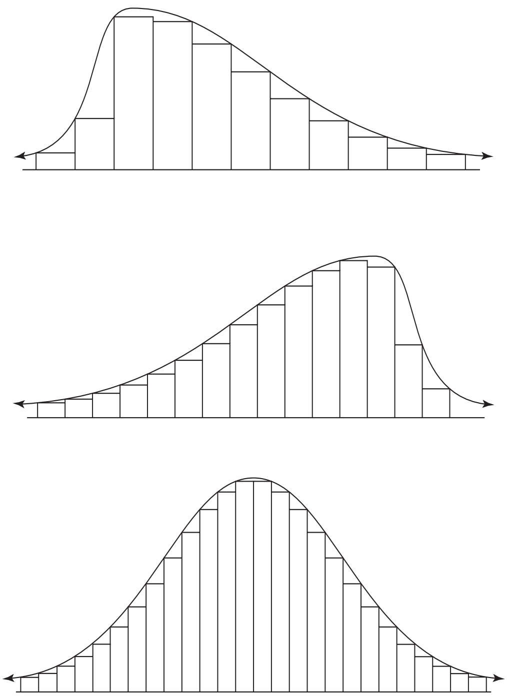


Figura 5-1: Dados distorcidos para a direita, para a esquerda e dados simétricos

LEMBRE-SE

A média de um conjunto de dados é influenciada por valores discrepantes, mas isso não acontece com a mediana. Se alguém lhe informar o valor médio, pergunte também pela mediana, para que assim você possa comparar as duas estatísticas e entender melhor o que realmente está se passando no gráfico e o que é realmente normal.

Contabilizando a variação

Sempre existe variação em um conjunto de dados, independente da característica que você esteja medindo, pois nem todos os indivíduos terão o mesmo exato valor para todas as variáveis. A variabilidade é o que faz a estatística ser como é. Por exemplo, os preços de casas variam de casa para casa, de ano a ano e de estado para estado. A renda familiar também varia de família para família, de país para país e de ano a ano. O número de jardas alcançadas por um quarterback em um jogo varia de jogador para jogador, de jogo para jogo e de temporada para temporada. O tempo que você leva para chegar ao trabalho varia dia a dia. O truque para lidar com a variação é o de ser capaz de medir a variabilidade da melhor forma possível.

Entendendo o significado do desvio padrão

O desvio padrão é de longe a medida mais utilizada para a variabilidade. O desvio padrão representa a distância normal de qualquer ponto no conjunto de dados até o centro. Ele é, aproximadamente, a distância média do centro e, neste caso, o centro é a média. Na maioria das vezes, você não ouve falar do desvio padrão propriamente dito, quando ele é informado (o que não ocorre com muita frequência), provavelmente está naquelas letrinhas miúdas, geralmente em parênteses como "($s = 2{,}68$)".

PAPO TÉCNICO

O desvio padrão de uma população inteira de dados é denotado pela letra grega σ. O desvio padrão de uma amostra da população é denotado pela letra s. Já que, na maioria das vezes, o desvio padrão da população não é um valor conhecido, quaisquer fórmulas envolvendo o desvio padrão lhe deixariam na mão sem nada para colocar em seu lugar. Mas não tema. Quando estiver em Roma, faça como os romanos, certo? Então, quando você tiver que lidar com estatística, faça como os estatísticos – sempre que eles se veem sem saída por causa de um valor desconhecido, eles apenas o estimam e seguem em frente! Assim, *s* é utilizado para estimar nos casos em que é desconhecido.

Neste livro, quando utilizo o termo desvio padrão, refiro-me a *s*, ou seja, o desvio padrão da amostra (caso me refira ao desvio padrão da população, eu lhe avisarei!).

Calculando o desvio padrão

A fórmula para o desvio padrão é $s = \sqrt{\dfrac{\sum (x - \overline{x})^2}{n-1}}$.

Para calcular o desvio padrão amostral, *s*, siga os passos abaixo:

1. **Encontre a média do conjunto de dados.**

   Para encontrar a média, some todos os números e divida o resultado pela quantidade de números contidos pelo conjunto de dados, *n*.

2. **Subtraia o valor da média para cada um dos números.**
3. **Eleve ao quadrado cada resultado obtido no passo anterior.**
4. **Some todos os resultados obtidos no passo 3.**
5. **Divida a soma encontrada no passo 4 pela quantidade de números no conjunto de dados, menos 1 (n -1).**
6. **Descubra a raiz quadrada do número resultante.**

PAPO TÉCNICO

Na fórmula, dividimos por n -1 ao invés de n para que assim o desvio padrão da amostra tenha propriedades que funcionem com todas as teorias (acredite em mim, é mais do que você quer saber sobre esse assunto). Por exemplo, ao dividir por $n - 1$, garantimos que o desvio padrão médio não é tendencioso. Caso você ainda não esteja tão confuso, aqui vai mais: se você conseguir os dados de uma população inteira e quiser encontrar o desvio padrão dessa população, σ, utilize a mesma fórmula usada para s, mas divida por n e não por $n - 1$!

Veja este pequeno exemplo. Suponha que você tenha quatro números: 1, 3, 5 e 7. A média é $16 \div 4 = 4$. Subtraindo a média de cada número, você tem $(1-4) = -3$, $(3-4) = -1$, $(5-4) = +1$ e $(7-4) = +3$. Elevando os resultados obtidos ao quadrado, temos 9,1,1 e 9. A soma desses números é 20. Neste exemplo, $n = 4$ e, portanto, $n - 1 = 3$, assim, você divide 20 por 3 e chega no valor 6.67. Por fim, descubra a raiz quadrada de 6,67, que é 2,58, e esse é o desvio padrão para esse conjunto de dados. Portanto, para o conjunto de dados 1, 3, 5, 7, a distância normal dos dados com relação ao centro do conjunto de dados é 2,58.

LEMBRE-SE

Já que o cálculo do desvio padrão envolve muitos passos, na maioria dos casos, você provavelmente precisará que um computador calcule-o para você. Mas, saber como calcular o desvio padrão lhe ajudará a interpretar melhor esta estatística e também poderá lhe ajudar a perceber quando a estatística estiver errada.

Interpretando o desvio padrão

É difícil interpretar o desvio padrão como sendo um número sozinho. Um desvio padrão pequeno, basicamente, significa que os valores do conjunto de dados estão, na média, próximos do centro desse conjunto, enquanto um desvio padrão grande significa que os valores do conjunto de dados estão, na média, mais afastados do centro.

Um desvio padrão pequeno pode ser um bom objetivo em determinadas situações, onde os resultados são restritos (por exemplo, na produção e no controle de qualidade de uma indústria). Uma determinada peça de carro que deve ter centímetros de diâmetro para encaixar perfeitamente não pode apresentar um desvio padrão muito grande. Um desvio padrão grande, nesse caso, significaria que acabariam sendo jogadas fora, pois ou não se encaixariam adequadamente ou os carros teriam problemas.

Em situações em que você apenas observa e registra os dados, um desvio padrão grande não é necessariamente algo ruim; ele apenas reflete uma grande variabilidade dentro do grupo que está sendo estudado. Por exemplo, se você observar os salários de todos os funcionários em uma empresa, incluindo todos desde o estagiário até o diretor, o desvio padrão poderá ser muito grande. Por outro lado, se você estreitar o grupo e observar somente os estagiários ou os executivos, o desvio padrão será menor, pois os indivíduos dentro de cada um desses dois grupos têm salários que variam menos.

Observe as unidades no momento em que for determinar se um desvio padrão é grande. Por exemplo, um desvio padrão de 2 em unidades de anos é equivalente a um desvio padrão de 24 em unidades de meses. Além disso, também observe o valor da média quando for colocar o desvio padrão em perspectiva. Se o número médio de comunidades virtuais em que um usuário participa é de 5,2 e o desvio padrão é de 3,4, há muita variabilidade se falarmos de um modo geral. Mas se levarmos em consideração a idade dos usuários dessas comunidades, onde a média é 25,6 anos, o desvio padrão de 3,4 seria comparativamente menor.

Outra maneira de interpretar o desvio padrão é utilizando-o em conjunto com a média para descrever a posição da maior parte dos dados. Se os dados estiverem distribuídos em uma curva com formato de sino (com muitos dados próximos do centro e com cada vez menos valores conforme você se afasta do centro), é possível usar algo chamado de regra empírica para interpretar o desvio padrão (veja o Capítulo 4). A regra empírica diz que cerca de 68% dos dados deveriam estar dentro de um desvio padrão da média; cerca de 95% dos dados deveriam estar dentro dois desvios padrões da média e cerca de 99% dos dados deveriam estar dentro de três desvios padrões da média.

Em um estudo sobre como as pessoas fazem amigos virtuais utilizando comunidades na internet, por exemplo, a média de idade dos usuários desses grupos foi informada como sendo de 31,65 anos, com um desvio padrão de 8,61 anos. Os dados estavam distribuídos em uma curva com formato de sino. De acordo com a regra empírica, cerca de 68% dos usuários dessa comunidade tinham idades dentro de 1 desvio padrão (8,61 anos) da média (31,65 anos). Portanto, cerca de 68% dos usuários tinham idades entre 31,65 – 8,61 anos e 31,65 + 8,61 anos, ou seja, entre 23,04 e 40,26 anos. Cerca de 95% do usuários tinham idades entre 31,65 – 2(8,61) e 31,65 + 2(8,61) ou entre 14,43 e 48,87 anos. Por fim, cerca de 99% das idades dos usuários dessa comunidade eram de 31,65 – 3(8,61) e 31,65 + 3(8,61), ou seja, entre 5,82 e 57,48 anos (para mais aplicações da regra empírica, veja o Capítulo 8).

A maioria das pessoas não se importa em tentar contabilizar 99% dos valores de um conjunto de dados, geralmente, elas se satisfazem com 95%. Contabilizar um desvio padrão para cada um dos lados da média simplesmente para abranger 4% a mais dos dados (99% - 95%), para muitas pessoas, não parece valer a pena.

Entendendo as propriedades do desvio padrão

Aqui estão algumas propriedades que podem lhe ajudar a interpretar um desvio padrão:

- O desvio padrão nunca pode ser um número negativo (graças ao modo como ele é calculado e pelo fato de que ele mede uma distância; e distâncias nunca são números negativos).
- O menor valor possível para o desvio padrão é 0, e isso acontece somente em situações planejadas onde cada número do conjunto de dados é igual (sem desvios).
- O desvio padrão é influenciado por valores discrepantes (números extremamente baixos ou extremamente altos dentro de um conjunto de dados). Isso porque o desvio padrão baseia-se na distância dos dados com relação à média.

O desvio padrão tem a mesma unidade dos dados originais.

Fazendo propaganda do desvio padrão

O desvio padrão não aparece muito na mídia e isso é um grande problema. Se você descobrir onde fica o centro dos dados sem saber o quanto tais dados variam, você tem apenas uma parte da história. De fato, você pode estar perdendo a parte mais interessante da história. A variedade é o tempero da vida, assim, sem uma indicação de quanto diversificado e variado são os dados, você não saberá o verdadeiro tempero dos dados.

Sem saber o desvio padrão, você não consegue saber se todos os dados estão próximos da média (assim como os diâmetros das peças de carros que apresentam defeitos mesmo quando tudo está funcionando corretamente) ou se os dados estão muito espalhados (assim como os salários dos jogadores da NBA). Se alguém lhe disser que a média dos salários iniciais para alguém que trabalha na Statistix & CIA é de $ 70.000, você pode pensar: "Nossa, que maravilha!". Mas, e se o desvio padrão para os salários iniciais na Statistix & CIA for de $20.000 e se utilizarmos a regra empírica, supondo que a distribuição dos salários tem um formato de sino, você poderia ganhar qualquer valor entre $30.000 a $ 110.000 (ou seja, $70.000 mais ou menos dois desvios padrões iguais a $20.000 cada). A Statistix & CIA apresenta uma grande variedade em termos de valores pagos a seus funcionários, portanto, o salário inicial médio de $70.000, no final das contas, não diz muito, não é mesmo? Por outro lado, se o desvio padrão fosse apenas de $5.000, você poderia ter uma ideia muito melhor do que esperar de um salário inicial nessa empresa.

LEMBRE-SE

Sem o desvio padrão não é possível comparar dois conjuntos de dados de maneira eficiente. E se os dois conjuntos de dados tiverem, aproximadamente, a mesma média e a mesma mediana; isso significa que os dados serão todos os mesmos? Não mesmo. Por exemplo, os conjuntos de dados 199, 200, 201 e 0, 200, 400 têm a mesma média, 200, e a mesma mediana, 200. No entanto, eles possuem desvios padrões muito diferentes. O primeiro conjunto de dados tem um desvio padrão muito pequeno comparado ao segundo conjunto de dados.

Jornalistas normalmente não informam o desvio padrão. A única razão para isso, a meu ver, é que as pessoas não perguntam por ele – talvez o público ainda não esteja pronto para o desvio padrão. Mas as referências ao desvio padrão podem se tornar mais comuns na mídia se cada vez mais pessoas começarem a descobrir o que um desvio padrão pode lhe revelar a respeito de um conjunto de resultados. E, em muitos ambientes de trabalho, o desvio padrão é, frequentemente, usado e informado, pois essa estatística é um padrão e também um modo muito bem aceito de medir a variação.

Abaixo à amplitude

Muitas vezes, a mídia relata a amplitude de um conjunto de dados como sendo uma maneira de medir a variabilidade. A amplitude é o maior valor de um conjunto de dados menos o menor valor desse conjunto. É muito fácil encontrar a amplitude; tudo o que você precisa fazer é colocar os números em ordem (do menor para o maior) e fazer uma rápida subtração. Talvez por isso a amplitude seja tão utilizada; mas não por seu valor interpretativo.

A amplitude de um conjunto de dados é quase inútil. Ela depende apenas de dois números de um conjunto de dados, que podem refletir valores extremos (valores discrepantes). Meu conselho é ignorar a amplitude e encontrar o desvio padrão, a forma mais eficaz de medir a variabilidade presente em um conjunto de dados.

Os salários da NBA, como esperado, têm uma grande variabilidade. Os salários para um único time, os Los Angeles Lakers, para a temporada de 2001 – 2002 são um exemplo típico. Veja a Tabela 5-2 para os salários dos 13 jogadores do time. O salário médio é $4.168.069,77 e a mediana é $ 1.400.000. Os salários variam do mais alto, que é $21.428.572 (Shaquille O'Neal), ao mais baixo, $465.850 (Stanislav Medvedenko), com uma amplitude de $21.428.572 - $ 465.850 = $20.962.722. Nossa – é uma amplitude e tanto! Isso, certamente, mostra uma grande diferença entre o jogador com o salário mais alto e o jogador com o salário mais baixo. Mas isso significa muito em termos da variabilidade total entre os salários de todo o time? Na verdade não. O desvio padrão se baseia em todos os salários do time (não apenas os maiores e os menores) o desvio padrão tem muito mais significado estatístico do que a amplitude.

Quando você se deparar com estatísticas de síntese, procure o desvio padrão para saber qual é a variação dos dados. Se essa informação não estiver disponível, pergunte por ela ou vá até a fonte (a comunicação de imprensa, o artigo de jornal ou os próprios pesquisadores), onde você possa realmente encontrá-la. Não confie muito na amplitude; ela é uma estimativa muito aproximada da variabilidade para ser levada em consideração.

Determinando sua posição: Percentil

Todos querem saber como estão em relação aos outros. Na escola, a nota que você consegue em uma prova não é tão importante quanto o como essa nota se compara com a das outras crianças da sua classe. Exames como o GRE e o ACT (exames necessários para o ingresso nas universidades americanas) geralmente mantêm o mesmo total de pontos todos os anos, enquanto o desempenho dos alunos varia conforme as provas mudam de um ano para o outro. Portanto, juntamente com suas notas, você sempre recebe um relatório com a média de sua pontuação em relação a dos outros. Em outras palavras, você descobre sua posição relativa dentro do grupo.

Entendendo o percentil

A maneira mais comum de informar a posição relativa é por meio do uso do percentil. O percentil é a porcentagem de indivíduos dentro de um conjunto de dados que estão abaixo de você. Se você estiver no percentil 90º, por exemplo, isso significa que 90% das pessoas que fizeram o exame com você conseguiram menos pontos do que você. E, também, significa que 10% conseguiram mais pontos do que você, uma vez que a soma do total dever ser igual a 100% (todos que fizeram o exame têm que aparecer em algum lugar com relação a você, certo?).

O percentil não representa a sua pontuação propriamente dita. Suponha que sua pontuação no GRE tenha ficado no percentil 80º. Isso não significa que você tenha acertado 80% da prova, mas sim que 80% dos alunos que também fizeram a prova pontuaram menos do que você e 20% dos alunos pontuaram mais do que você.

Calculando o percentil

Para calcular o *k*º percentil (onde *k* é qualquer números entre um e cem), siga os passo abaixo:

1. **Ordene todos os números do conjunto de dados do menor para o maior.**

2. **Multiplique a porcentagem, *k*, pela quantidade total de números, *n*.**

3. **Arredonde o resultado para o número inteiro mais próximo.**

4. **Conte os números da esquerda para a direita (do menor para o maior) até encontrar o valor calculado no terceiro passo.**

Por exemplo, suponha que você tenha as pontuações de 25 testes e quando ordenados do menor para o maior, ficam desta forma: 43, 54, 56, 61, 62, 66, 68, 69, 69, 70, 71, 72, 77, 78, 79, 85, 87, 88, 89, 93, 95, 96, 98, 99, 99. Suponha ainda que você queira encontrar o percentil 90º para essas pontuações. Já que os dados já estão ordenados, o próximo passo é multiplicar 90% pela quantidade total de pontuações, ou seja, 90% × 25 =

0,90 × 25 = 22,5. Arredondando o resultado obtido para o número inteiro mais próximo, chegamos ao número 23. Isso significa que você deve contar da esquerda para a direita (do menor para o maior número do conjunto de dados) até encontrar o 23º número do conjunto de dados. No exemplo acima, o 23° é o número 98, que, nesse caso, será o 90° percentil.

O percentil 50° é o ponto no conjunto de dados onde 50% dos dados encontram-se abaixo dele e 50% se encontram acima dele. Você talvez o reconheça por outro nome – a mediana. Na realidade, a mediana é um percentil especial; ela é o percentil 50°.

Um percentil alto nem sempre significa algo positivo. Por exemplo, se sua cidade está no percentil 90° em termos de taxa de criminalidade, quando comparada com outras cidades do mesmo tamanho, isso significa que 90% das cidades semelhantes à sua apresentam uma taxa de criminalidade mais baixa do que a da sua cidade, o que não é uma boa notícia para você.

Interpretando o percentil

O governo americano sempre usa o percentil ao resumir seus dados. Por exemplo, o U.S Census Bureau informou que a renda familiar mediana em 2001 era de $ 42.228. O Bureau também informou vários percentis para a renda familiar, incluindo o 10°, 20°, 50°, 80°, 90° e 95°. A Tabela 5-3 mostra os valores para cada um desses percentis.

Tabela 5-3 Renda Familiar nos Estados Unidos em 2001

Percentil	*Renda Familiar em 2001*
10°	$ 10.913
20°	$ 17.970
50°	$42.228
80°	$83.500
90°	$116.105
95°	$150.499

Ao observar esses percentis é possível verificar que as rendas mais baixas não são tão diferentes quanto as rendas mais altas. A diferença entre o 50° percentil e o 20° percentil é, aproximadamente, $25.000, enquanto a diferença entre o 50° percentil e o 80° percentil é mais de $41.000. E a diferença entre o 10° percentil e o 50° percentil é apenas de cerca de $31.000, enquanto a diferença entre os percentis 90° e 50° é de $74.000.

PAPO TÉCNICO

Por meio desses percentis e da maneira com eles se distribuem entre os dados, você pode dizer que esse conjunto de dados, se fosse retratado por um histograma, seria distorcido para a direita (o histograma é basicamente um gráfico de barras que divide os dados em grupos e mostra o número de indivíduos em cada grupo. Veja o Capítulo 4 para mais informações sobre histogramas). Isso porque as rendas mais altas estão mais espalhadas e mais fora do trilho do que as rendas mais baixas, que se apresentam mais agrupadas. Nesse informe, a média não foi mostrada, pois ela seria muito influenciada por esses valores discrepantes (as famílias com rendas muito altas), que elevariam muito a média, inflando artificialmente a descrição geral das rendas familiares nos Estados Unidos.

Os percentis realmente aparecem com frequência na mídia e em muitos documentos públicos; eles podem transmitir informações interessantes acerca de dados, incluindo a regularidade ou irregularidade de sua distribuição, a simetria e alguns marcos importantes como, por exemplo, a mediana. Os percentis também podem lhe dizer onde você (sua pontuação em um teste, sua renda familiar e etc.) encontra-se dentro de um conjunto de dados. Às vezes, o valor da média não é importante, desde que você saiba a que distância (acima ou abaixo) da média você está. Para mais informações sobre essa e outras aplicações dos percentis, veja o Capítulo 8.

LEMBRE-SE

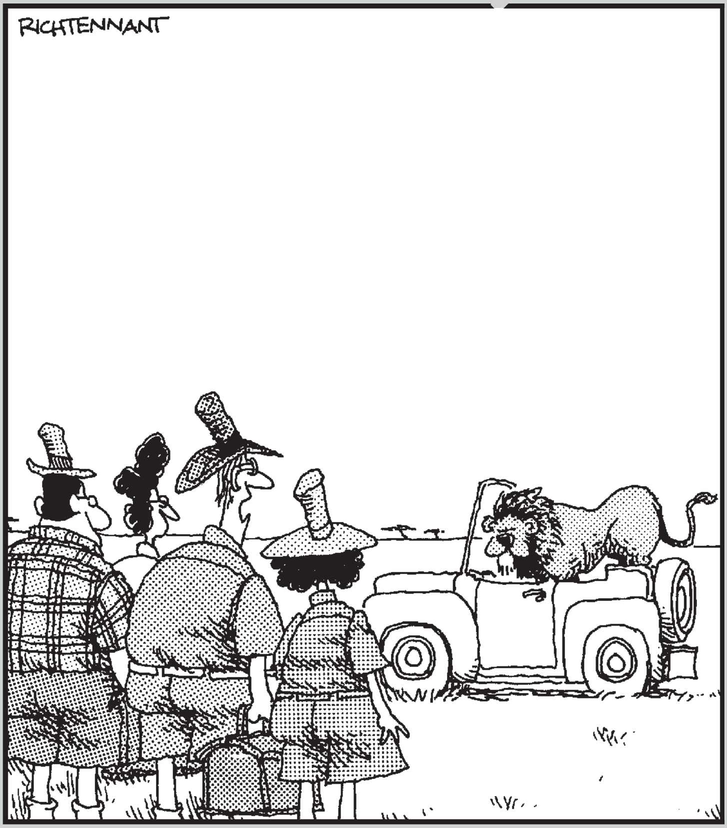
Independente do tipo de dado e do tipo de estatística que esteja sendo resumido, lembre-se de que a estatística de síntese não lhe diz tudo sobre os dados. Porém, se essas estatísticas forem bem escolhidas e não forem enganosas, podem lhe dizer muito e de forma rápida. Erros de omissão podem ocorrer, portanto fique atento a essas estatísticas menos conhecidas, que podem lhe dar importantes pistas para a compreensão da história real por trás dos dados.

Parte III

Determinando as Probabilidades

A 5ª Onda

por Rich Tennant



"Ok – vamos usar as probabilidades estatísticas dessa situação. Somos 4 e ele é apenas um. Philip provavelmente irá começar a gritar, Nora provavelmente irá desmaiar, você provavelmente irá me xingar por ter deixado o jipe aberto e há uma boa probabilidade de que eu corra como frangote se ele vier em nossa direção".

Nesta Parte...

Preparem seus dados! Nesta parte, você irá descobrir alguns segredos das mesas de jogos (e a regra número um é sair enquanto você estiver em vantagem!). Você também verá os fundamentos da probabilidade e, sendo assim, entender o que está contra você em uma mesa de jogo ou em qualquer outro tipo de situação que envolva possibilidades e incertezas. E talvez você se surpreenda ao descobrir que probabilidade e intuição nunca se misturam!

Capítulo 6

Quais são as Chances? Entendendo Probabilidade

Neste Capítulo

- Usando a probabilidade em sua vida pessoal e no ambiente de trabalho
- Entendendo com a probabilidade funciona
- Percebendo como a probabilidade vai contra sua intuição
- Fazendo a relação entre probabilidade e estatística

Neste capítulo você vai descobrir com a probabilidade é utilizada em sua vida pessoal e profissional diariamente e investigará algumas das regras da probabilidade. Também verá como a probabilidade e a intuição nunca se misturam, descobrirá maneiras de evitar alguns equívocos sobre a probabilidade e saberá o que a probabilidade tem a ver com a estatística.

Tendo uma Chance com a Probabilidade

Você alguma vez já disse: "Quais são as chances de isso acontecer?". Você lê, por exemplo, sobre dois tornados que devastaram a mesma cidadezinha do Kansas dentro de um período de 50 anos, revê um amigo no avião com quem você não conversava há anos, dois pneus do seu carro furam no mesmo dia, seu time, que era o lanterninha, acaba ganhando o campeonato. Coisas estranhas acontecem e, às vezes, esses eventos deixam-lhe com as dúvidas: "Quais são as probabilidades? Quem poderia ter previsto isso? Quais são as chances de que esse evento volte a acontecer?". Todas essas perguntas têm a ver com probabilidade.

Mas probabilidade não se trata apenas de analisar as esquisitices da vida (embora isso seja realmente um passatempo muito divertido para quem gosta de probabilidade). A probabilidade trata, na verdade, de lidar com o desconhecido de uma maneira sistemática, por meio da identificação das possibilidades, da descoberta das situações mais prováveis ou por meio de um plano B, caso as situações mais prováveis não ocorram.

A vida é uma sequência de eventos imprevisíveis, mas a probabilidade pode ser utilizada para ajudar na previsão da possibilidade de que certos eventos venham a ocorrer. Aqui vão alguns dos modos mais mundanos com os quais a probabilidade pode cruzar seu caminho diariamente:

- A previsão do tempo diz que há 80% de chances de chover hoje, portanto, você acaba levando seu guarda-chuva para o trabalho.
- Você sabe por experiência que se ficar um pouquinho acima do limite de velocidade, suas chances de passar por mais faróis abertos de uma vez aumentam (desde que você não leve uma multa por fazer isso).
- Em seu caminho para o trabalho, você começa a se perguntar se seu assistente vai lhe telefonar dizendo que está doente, pois hoje é sexta-feira e 75% das vezes que ele fica doente são na sexta-feira (você também pode ponderar as chances de que ele lhe telefone dizendo que encontrou outro emprego, acontecimento com chances muito menores de ocorrer, você supõe).
- Você compra um bilhete de loteria na hora do almoço, pois pensa "Alguém tem que ganhar e esse alguém pode muito bem ser eu!" (a propósito, suas chances de ganhar a mega-sena são de 1 em 50.063.860, portanto não adianta cruzar os dedos).
- Na TV, você fica sabendo da mais recente descoberta médica, que diz que se você tirar um cochilo revigorante durante o dia, você reduzirá suas chances de insônia em 35% (mas você não consegue terminar de ver a reportagem, pois já pegou no sono).
- Você termina o dia assistindo à vitória do seu time do coração e sonha com as chances de que ele vença o campeonato.

A probabilidade também é usada em quase todos os ambientes de trabalho, de empresas de marketing a firmas de investimento, de agências governamentais a indústrias e de hospitais a restaurantes. A lista a seguir inclui apenas alguns dos muitos exemplos de como a probabilidade é utilizada no trabalho:

- Uma pequena empresa realiza uma pesquisa de mercado para saber se os consumidores gostam o bastante de um produto, para que a empresa possa oferecê-lo em uma rede de lojas. Se a empresa acertar, ela pode ganhar rios de dinheiro, mas se errar, a empresa pode ir para o buraco.
- Uma empresa que faz batatas chips tem que garantir que as embalagens estejam sendo enchidas de acordo com as especificações: muito poucas batatas e eles terão problemas por representar incorretamente seu produto; batatas demais e eles terão prejuízo. Assim, eles coletam amostras de embalagens e, baseando-se nelas, descobrem a probabilidade de que algo de errado está acontecendo com as máquinas.
- O senhor Sou Otimista da Silva decide investir na ideia de candidatar-se a governador, mas antes de passar pela dura tarefa de arrecadar milhões para sua campanha, ele realiza uma enquete para determinar suas chances de ganhar a eleição.
- Uma empresa farmacêutica tem um novo medicamento para a hipertensão. Baseada em testes clínicos com voluntários, a empresa

determina as probabilidades para a melhora da condição e para o desenvolvimento de determinados efeitos colaterais nas pessoas que utilizaram o medicamento.

- A engenharia genética utiliza a probabilidade para prever os padrões e resultados genéticos em uma variedade de áreas que vão desde o desenvolvimento de novas plantas até a identificação de doenças hereditárias antes que elas ocorram.

- Um gerente de restaurante pensa em probabilidade quando necessita saber quantos clientes entrarão em seu restaurante em um determinado horário para, assim, deixar tudo preparado.

- Um corretor de ações usa a probabilidade todos os dias para tomar decisões. Ele constantemente se pergunta se o valor de determinada ação irá subir ou cair, se ele deve comprar ou vender e o que ele deve dizer a seus clientes.

Ganhando território: Fundamentos da Probabilidade

A probabilidade está em todo lugar, ainda que, às vezes, possa ser difícil entendê-la, pois ela pode parecer estar contra as expectativas. O primeiro passo para ganhar o território da probabilidade é entender algumas de suas regras básicas e como essas regras são aplicadas. Quando os estatísticos falam de probabilidade, eles se referem à probabilidade de um resultado, que é um determinado resultado de um processo aleatório que está sendo estudado. Mas você deve estar se perguntado: “o que é um processo aleatório?” É qualquer processo para o qual o resultado não é fixo, podendo variar aleatoriamente. Por exemplo, se você jogar um dado de seis lados uma vez, o resultado (o número que sair na face superior do dado) será um dos seis possíveis números: 1, 2, 3, 4, 5 ou 6.

Entendendo as regras

DICA

Considere as seguintes regras básicas da probabilidade:

- A probabilidade de um resultado é a porcentagem de vezes que se espera obter esse resultado. Isso pode ser calculado, dividindo o número de maneiras que o resultado pode ocorrer pelo número total de possíveis resultados. Por exemplo, a probabilidade do número 1 aparecer quando um dado é lançado é de 1 em 6, ou seja, 1/6 (ou 16,7%).

- Toda probabilidade é um número (uma porcentagem) entre 0% e 100% (observe que os estatísticos, com frequência, representam as porcentagens como proporções – números entre 0 e 1). Se um resultado tem uma probabilidade de 0%, ele nunca vai acontecer, independente de qualquer coisa. Se um resultado tem uma probabilidade de 100%, ele sempre irá acontecer, independente de qualquer coisa. A maioria das probabilidades não é de 0% nem de 100%, mas se enquadram em algum lugar entre esses dois números.

- A soma das probabilidades de todos os possíveis resultados é 1 (ou 100%).
- Para chegar à probabilidade de obter um de um conjunto de resultados, você deve somar as probabilidades de cada um dos resultados individualmente. Por exemplo, a probabilidade de tirar um número ímpar (1, 3, ou 5) com o único lançamento de um dado é a soma das probabilidades de tirar um 1, um 3 e um 5: 1/6 + 1/6 + 1/6 = ½, ou 50%.
- O complemento de um evento é todos os resultados possíveis exceto os que constituem o evento. A probabilidade do complemento de um evento é 1 menos a probabilidade do evento. Por exemplo, tirar um 1, 2, 3, 4, ou 5 é o complemento de tirar 6 em apenas um lançamento de um único dado, portanto a probabilidade de tirar 1, 2, 3, 4 ou 5 é 1 menos a probabilidade de tirar um 6, ou seja, 1 – 1/6 = 5/6.

DICA

Quando o complemento de um evento é complicado, geralmente é mais fácil encontrar a probabilidade do evento propriamente dito e subtrair 1 da probabilidade calculada. E por que fazer essa subtração? Pois a soma de todas as probabilidades de todos os resultados é igual a 1, assim a probabilidade do complemento de um evento mais a probabilidade do evento deve ser igual a 1.

Jogando os dados

No jogo de craps, dois dados são lançados e o número 7 tem um importante papel nesse jogo. No craps, cada resultado compõe-se dos dois números no dado (por exemplo, a combinação 6,2 é um resultado). Os números dos dois dados são somados (veja a Tabela 6-1) e o resultado 7 é o mais frequente e, por isso, ele tem a maior probabilidade de acontecer. O shooter é o jogador que lança os dados e qualquer resultado que ele obtenha é chamado de ponto (por exemplo, a soma da combinação 6,2 resulta no ponto 8). Se o ponto for maior do que 7, o shooter termina sua rodada e todos que apostaram perdem. Se o ponto não for maior do que 7, o shooter continua lançando os dados até que saia um ponto 7 ou o primeiro ponto (nesse caso, 8). Qualquer jogador na mesa pode apostar que o ponto 7 irá ou não aparecer antes do ponto aparecer novamente. E é por isso que todos ao redor da mesa ficam alvoroçados e torcem para o jogador (shooter). Eles esperam que o shooter traga-lhe boa sorte lançando as combinações em que eles apostaram.

Você pode usar as regras de probabilidade listadas na seção anterior para verificar os resultados das somas dos dois dados e designar as probabilidades para elas. Você sabe quais são as outras somas com maior probabilidade de sair?

Quando dois dados são lançados, cada um deles tem seis resultados possíveis; juntos, esses seis resultados possíveis para cada dado produzem 36 (6 × 6) combinações possíveis, ou seja, 36 pares possíveis. Pelo fato de que nesse exemplo um resultado é a soma de números obtidos a partir de dois dados, você tem onze diferentes resultados possíveis, que variam de 2 (a soma de 1 + 1) a 12 (a soma de 6 + 6). A Tabela 6-1 mostra os 36 resultados possíveis para os lançamentos dos dados e também as 11 diferentes somas.

Tabela 6-1 Resultados para a Soma de Dois Dados

Resultado do lançamento do dado	Soma	Resultado do lançamento do dado	Soma	Resultado do lançamento do dado	Soma	Resultado do lançamento do dado	Soma	Resultado do lançamento do dado	Soma	Resultado do lançamento do dado	Soma
1,1	2	2,1	3	3,1	4	4,1	5	5,1	6	6,1	7
1,2	3	2,2	4	3,2	5	4,2	6	5,2	7	6,2	8
1,3	4	2,3	5	3,3	6	4,3	7	5,3	8	6,3	9
1,4	5	2,4	6	3,4	7	4,4	8	5,4	9	6,4	10
1,5	6	2,5	7	3,5	8	4,5	9	5,5	10	6,5	11
1,6	7	2,6	8	3,6	9	4,6	10	5,6	11	6,6	12

Você pode usar a primeira regra da probabilidade (veja a seção “Entendendo as regras”) para calcular as probabilidades para cada uma das 11 possibilidades de soma. A lista de todos os resultados e suas probabilidades é chamada de modelo de probabilidade. Por exemplo, a soma 7 pode acontecer de seis maneiras diferentes: (1,6), (2,5), (3,4), (4,3), (5,2) e (6,1). Com 36 possíveis combinações para os dois dados, a probabilidade de que a soma seja 7 é 6 ÷ 36, ou 1/6. Da mesma forma, você pode descobrir as probabilidades ao obter somas de 2 a 12. O modelo de probabilidade para a soma dos dois dados é mostrada na Tabela 6-2. Você pode ver que as duas somas que têm a segunda maior probabilidade de ocorrer, 5/36 são 6 e 8. Observe que a soma de todas as probabilidades na Tabela 6-2 é igual a 1. Observe também que as probabilidades aumentam gradualmente, conforme as somas dos dados passam de 2 para 3 para 4, 5, 6 e chega ao clímax quando a soma dos dados é 7 (isso porque o número de combinações cujas somas podem resultar em 7 é maior do que o número de combinações para quaisquer outras somas). As probabilidades diminuem gradualmente, conforme as somas vão de 8 a 9 até chegar ao 12.

Tabela 6-2 Modelo de Probabilidade para a Soma de Dois Dados

Soma dos Dados	*Probabilidade*
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

Para ganhar em jogos de azar é necessário basear-se em probabilidades. No jogo de craps, por exemplo, você pode apostar nas somas que vão aparecer para uma dada jogada. Se, em uma dada jogada, você apostar que os dados do shooter irão somar 2 e isso realmente acontecer, você irá ganhar mais do que se tivesse apostado na soma 8. Por quê? Pois, segundo a Tabela 6-2, conseguir uma soma igual a 2 com dois dados é muito menos provável do que conseguir uma soma igual a 8 com dois dados. É por isso que eles chamam isso de jogos de azar (para mais sobre probabilidade e jogos de azar, veja o Capítulo 7).

As probabilidades para a soma de dois dados foram relativamente simples de serem calculadas. Entretanto, outras probabilidades podem ser mais complexas, por exemplo, as probabilidades para diferentes mãos do poker tais como um full house, um straight flush ou dois pares. O que é importante lembrar, no entanto, é que o rank das mãos do poker está diretamente relacionado com a probabilidade de conseguir essas mãos; a mão mais forte do poker é o royal flush (10, Valete, Rainha, Rei e Às todos do mesmo naipe). A razão para o royal flush ser a mão mais forte do poker é o fato de ela ter a menor probabilidade de ocorrer.

Modelos e Simulações

Nem todas as probabilidades podem ser calculadas por meio da matemática. Nos casos em que a matemática não funciona, outros métodos podem ser utilizados para estimar as probabilidades ou se podem usar as probabilidades conhecidas para fazer previsões a respeito do mundo. Por exemplo, modelos computacionais complicadíssimos são utilizados para prever a probabilidade de um furacão chegar à costa dos Estados Unidos e, caso chegue, quando e onde. Esses modelos computacionais baseiam-se em dados sobre o comportamento de outros furacões, além de dados sobre as atuais condições climáticas e outras variáveis. Os cientistas colocam todas essas informações em sofisticados modelos matemáticos que tentam prever o que os furacões farão. Ainda há muito trabalho a ser feitos nessa área, mas se tem alcançado progresso a todo momento. Modelos como esse salvariam vidas, propriedades e milhões de dólares se as pessoas pudessem saber antecipadamente o que esperar e se preparar da maneira adequada.

Outros modelos baseiam-se em dados observacionais. O departemento American Comunity Survey do U.S. Census Buerau fez um levantamento sobre as residências em Columbus, no estado de Ohio, em 2001, para ter uma ideia da constituição da comunidade. Uma das características examinadas foi a composição familiar (casais, outras famílias, pessoas que vivem sozinhas e outros lares não familiares). Os dados estão resumidos na Figura 6-1. Essas estatísticas de uma amostra de residências podem servir para um modelo de probabilidade para a constituição de todas as residências em Columbus, Ohio, em 2001.

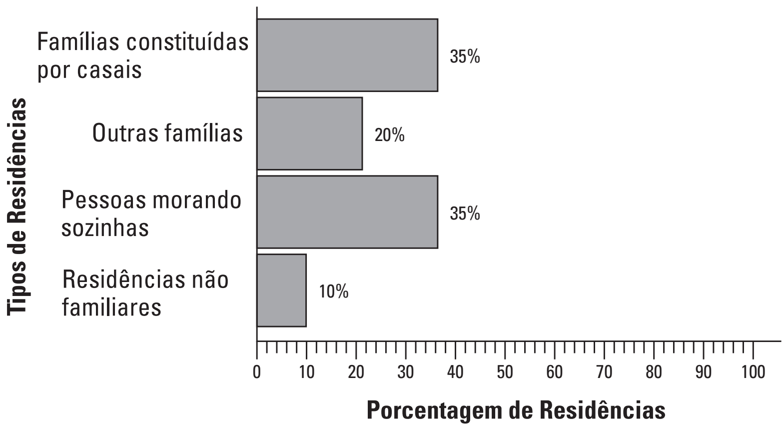


Figura 6-1: A constituição das residências em Columbus, Ohio, em 2001.

Por exemplo, pelo fato de 35% da amostra de residências serem constituídas por casais, é possível dizer que a probabilidade de que uma residência selecionada aleatoriamente em Columbus seja de um casal é de 35%. Você também pode utilizar as regras da probabilidade para fazer outras afirmações sobre as residências em Columbus em 2001. Por exemplo, qual a probabilidade de que uma residência selecionada aleatoriamente seja constituída por algum tipo de família? A resposta seria a soma das probabilidades de selecionar as residências de casais (35%) e das residências que se enquadram na categoria "outras famílias" (20%). Assim, a probabilidade de que uma residência selecionada aleatoriamente em Columbus, Ohio, em 2001, contenha uma família é de 35% + 20% = 55% (e, consequentemente, a probabilidade de selecionar uma residência não familiar é de 100% - 55% ou 45%).

O modelo de probabilidade ilustrado na Figura 6-1 não deve ser usado para outras comunidades fora de Columbus em Ohio, pois nessa pesquisa a amostra de residências foi selecionada apenas em Columbus. O uso desses dados para tirar conclusões sobre a população de outro lugar que não o local de onde a amostra foi coletada não é válido (veja o Capítulo 16, para saber mais sobre pesquisas e sobre o que elas podem e não podem dizer sobre as populações).

As simulações são outra maneira de estimar a probabilidade quando na falta de uma fórmula matemática. Em uma simulação, um processo é repetido várias e várias vezes sob as mesmas condições (geralmente com o uso do computador), e os resultados são registrados todas as vezes. A probabilidade de qualquer resultado é estimada pela porcentagem de vezes da ocorrência desse resultado durante as simulações. Por exemplo, um torcedor fanático com muito tempo livre simulou milhares de torneios do NCAA em seu computador e usou essas simulações para prever que os Duke iriam vencer o campeonato de basquete da NCAA em 2002 com uma probabilidade acima de 95%. Como a sorte é algo imprevisível, a previsão falhou (os Duke foram eliminados logo no início do campeonato), provando que a única coisa de que se pode ter certeza é da incerteza.

Interpretando a Probabilidade

Uma probabilidade pode ser interpretada de duas maneiras: como uma possibilidade a curto prazo ou como uma porcentagem a longo prazo. A curto prazo, a probabilidade de um evento é a porcentagem de chances de que esse evento venha a acontecer em uma próxima tentativa. Por exemplo, o meteorologista diz que a probabilidade de chover amanhã é de 40% ou, então, de que a média de rebatidas de um determinado jogador de beisebol é 0, 291 (o que significa que ele tem em média 29,1% de chances de fazer um ponto da próxima vez em que ele for rebater a bola).

A probabilidade também significa a porcentagem de vezes que um evento irá ocorrer a longo prazo (durante um longo período de tempo com a repetição de testes sob as mesmas condições). Assim, os 40% de chances de chover amanhã podem significar que se você observar os dados obtidos a partir de um grande número de dias semelhantes ao tipo de dia esperado para amanhã, vai descobrir que choveu em 40% desses dias. A média 0, 291 de rebatidas de certo jogador pode ser interpretada como a proporção de vezes que ele acertou a bola entre todas as vezes em que ele foi o rebatedor (neste caso, supõe-se que ele acertou a bola 291 vezes em 1000 rebatidas).

Evitando Mal Entendidos sobre a Probabilidade

As regras básicas da probabilidade parecem ser bastante simples, mas a probabilidade pode muitas vezes ir contra as expectativas. Nesta seção, você ficará por dentro dos mal entendidos mais comuns com relação à probabilidade.

Parecendo mais provável

Se você tivesse que escrever uma sequência do que pensasse ser os resultados de seis lançamentos de uma moeda não viciada (ou seja, uma moeda que não tenha sido adulterada), provavelmente você não escreveria algo como ca co co co co ca (em que "Ca" significa cara e "Co" significa coroa), pois isso não parece muito "aleatório". Entretanto, essa sequência exata de cara e coroa tem as mesmas chances de acontecer do que qualquer outra sequência. Isso porque a probabilidade de sair uma cara é a mesma probabilidade de sair uma coroa em cada lança individual. Agora, se você comparar a probabilidade de sair duas caras (em seis lances) com a probabilidade de sair seis caras (em seis lances), você vai obter diferentes valores. A probabilidade de sair duas caras (em seis lances) é maior, pois você tem mais maneiras de conseguir realizar esse feito do que conseguir tirar cara em todos os lances.

Na loteria, a sequência 1, 2, 3, 4, 5, 6 tem tantas chances de ganhar quanto qualquer outra combinação de seis números, ainda que ela não pareça ocorrer com frequência. Esse fato faz com que você perceba que todas as outras combinações simplesmente são tão improváveis de serem sorteadas quanto essa combinação. No entanto, se você apostar nessa sequência e ganhar, provavelmente não terá que dividir o prêmio com ninguém mais.

Previsão a longo e curto prazo

A probabilidade funciona muito bem para prever comportamentos a longo prazo, mas não funciona tão bem assim para a previsão de resultados a curto prazo. A longo prazo, você sabe que um evento irá ocorrer em algum momento, a não ser quando sua probabilidade for igual a 0 e, dependendo do tamanho da probabilidade, você pode, até mesmo, fazer uma ideia de quanto terá que esperar. Mas você não saberá exatamente quando o evento irá ocorrer. É isso que faz com que a probabilidade seja tão interessante e com que os apostadores voltem sempre a arriscar.

Por exemplo, se eu jogar uma moeda não viciada seis vezes e tirar seis caras sucessivamente, o que você acha que vai sair na próxima vez em que eu lançar a moeda, cara ou coroa? Você pode achar que, desta vez, deve haver uma maior chance de que saia coroa. No entanto, a probabilidade de tirar coroa no próximo lançamento ainda é de ½, a mesma que havia sido nos lançamentos anteriores. Você sabe que se a moeda for lançada muitas vezes, podemos esperar que 50% dos resultados sejam caras e 50% seja coroa. Mas não é possível prever quando esses resultados irão aparecer durante os lançamentos (portanto, ainda que pareça muito provável que saia coroa no próximo lançamento, a probabilidade de sair cara ou coroa ainda é de 50%). No final, as coroas irão começar a aparecer, mas não é possível dizer quando.

Uma em duas

Outro mal entendido é acreditar que todas as situações envolvendo apenas dois possíveis resultados seja uma situação de "uma em duas" (em outras palavras, 50% de chances para que um dos dois eventos ocorra, justamente como se fosse o resultado de um único lançamento de uma moeda não viciada). Muitas pessoas pensam que só porque há apenas dois resultados possíveis, cada resultado deve ter uma em duas chances de ocorrer, porém nem sempre é assim que acontece. Nem todas as situações são como o lançamento de uma moeda não viciada. Muitas situações têm uma probabilidade mais alta de obter um resultado sobre o outro.

Por exemplo, pense no sinal eletrônico para pedestres em uma rua muito movimentada. Esse sinal irá 50% das vezes permitir a passagem dos pedestres? A resposta é não. Quando a rua está movimentada, esse sinal irá parar o tráfico com menos frequência e os pedestres terão que esperar uma pouco mais até a próxima vez, se quiserem atravessar. Usando um exemplo esportivo, pense em uma jogadora de basquete parada na linha de lance livre. Suas chances de fazer uma cesta são uma em duas? (afinal de contas, ela pode ou não fazer a cesta). Suas chances apenas serão meio a meio se a porcentagem de todos seus arremessos livres for de 50% em muitas tentativas. Muito provavelmente será algo mais alto do que isso.

Interpretando eventos raros

A probabilidade pode ser tornar motivo de controvérsias, especialmente em se tratando de eventos raros. Um evento raro tem uma pequena probabilidade de acontecer, mas o que isso significa? Isso significa que, para qualquer situação ou pessoa, é improvável que certo evento aconteça, no entanto, se determinada situação se repetir várias e várias vezes ao longo de um período ou com um grande número de pessoas, esse evento pode acontecer para alguém em algum lugar e em algum momento. Isso vem à tona em situações em que você tem um grupo de pessoas com uma doença rara em uma cidade e precisa descobrir se algo causou aquela doença (o ar, a água, o solo, etc.) ou se ela apenas ocorreu por acaso (algo que a maioria das pessoas não leva em consideração).

Pelo fato de um evento não parecer muito provável, as pessoas naturalmente querem achar uma causa para a ocorrência desse evento. Em alguns casos, elas podem estar certas; em outros, esse evento pode ser apenas um fenômeno de possibilidades aleatórias. Três anos seguidos de aumento das temperaturas médias indicam um aquecimento global? Se em uma fazenda, duas vacas dão à luz a dois bezerros com duas cabeças no mesmo ano, isso significa que as vacas têm um terrível problema? Quantos pneus precisam estourar para que se faça um recall? Observar algo depois do fato e dizer, "Qual era a chance disso acontecer?" é diferente de saber, antes do fato, que o mesmo evento poderia ocorrer em algum lugar e em algum momento.

Por exemplo, se você lança uma moeda não viciada por várias vezes, em determinado momento, você se depara com uma série de caras, apenas por coincidência. E pelo fato de que a moeda não é viciada, a única explicação para o ocorrido é a coincidência. Entretanto, a mídia pode tentar estabelecer um padrão para os eventos com duas ou mais ocorrências, tais como o desaparecimento de crianças pelo país, incêndios em danceterias ou ocorrências de doenças raras em uma única cidade. Não estou dizendo que não se deve investigar tais acontecimentos, mas estou dizendo que a mídia precisa se conscientizar de que, às vezes, os eventos apenas acontecem por acaso, sem nenhuma grande história por trás deles. Também é interessante notar que as pessoas veem a probabilidade de um evento raro de maneiras diferentes, dependendo se o evento é algo bom, como ganhar na loteria ("Alguém vai ganhar e por que não eu?) ou algo ruim, como ser atingido por um raio em um torneio de golfe ("As chances de que isso aconteça são uma em um milhão. Isso nunca vai acontecer comigo!"). Talvez seja apenas a natureza humana. Observação mental: a natureza humana não corresponde às leis da probabilidade.

Para evitar alguns dos mal entendidos mais comuns com relação à probabilidade, tenha isto em mente:

- A probabilidade não é eficiente para a previsão de comportamentos a curto prazo. Ela é eficiente para a previsão de comportamentos a longo prazo.
- Nos casos em que há apenas dois resultados possíveis, cada resultado não tem necessariamente 50% de chances de ocorrer.
- Se vários eventos raros ocorrem em algum lugar, eles podem ter sido causados simplesmente pela casualidade. Os eventos raros irão acontecer com alguém em algum lugar e em algum momento, desde que haja tempo e pessoas o suficiente.
- Você não consegue manter sua sorte o tempo todo se o mesmo processo estiver sendo repetido várias e várias vezes sob as mesmas condições (como em uma situação de jogo). A probabilidade não tem memória.
- As sequências de resultados que "pareçam mais aleatórias", com frequência, têm a mesma probabilidade de sequências que as que não parecem tão aleatórias. Você pode pensar que a sequência ca co co co co ca tem uma chance menor de ocorrer do que a sequência ca co co co ca co, pois a primeira não parece "tão aleatória" quanto a segunda. Na realidade, elas têm a mesma probabilidade de ocorrer, pois cada um dos resultados contém quatro coroas e duas caras (e, aqui, a ordem não faz diferença quando se calcula a probabilidade).

Unido a Probabilidade à Estatística

Você pode estar pensando, "Probabilidade é interessante, mas o que ela tem a ver com a estatística?" Boa pergunta! Pode não parecer tão óbvio, mas a probabilidade e a estatística caminham juntas. Os dados são coletados a partir de uma amostra de indivíduos, para que, depois, a estatística seja calculada, a fim de resumir aqueles dados. Mas você não para aqui. O próximo passo é usar essas estatísticas para fazer prever, generalizar, concluir ou decidir algo a respeito da população da qual se tirou a amostra. É aí que entra a probabilidade.

Estimando

Os dados são coletados para ajudar a estimar as proporções ou médias da população. Por exemplo, para estimar as chances de um paciente sofrer um ataque cardíaco, primeiramente, os médicos reúnem algumas informações sobre o paciente, tais como peso, índice de massa corporal, sexo, tendência genética, hábitos alimentares, frequência de exercícios físicos, entre outras. Depois, comparam essas informações a dados coletados de uma amostra de pessoas com características semelhantes às desse paciente e, então, estabelecem a probabilidade do paciente (o nível de risco) ter um ataque cardíaco dentro de um dado período. Engenheiros estimam o número médio de carros que transitarão em determinado trecho de uma rodovia durante os horários de pico por meio do registro de dados de trânsito, usando sensores no asfalto. Depois que os dados são coletados, a probabilidade é utilizada para determinar o quanto a informação das amostras pode variar de amostra para amostra, de dia para dia, de hora para hora, etc.

Prevendo

A estatística está envolvida em todo tipo de previsão – desde a previsão do tempo e do tamanho de uma população até a previsão da difusão de uma doença ou dos valores futuros da bolsa de ações. Os dados são coletados ao longo de um período e são analisados a fim de encontrar um modelo que não apenas sirva os dados, mas que também permita com que se façam algumas previsões. A probabilidade ajuda às pessoas a avaliar a precisão esperada para as previsões, de acordo com os dados que se têm à mão. A probabilidade também ajuda os cientistas a determinar quais serão os cenários mais prováveis, de acordo com os dados coletados.

Por exemplo, o U.S. Census Bureau disponibiliza suas projeções para o tamanho da população americana. Atualmente, é possível encontrar projeções até 2100. Em 2000, a população projetada para 2003 era de 282.798.000 e, até maio de 2003, a população americana (como se pode ver no site do U.S. Census Buerau) já era de 291.065.455. Portanto, a projeção já tinha uma diferença de, aproximadamente, 8,3 milhões de pessoas, faltando ainda mais da metade do ano. Mas esse número representava apenas 2,8% da população total nessa situação. Estimar a população total dos Estados Unidos para o futuro não é uma tarefa fácil. Já é difícil contar o número de pessoas vivendo aqui e agora! (A propósito, segundo o Census Bureau, o tamanho da população americana será de 570.954.000 pessoas em 2100).

Decidindo

Muitos processos de tomada de decisão envolvem estatística e probabilidade. Os tratamentos médicos são geralmente decididos levando em consideração a porcentagem de pessoas que melhoraram ao utilizar certo tratamento, quando comparadas com outras. A probabilidade de que outra pessoa melhorará usando certo tratamento seria estimada a partir da porcentagem de outros pacientes que tiveram melhoras com aquele tratamento. A maioria dos termos de responsabilidade que você precisa assinar antes de submeter-se a uma cirurgia enfatiza os possíveis efeitos colaterais ou complicações e dá algumas indicações da frequência que eles ocorrem (veja o Capítulo 17, para saber mais sobre estudos médicos).

Verificando a qualidade

Outras decisões que envolvem a probabilidade ocorrem durante os processos de manufatura. Muitas indústrias costumam conduzir procedimentos de controle de qualidade; ou seja, elas coletam amostras dos produtos que saem das linhas de produção e avaliam a qualidade desses produtos de acordo com um conjunto de especificações. A probabilidade é usada para decidir se e quando o fabricante precisa interromper a linha de produção em virtude de um problema com a qualidade de seus produtos. As diferenças entre o produto coletado e as especificações pode acontecer apenas por causa da variabilidade aleatória ou por causa de uma amostra não representativa; Em contrapartida, essas diferenças podem significar que há algo de errado com o processo. Interromper a linha de produção sem necessidade custa dinheiro e tempo para a indústria, mas não interrompê-la quando necessário custará a reputação da empresa junto aos consumidores. Assim, a probabilidade é utilizada para a tomada de decisões muito importantes no âmbito industrial (veja o Capítulo 19, para saber mais sobre controle de qualidade).

LEMBRE-SE

A probabilidade é usada para avaliar a precisão da generalização dos resultados feita a partir de uma amostra da população. Ela é utilizada para determinar qual conclusão é mais provável e o motivo de ser assim. Ao tomar uma decisão sobre uma situação com um resultado desconhecido, você usa a probabilidade para avaliar a evidência que foi coletada, fazer uma escolha baseando-se nessa avaliação e saber as chances de que você tenha tomado a decisão certa ou errada (veja o Capítulo 14, para mais informações).

Capítulo 7

Jogando para Vencer

Neste Capítulo

- Porque os cassinos ganham dinheiro
- Entendendo a probabilidade por de trás dos jogos
- Pegue seu dinheiro e caia fora: dicas para os jogadores

Las Vegas é um dos lugares mais emocionantes do mundo. No entanto, tenho a sensação de que os baratíssimos restaurantes "coma a vontade" e os majestosos gladiadores romanos que vagam pelo Caesar's Palace não são as principais atrações (embora eu realmente os recomende!). Las Vegas é conhecida como o paraíso dos jogos de azar; o lugar certo para quando você se sentir com sorte e estiver com vontade de ganhar muito dinheiro. O fato de que a grande maioria das pessoas que aposta suas fichas em Las Vegas volta para casa como perdedoras não é tão importante para os vários ganhadores em potencial que diariamente lotam aviões rumo a Nevada, cheios de convicção e esperança. Afinal de contas, alguém tem que ganhar, não é mesmo? E esse alguém pode ser você. Embora eu não possa prometer que este capítulo irá lhe transformar em um grande vencedor em Las Vegas (ou em qualquer outro lugar que você escolha tentar a sorte), eu posso dizer que este capítulo lhe ajudará a entender o que você tem que enfrentar, lhe dará dicas para ganhar o máximo possível e, caso tudo falhar, pelo menos lhe ajudará a não perder tanto.

Apostando na Banca: Por que os Cassinos Ainda estão no Páreo

Os cassinos são lugares lindos; eles possuem ambientes encantadores, com máquinas de cores vibrantes, luzes que piscam, músicas contagiantes, garçons e carteadores alegres, sem relógios ou janelas em nenhum lugar (eles não querem que você se dê conta de quanto tempo você está gastando lá dentro). Tudo é muito bem planejado, desde a disposição do prédio (para ir a qualquer banheiro, você passa por centenas de caça-níqueis) até a estampa escolhida para o carpete (as estampas são intencionalmente confusas e incômodas para os olhos, pois os donos de cassinos não querem que você olhe para baixo, mas sim que você olhe para o que está acontecendo e desloque-se sempre em busca de sua próxima aventura).

A indústria do jogo tem agido cientificamente, e as pessoas que administram as casas são muito boas no que fazem. Eles oferecem muita diversão e uma chance de ganhar muito dinheiro; tudo o que você precisa fazer é jogar. Muitas pessoas se deixam levar pela chance de ganhar milhares de dólares ou um carro novo com apenas um puxão em uma máquina caça-níquel ou com apenas uma mão perfeita no blackjack (vinte um). É claro que essa chance existe, mas se todos ganhassem muito dinheiro, os cassinos não conseguiriam se manter abertos. Então, eles precisam encontrar maneiras de pegar o seu dinheiro e fazem isso estabelecendo regras para os jogos que lhe colocam em ligeira vantagem durante cada rodada; assegurando que quando alguém ganhar, ele ou ela irá ganhar alto (e, então, o falatório irá correr pelo cassino); e encoraja-lhe a permanecer no cassino o máximo de tempo possível. Eles sabem que quanto mais você jogar, mais são as chances que eles têm de ficar com seu dinheiro.

Centenas de livros sobre como quebrar a banca têm sido escritos. Cada autor tenta lhe convencer de que sua estratégia irá fazer um grande vencedor. A verdade é que o melhor que esses livros podem fazer é ajudar-lhe a não perder muito, pois da maneira como os jogos são estabelecidos, a banca (o cassino) sempre tem vantagem. Essa vantagem é menor em alguns jogos, tais como o Blackjack, e maiores em outros, com as máquinas de caça-níquel, que, segundo rumores, são responsáveis por mais de 80% do faturamento total de alguns cassinos. A coisa mais importante a ser lembrada é que em qualquer situação de jogo (jogo é o eufemismo utilizado pelos cassinos para se referir à jogatina) é a de sair fora enquanto você estiver ganhando. Se todo mundo fizesse isso, os cassinos teriam falido. Mas é claro que isso não acontece, pois sair do jogo enquanto se está ganhando é algo muito difícil de fazer.

No outro lado da moeda, outra boa estratégia é cortar seus prejuízos e sair antes de perder demais, ao invés de pensar que as probabilidades irão virar a seu favor e você irá recuperar todo o dinheiro que perdeu. Essa abordagem quase sempre se torna uma situação em que todos perdem. Os cassinos apostam que você vai ficar e jogar independente do que aconteça, o que aumenta as chances deles de, no final, tirar mais e mais o seu dinheiro. Olhando os extravagantes cassinos que estão sendo construídos em Las Vegas, parece que as apostas deles têm valido a pena.

LEMBRE-SE

O que você pode fazer para diminuir suas chances de perder ou de perder muito? Estabeleça alguns limites para si mesmo antes mesmo de entrar em uma situação de jogo (por exemplo, pare de jogar enquanto ainda estiver ganhando ou quando já tiver perdido uma determinada quantia) e obedeça a esses limites. Enquanto estiver na jogatina (desculpe, no jogo), pare enquanto estiver em vantagem ou antes de ficar muito para trás. Estabeleça seus limites antes de começar.

Saber Um Pouco de Probabilidade Dá uma Mega Ajuda

As duas ferramentas mais importantes para você usar em um jogo são a informação sobre suas chances de ganhar e a verdadeira compreensão do que elas significam. A probabilidade certamente vai contra sua intuição, mas você não deve deixar que sua intuição cruze seu caminho e lhe faça perder dinheiro (veja no Capítulo 6 uma discussão mais detalhada sobre o uso da probabilidade na estatística).

Aqui vão alguns dos mal entendidos mais comuns sobre a probabilidade:

- Qualquer situação que tenha apenas dois resultados possíveis é uma situação em que as chances são 1 em 2 (50% de chance de ganhar, 50% de chance de perder).
- A sequência de números 1-2-3-4-5-6 nunca irá ser sorteada, pois não é aleatória o suficiente.
- Comprar 100 bilhetes de loteria ao invés de um é uma grande ideia, pois 100 bilhetes lhe dão mais chances de ganhar.
- Se um casal tem três filhas meninas, a chance de que seu próximo filho seja um menino é muito alta.
- Quanto mais você jogar nas máquinas caça-níquel, mais chances você tem de ganhar.

Nesta seção, dividi cada um desses mal entendidos para que você possa ficar informado sobre a realidade de qualquer jogo de azar, compreender melhor o que você deve esperar e preparar-se adequadamente. Este capítulo poderá arruinar a mágica e a emoção causada pelos jogos, mas, mais uma vez, isso provavelmente explica o porquê de você não encontrar estatísticos (certamente nenhum que eu conheça) que sejam jogadores profissionais ou compulsivos. Na verdade, Las Vegas não se importará se os estatísticos não realizarem mais sua conferência lá, pois eles não gastaram muito dinheiro nos cassinos da última vez em que estiveram na cidade (Eu mesma só joguei na máquina de caça-níquel uma moeda de cada vez. Pelo menos meus $20 irão durar bastante).

Tendo 50% de chance

Você vai lançar uma moeda não viciada (ou seja, uma moeda que não tenha sido adulterada), um de seus lados é cara, o outro coroa. Quais são as chances de sair cara? Cinquenta por cento. Quais são as chances de sair coroa? Cinquenta por cento. Se você fosse apostar no resultado do lançamento da moeda, você teria 50% de chances de vencer (ou de perder). Por que isso acontece? Porque você tem apenas dois resultados possíveis, cara ou coroa, e cada um desses resultados tem chances iguais de ocorrer. Você conhece um casal que vai ter um filho. Eles, é claro, podem ter um menino ou uma menina. Cada um desses resultados tem a mesma probabilidade de acontecer, portanto a chance de ter uma menina (versus a chance de ter um menino) é de 50%. Por outro lado, se você comprar 1.000 bilhetes de rifa de uma moto, você tem apenas dois resultados possíveis, ganhar ou perder a moto. Isso significa que a chance de você ganhar a moto é de 50%. Não. Por que não? Porque você não foi a única pessoa que comprou os bilhetes!

Quatro regras de probabilidade podem ajudar a desfazer um pouco essas ideias:

- A probabilidade de que certo resultado irá acontecer é a porcentagem de vezes que se espera que o resultado apareça a longo prazo, caso exatamente as mesma condições estejam sendo repetidas.

- Qualquer probabilidade é um número entre 0 e 10. Uma probabilidade igual a 0 significa que o resultado nunca será possível. Uma probabilidade igual a 1 significa que o resultado é certo

- A soma das probabilidades para todos os resultados possíveis deve ser igual a 1. Isso significa que a probabilidade de um resultado não sair é igual a 1 menos a probabilidade de que o resultado realmente ocorra.

- A probabilidade de um evento (uma combinação de resultados) é igual à soma das probabilidades dos resultados individuais que compõem o evento

No exemplo do lançamento de uma moeda, apenas dois resultados são possíveis, cara ou coroa. O número de maneiras com que o resultado cara pode ocorrer é de apenas um e o número de maneiras com que o resultado coroa pode ocorrer também é de apenas um. O número total de possíveis resultados é 2: cara ou coroa. Sendo assim, nesse caso, a chance de se obter cara é ½ ou 50%; e o mesmo também é verdadeiro para coroa. Essa é uma situação de 50% de chances para cada resultado.

Entretanto, se você observar atentamente a primeira regra, ela explica o porquê nem todas as situações com apenas dois resultados possíveis têm 50% de chance para cada um deles. É necessário levar em consideração as possíveis maneiras de se obter cada resultado. No exemplo da rifa da moto, você pode tanto perder quanto ganhar. O número de maneiras com que você pode ganhar essa rifa é de apenas um, pois os organizadores da rifa irão sortear apenas um número. O número de maneiras com que você pode perder é de 999, pois todos os números restantes serão perdedores. O número total de resultados é 1.000, o que significa que suas chances de ganhar é de $1 \div 1.000 = 0{,}001$ e sua chance de perder é $999 \div 1000 = 0{,}999$. Certamente, só há dois resultados possíveis com relação ao bilhete da rifa (ganhar ou perder), porém cada um desses resultados não tem as mesmas chances de acontecer, portanto essa definitivamente não é uma situação de 50% de chances para cada resultado.

Muito poucas probabilidades na vida são, realmente, meio a meio. Para estar em situações em que as chances para ambos os resultados seja de 50%, é necessário que se tenham apenas dois resultados possíveis e a probabilidade para cada um deles dever ser a mesma, ou seja, 50%. Na maioria das situações com apenas dois resultados possíveis, não é provável que os dois resultados possam ocorrem igualmente.

Escolhendo os números vencedores

Então, você está decidido a apostar na Powerball Lottery. Você ficou sabendo que o prêmio já está acumulado em $200 milhões – talvez você compre mais bilhetes do que o de costume para aumentar suas chances de ganhar e está pronto para escolher os números certos (nesse tipo de loteria americana, você tem que escolher cinco números diferentes de 1 a 53 e, separadamente, você ainda tem que escolher apenas um número de 1 a 42, que pode ser igual a algum dos números escolhidos para formar a primeira sequência). Para ganhar o prêmio máximo, é preciso acertar todos os cinco números (independente da ordem) mais o número escolhido separadamente (número Poweball). Que combinação você dever fazer? O número da camisa de futebol do seu irmão, o aniversário da sua mãe, quatro números de seus documentos, a idade em mês de seu cachorro e o número com que você sonhou a noite passada? Essas opções soam tão boas quanto qualquer outra, pois qualquer combinação que você escolher tem justamente a mesma probabilidade de sair do que qualquer outra.

Isso faz sentido até você pensar na combinação 1, 2, 3, 4, 5,com o número Powerball sendo o 6. Parece que essa combinação nunca irá sair, pois esses números não são nada aleatórios. Bom, esse é um caso no qual sua intuição pode tirar o melhor de você. De fato, essa combinação tem a mesma chance de qualquer outra. Vamos pegar um pequeno exemplo, em que os possíveis números sejam 1, 2, 3, 4 e você tenha que escolher dois deles. Os seis resultados possíveis são 1-2; 1-3; 1-4, 2-3; 2-4; 3-4. Sua chance de ganhar é 1 em seis. Observe que a combinação 1-2 tem a mesma chance das outras combinações de ser sorteada. O mesmo ocorre com a combinação 1-2-3-4-5 com o número Powerball sendo o 6. Combinações como 23-26-05-24-18 com o Powerball sendo o 12 podem parecer mais fáceis de ocorrer, mas lembre-se de que você tem que acertar número por número para ganhar o grande prêmio.

FIQUE ATENTO

Uma combinação como 1-2-3-4-5 com o número Powerball sendo o 6 parece difícil de ser sorteada, mas, na verdade, ela tem a mesma chance que qualquer outra combinação. O que essa combinação deve fazer você perceber é o quanto é pequena a chance de ganhar o prêmio máximo (a real chance de ganhar o prêmio máximo em uma loteria como a descrita acima é de 1 em 120.526.770).

PAPO TÉCNICO

As loterias normalmente se referem à probabilidade de ganhar como as chances de ganhar e, nesse caso, esses dois termos assumem o mesmo significado. Entretanto, a maneira como os prêmios e as chances são apresentados pelas apostas esportivas (corrida de cavalos, jogos de futebol, lutas de boxe e outras) é diferente do que está descrito aqui e essa forma mais complexa de chances de aposta está além do intuito deste livro.

Comprando bilhetes de loteria – menos pode ser mais

Um bilhete de loteria Powerball custa apenas um dólar e lhe dá uma chance de ganhar um prêmio multimilionário. Você sabe que alguém, em algum momento, vai ganhar, portanto, decide tentar a sorte. Afinal de contas quem não arrisca não petisca. Desde que você realmente entenda suas reais chances de ganhar e de perder, comprar poucos bilhetes pode ser uma diversão barata.

O problema acontece quando as pessoas compram muitos bilhetes, acreditando que as suas chances de ganhar serão bem maiores se comprarem mais bilhetes. Comprar uma centena de bilhetes realmente aumenta suas chances de ganhar uma centena de vez, mas você tem que se lembrar de que as chances de ganhar o prêmio máximo são muito pequenas – quase zero. E cem vezes um número próximo de zero ainda é muito próximo de zero. Se você não pode se dar ao luxo de perder cem dólares (pois as chances são realmente de que você perca), não aposte tudo na loteria.

Para colocar a probabilidade de ganhar o prêmio máximo na Powerball em perspectiva, veja a Figura 7-1. Ela mostra os prêmios e as chances de ganhar cada um deles (a maioria das loterias Poweball oferece os mesmo prêmios e as mesmas chances). Os círculos cinzas indicam as cinco bolas escolhidas de 1 a 53 e o círculo preto indica o número Powerball. As chances de ganhar o prêmio de $3 são de 1 em 70, ou 0.0142 (cerca de 15%). Observe que isso se parece mais com uma probabilidade do que com chances reais, mas vamos deixar isso para lá. Conforme você sobe na escala de prêmios, suas chances de ganhar diminuem de maneira exponencial. As chances de acertar 4 dos 5 números, por exemplo, são de 1 em cerca de 12.000; as chances de acertar 5 de 5 são 1 em cerca de 3 milhões e, por fim, as chances de acertar os 5 números mais o Powerball são de 1 em mais de 120 milhões.

Você Pode Ganhar $10 Milhões ou Mais!

Acertos	Prêmios	Chances
○○○○○ + ●	Prêmio Máximo	1:120,526,770
○○○○○	$100,000	1:2,939.677
○○○○ + ●	$5,000	1:502,195
○○○○	$100	1:12,249
○○○ + ●	$100	1:10,685
○○○	$7	1:261
○○ + ●	$7	1:697
○ + ●	$4	1:124
●	$3	1:70

Figura 7-1: Prêmios e chances de ganhar na Powerball.

PAPO TÉCNICO

Por que as chances diminuem tão drasticamente apenas por ter que acertar um número a mais? Um pequeno exemplo pode ilustrar o que ocorre. Suponha que você tenha que escolher dois números de 0 a 9. Se você tiver que acertar apenas um número, suas chances são de 1 em 10. Se você tiver que acertar os dois números, suas chances caem para 1 em 45. Isso acontece, pois você tem 10 possibilidades para o primeiro número vezes 9 possibilidades para o segundo número (já que não pode haver números repetidos), totalizando 90 possibilidades. Mas, agora, você deve dividir essas 90 possibilidades por 2, pois a ordem dos números não interfere no resultado (por exemplo, não é preciso contar as combinações 1-0 e 0-1 separadamente). As chances mudam drasticamente por causa desse efeito multiplicador. A necessidade de ter que acertar o número Powerball além de ter que acertar a combinação dos cinco números multiplica as chances por 42 (pois você tem 42 possibilidades para o número Powerball).

A loteria informa que as chances de ganhar qualquer um dos prêmios (não apenas o prêmio máximo) é de 1 em 36. Devido ao fato de que todas as chances de ganhar é a mesma de ganhar qualquer prêmio, some todas as probabilidades e você chega ao resultado 1 em 36. Aqui, você está utilizando a quarta regra da probabilidade (veja a seção Tendo uma chance de 50%, no início do capítulo).

LEMBRE-SE

Antes de jogar qualquer jogo de probabilidade, sempre analise as chances ou probabilidades de ganhar. E não gaste mais do que você possa perder. Com jogos que oferecem grandes prêmios, as chances de ganhar são extremamente pequenas e o fato de comprar muito mais bilhetes ou de jogar muito mais vezes não aumenta significativamente suas chances a ponto de justificar o gasto extra. Como diz o ditado, a melhor maneira de multiplicar seu dinheiro é dobrando-o ao meio e guardando-o no bolso!

Menino ou menina!

Um casal já teve três meninas e, agora, está esperando seu quarto filho. Eles realmente querem um menino dessa vez. Amigos e parentes acham que as chances de que eles tenham um menino é mais alta dessa vez, pois eles já tiveram três meninas. Eles estão certos? Essa situação é semelhante a uma mesa de craps, em que as pessoas ao redor estão torcendo para o shooter (o jogador que lança os dados), pois a sorte está a seu lado, e nada pode dar errado. Mas será que a sorte fica mesmo a favor ou contra alguém? E será que o fato de uma série de eventos ter ocorrido sucessivamente aumenta ou diminui as chances de que esse evento volte a acontecer?

Em muitas situações, especialmente em jogos, a sorte não acompanha nem abandona ninguém, pois a cada vez que você joga, tudo volta a zero e o resultado da última vez não tem nada a ver com o resultado da próxima. Em outras palavras, quando os eventos são independentes um do outro, a probabilidade de um evento continuar a ocorrer é sempre a mesma a cada jogada.

Para o casal do nosso exemplo, essa não é uma boa notícia, mas a probabilidade de que eles tenham um menino ainda é de 50%, como sempre foi, independente do fato de que eles já tiveram três meninas. Da mesma forma, se uma moeda não viciada for lançada três vezes e o resultado for três caras, não espere que no quarto lançamento as chances de aparecer coroa sejam mais altas. As chances ainda são de 50%.

As probabilidades servem somente para prever comportamentos a longo prazo e não a curto prazo. Se você sabe que as chances de tirar cara são de 50%, isso significa que se a moeda for lançada várias vezes, pode-se esperar que metade dos resultados sejam cara e a outra metade coroa. Entretanto, não é possível prever quando essas caras e essas coroas irão aparecer. Elas não se igualam, conforme vão aparecendo; com o tempo, elas apenas se aproximam de suas reais probabilidades. Esse fenômeno é conhecido como a lei das médias, que será definido na seção a seguir. Muitas pessoas usam esse termo para definir por que a maré de sorte foi embora, mesmo quando essas "marés" nem existiam.

Tentando ganhar um níquel

A velha máquina de caça-níquel é uma arma poderosa. Essas máquinas são feitas de um material especial que faz com que o som das moedas caindo seja projetado bem alto. As máquinas piscam e apitam todas as vezes em que você ganha alguma coisa, elas apitam até mesmo quando você coloca seu dinheiro nela, só para encher ainda mais a atmosfera do cassino com aquele som de vitória. Alguns dizem que as máquinas mais fáceis (as que mais pagam prêmios) ficam próximas à entrada dos cassinos ou nas últimas fileiras. Outros dizem que as mais difíceis ficam próximas às mesas de blackjack, pois as pessoas não querem ser distraídas por aquele apito. Os donos de cassinos nunca revelam seus segredos e, portanto, ninguém pode dizer com certeza, mas uma coisa é certa: as máquinas de caça-níquel podem levar seu dinheiro embora rapidamente. Nenhuma habilidade é necessária, e cada puxão da alavanca leva apenas alguns segundos para fazer com que você gaste algo que possa variar de cinco centavos (como no meu caso) a milhares de dólares (como nas áreas de apostas mais altas em alguns cassinos). Tudo isso por um único objetivo: ganhar o grande prêmio que pode estar no próximo puxão.

Um dos mal entendidos mais comuns acerca da máquina caça-níquel é que quanto mais você jogar, maiores são suas chances de ganhar. Isso é no que os cassinos estão apostando. Na realidade, o contrário é que é verdadeiro, por causa da lei das médias. A lei das médias diz que com o tempo, as médias ficarão próximas ao valor esperado para elas. Em termos estatísticos, o valor esperado é uma média ponderada dos resultados, baseada em suas probabilidades.

No caso de qualquer jogo de cassino, a banca tem uma ligeira vantagem de ver qualquer rodada. Isso significa que, com o decorrer do tempo, espera-se que você tenha grandes chances de perder pouco e poucas chances de ganhar muito em todas as vezes que jogar. Os cassinos estabelecem suas chances e prêmios para que, com o decorrer do tempo, todas as médias fiquem a seu favor, mesmo levando em consideração o grande prêmio.

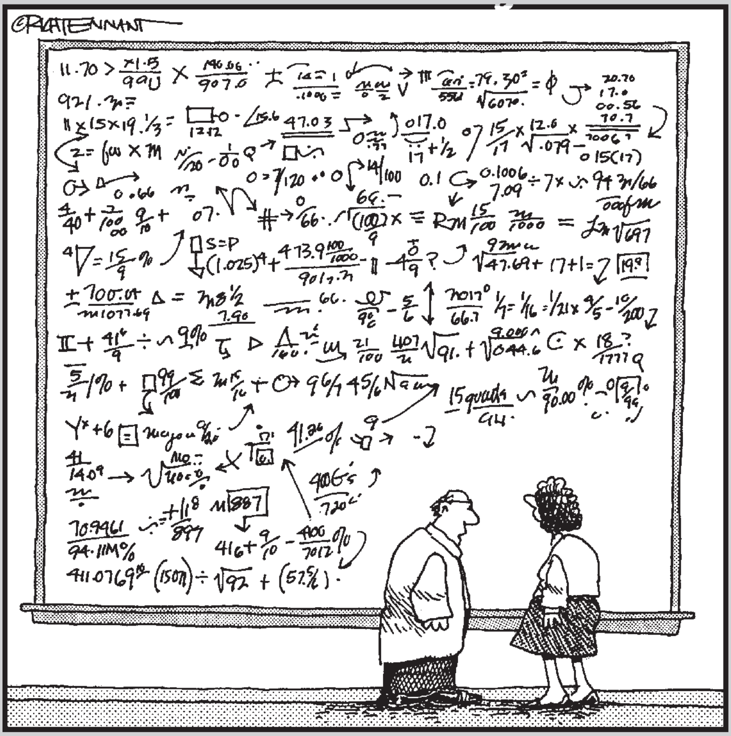
O que isso quer dizer é que você irá acabar perdendo um pouco a cada vez que jogar e, quanto mais você jogar, mais dinheiro se espera que você perca, pois seu valor total esperado é a soma dos valores esperados para cada vez que você jogar; e cada um desses valores esperado será um número negativo. Agora você já sabe por que os cassinos lhe oferece bebida de graça enquanto você está jogando e o porquê de você não ver relógios nas paredes nem janelas que lhe permitam ver a passagem do tempo enquanto você está sentado naquele banquinho da máquina caça-níquel. Os cassinos estão apostando que você irá esquecer tudo sobre a lei das médias, conforme ela vem chegando de mansinho perto de você.

DICA

A única maneira de derrotar as máquinas de caça-níqueis é sair do jogo enquanto você estiver em vantagem, antes que a lei das médias chegue. Lembre que os cassinos têm a probabilidade a seu favor, pois eles ficarão lá por mais tempo. Você não pode brincar com a Mãe-Natureza e nem com a probabilidade. Mas se você realmente ganhar, pegue seu dinheiro e vá embora!

Parte IV
Decifrando Resultados

A 5ª Onda por Rich Tennant



"Sobre o que nós estávamos falando mesmo?"

Nesta Parte...

Esta parte lhe ajudará a compreender as bases da estatística – as informações que ajudam a entender os assuntos mais profundos, aqueles que estão nos bastidores sempre que uma estatística é formulada. Você vai descobrir como medir a variabilidade de uma amostra para outra, como encontrar uma fórmula para medir a precisão de uma estatística e como medir a posição relativa de um indivíduo (a posição que um indivíduo ocupa em relação aos demais). Todos esses tópicos lhe ajudarão a se tornar mais seguro para interpretar e entender melhor os alicerces da estatística.

Capítulo 8

Medida da Posição Relativa

Neste Capítulo

- Investigando a curva com forma de sino
- Usando a regra empírica para a maioria dos casos
- Calculando e interpretando os escores padrões
- Descobrindo os percentis

A única maneira de realmente conseguir interpretar os resultados estatísticos é ter algo com o que os comparar, para que assim você possa colocá-los em algum tipo de perspectiva. Por exemplo, suponha que um estudante de fisioterapia faz um teste padrão para conseguir sua certificação e consegue uma pontuação de 235. O que essa pontuação significa nesse caso? Nada, se isso é só o que você tem para continuar. É preciso colocar essa pontuação em alguma perspectiva ao determinar sua posição relativa às outras pontuações. Uma lâmpada que dura mais de 1.200 horas é uma fenômeno da natureza ou apenas uma lâmpada padrão? Não é possível dizer sem saber o quanto a maioria das lâmpadas dura. Suponha que a nota média de um aluno no final de um curso de matemática tenha sido de 78; isso é um B ou um C? Depende da posição de sua nota quando comparada com as notas médias dos outros alunos (e de como está o humor do professor!).

Neste capítulo, você vai descobrir como encontrar e interpretar a posição relativa de resultados individuais; o objetivo aqui é descrever a posição de um indivíduo em relação a todos os outros indivíduos da população estudada. Nos Capítulos 9 e 10, discutirei como encontrar e interpretar a posição relativa de resultados a partir de uma amostra (por exemplo, a média amostral ou a proporção amostral). O objetivo neste caso será o de determinar a posição ocupada por sua média amostral ou proporção amostral, comparada à população de todos os valores possíveis da média amostral ou da proporção amostral.

Endireitando a Curva Em Forma de Sino

O primeiro passo para determinar onde um resultado em particular se encontra, é conseguir uma lista ou imagem de todos os possíveis resultados que a variável pode assumir dentro da população estudada e a frequência em que esse valores ocorrem; isso se chama distribuição.

Muitos tipos diferentes de distribuição são possíveis. Por exemplo, as notas de uma turma (vamos chamá-la de turma do Senhor Média) estão distribuídas uniformemente, com um número igual de escores em cada grupo (veja a parte superior da Figura 8-1), enquanto as notas de outra turma (a turma do Senhor Mediano) estão distribuídas de maneira polarizada, com notas que são ou muito altas ou muito baixas (como na parte inferior da Figura 8-1). (A maioria das distribuições tende a se enquadrar em algum lugar entre esses dois tipos). Observe que para qualquer distribuição, o total de todas as porcentagens tem de ser 100%, pois cada valor tem que aparecer em algum lugar na distribuição.

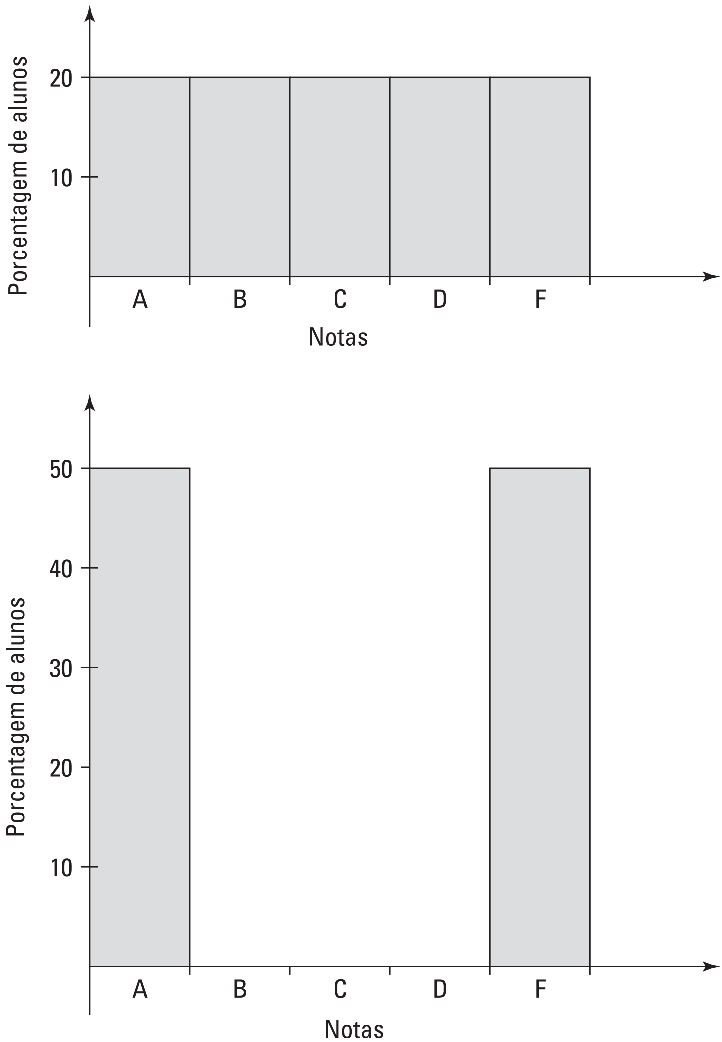


Figura 8-1: Distribuição das notas de duas turmas.

A curva em forma de sino descreve os dados de uma variável que tenha um número infinito (ou muito grande) de valores possíveis, e esses valores se distribuem entre a população, de maneira que quando são retratados por um histograma, resultam em uma figura semelhante a um sino. Isso basicamente significa que você tem um grande número de indivíduos próximos ao centro da distribuição e cada vez menos indivíduos nas duas extremidades. Muitas variáveis do mundo real (como, por exemplo, pontuação de exames padronizados, duração de produtos, alturas, pesos, (...) etc.) possuem distribuições parecidas com uma curva em forma de sino, o que faz com que a curva em forma de sino (ou, simplesmente, curva do sino) seja importante o suficiente para se destacar entre todos os outros possíveis tipos de distribuições.

Na estatística, uma distribuição cuja curva assume a forma de um sino é denominada distribuição normal. A Figura 8-2 mostra a imagem de uma distribuição normal. Neste exemplo, a variável é o número de horas de duração que se espera que uma lâmpada fabricada pela empresa Lights Out tenha (você gostaria de estar no lugar da pessoa que coletou os dados para fazer esse teste?).

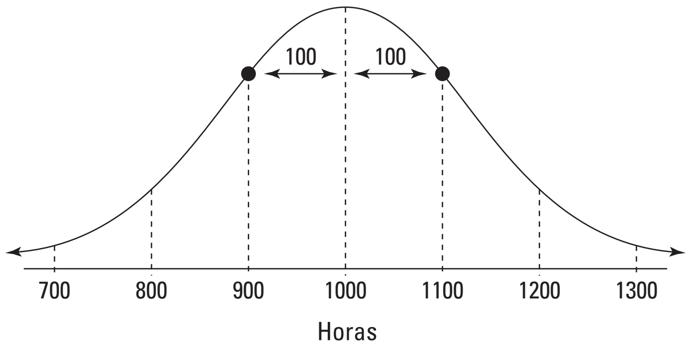


Figura 8-2: A distribuição da vida útil de lâmpadas fabricadas pela Lights Out.

Caracterizando a distribuição normal

Toda curva em forma de sino (distribuição normal) tem certas propriedades. Você pode utilizar essas propriedades para ajudar-lhe a determinar a posição relativa de qualquer resultado individual dentro da distribuição. A seguir, confira uma lista de propriedades compartilhadas por toda distribuição normal. Cada uma destas propriedades será explicada com mais detalhes nas seções a seguir.

- A forma da curva é simétrica.
- Possui uma elevação em seu centro e pontas que vão tanto para a direita quanto para a esquerda.
- A média encontra-se exatamente no meio da distribuição. A média da população é designada pela letra grega μ.

- A média e a mediana têm o mesmo valor graças à simetria da curva.
- O desvio padrão representa a distância mais comum entre a média e todos os demais dados. O desvio padrão da população é designado pela letra grega σ.
- Cerca de 95% dos valores se encontram dentro de dois desvios padrões da média.

Descrevendo a forma e o centro

Toda distribuição normal é simétrica, o que significa que se você dobrá-la exatamente ao meio, suas duas partes serão como imagens refletidas por um espelho. Devido à simetria da curva, a média (o ponto de equilíbrio) e a mediana (o ponto exato em que metade dos dados se encontram) são iguais e aparecem no centro da distribuição. A vida útil das lâmpadas mostradas na Figura 8-2 possui uma distribuição normal, com uma média (e mediana) de 1.000 horas (veja o Capítulo 5, para mais informações sobre média e mediana; veja o Capítulo 4, para mais informações sobre simetria).

Medindo a variabilidade

A forma e a média não são as únicas características importantes a serem consideradas quando se observa uma distribuição. A variabilidade dos valores também é extremamente importante, embora muitas notícias de jornais ignorem essa característica e apenas informem a média. Se voltarmos a observar a Figura 8-2, poderemos perceber que a maior parte da vida útil das lâmpadas fabricadas por aquela empresa varia de menos 700 horas até mais de 1.300 horas, com uma boa parte das lâmpadas com vida útil de 900 a 1.100 horas. Pensando como consumidor, você gostaria de se deparar com tamanha variabilidade quando for comprar um pacote de lâmpadas? Talvez não. Uma empresa concorrente (Lights Up) irá tentar fabricar lâmpadas com vida útil menos variável; a vida útil de seu produto ainda terá uma média de 1.000 horas, mas suas lâmpadas terão uma vida útil mais consistente, variando de 940 a 1.060 horas, com uma boa parte das lâmpadas com duração entre 980 e 1.020 horas (veja a Figura 8-3).

A variabilidade em uma distribuição é medida e marcada em termos numéricos como sendo um desvio padrão (veja o Capítulo 3, para conhecer as fórmulas usadas para o cálculo do desvio padrão). Em uma distribuição normal, o desvio padrão tem um significado especial, pois determina a distância da média até um local dentro da distribuição, conhecido como ponto sela. Toda distribuição normal possui dois pontos sela, cada um com a mesma distância da média. Para encontrar o ponto sela, vá do centro (a média) para a direita e para a esquerda até que a direção da curvatura se altere de uma concavidade para baixo para uma concavidade para cima.

Nas Figuras 8-2 e 8-3, os pontos sela estão marcados por dois pontos. O desvio padrão das lâmpadas da primeira empresa (veja a Figura 8-2) é de 100 horas. O desvio padrão das lâmpadas mais consistentes fabricada pela Lights Up (veja a Figura 8-3) é de 20 horas (para mais informações sobre desvio padrão, veja o Capítulo 5).

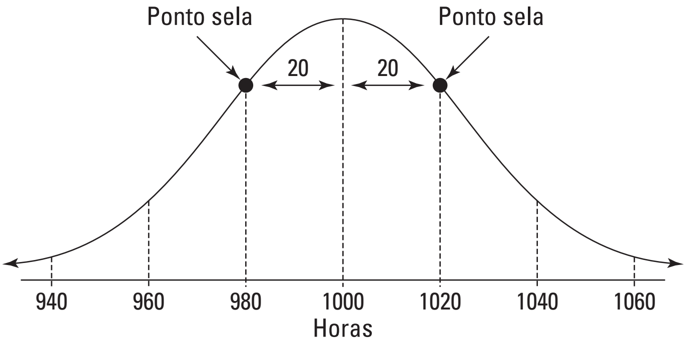


Figura 8-3: Distribuição da vida útil das lâmpadas da empresa Lights Up.

Antes de examinar os resultados, certifique-se de ter examinado a escala do eixo horizontal de qualquer distribuição e de saber qual é o desvio padrão. Dependendo da escala usada, uma distribuição pode parecer mais apertada ou mais espalhada do que realmente é. As Figuras 8-2 e 8-3, por exemplo, parecem ser semelhantes, mas suas escalas são muito diferentes. A melhor maneira de comparar a vida útil das lâmpadas fabricadas pelas duas empresas é colocando-as na mesma escala, como mostrado na Figura 8-4. Agora, você realmente pode observar o quanto a vida útil da lâmpada produzida pela Lights Out está distribuída de maneira mais espalhada do que a vida útil das lâmpadas produzidas pela Lights Up; a vida útil das lâmpadas feitas pela Lights Up está muito mais concentrada ao redor da média.

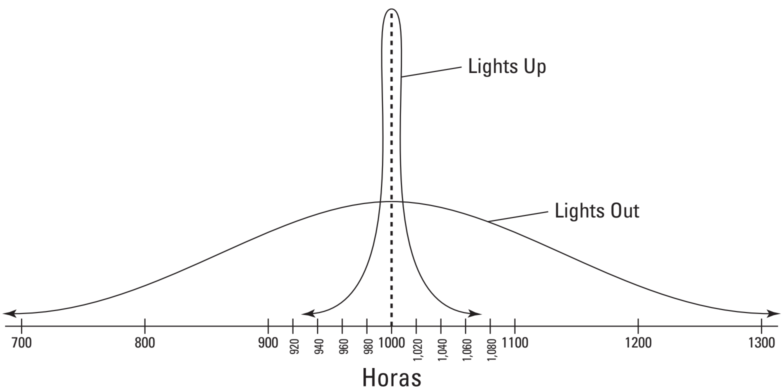


Figura 8-4: Variabilidade da vida útil das lâmpadas fabricadas pela Lights Out versus as lâmpadas da Lights Up.

Procurando pela maioria dos valores: A regra empírica

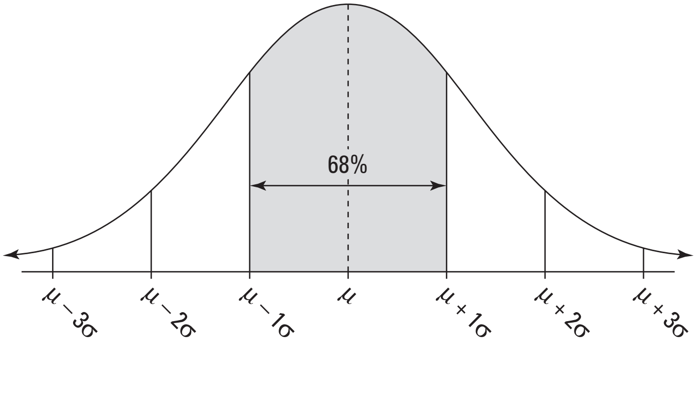
Desde que a curva de uma distribuição tenha a forma de uma montanha em seu centro – e a forma da distribuição normal decididamente satisfaz esse critério – é possível determinar a posição da maioria dos valores, usando as distâncias de 1, 2 ou 3 desvios padrões da média para estabelecer alguns marcos. A regra que lhe permite fazer isso se chama regra empírica.

A regra empírica diz que se a distribuição possui a forma de uma montanha, então:

- Cerca de 68% dos valores encontram-se dentro de 1 desvio padrão da média (ou seja, entre a média menos 1 vez o desvio padrão, e a média mais 1 vez desvio padrão). Na estatística, representamos isso da seguinte forma: $\mu \pm \sigma$.
- Cerca de 95% dos valores encontram-se dentro de 2 desvios padrões da média (ou seja, entre a média menos 2 vezes o desvio padrão, e a média mais 2 vezes o desvio padrão). Na estatística, representamos isso da seguinte forma: $\mu \pm 2\sigma$.
- Cerca de 99% (na verdade, 99,7%) dos valores encontram-se dentro de 3 desvios padrões da média (ou seja, entre a média menos 3 vezes o desvio padrão, e a média mais 3 vezes o desvio padrão). Na estatística, representamos isso da seguinte forma: $\mu \pm 3\sigma$.

Nas fórmulas da regra empírica, se você não souber a média da população e o desvio padrão, estime (substitua) o desvio padrão da população, σ, pelo desvio padrão da amostra, *s*. E você também pode estimar (substituir) a média da população, μ, pela média amostral, $\bar{x}$. Veja o Capítulo 3, para mais detalhes.

A Figura 8-5 ilustra a regra empírica. A razão para que 68% dos valores encontrem-se dentro de 1 desvio padrão da média é que a maioria dos valores em uma distribuição normal amontoam-se no centro, próximos à média (como mostra a Figura 8-5). Lembre-se que ela tem forma de sino. Adicionar 1 desvio padrão a mais para ambos os lados da média abrange mais valores, mas menos do que 30% mais (para um total de 95% dos valores), pois agora estamos considerando uma parte menor da concavidade e uma parte maior das pontas. Por fim, a adição de mais 1 desvio padrão para cada lado da média abrange o finalzinho das pontas, contabilizando 4,7% (quase todo o resto) dos valores restantes, para ir de 95% a 99,7% dos dados. A maioria dos pesquisadores fica com 95% da variação para informar seus resultados, pois, em sua visão, não parece valer a pena adicionar 3 desvios padrões para cada lado da média a fim de contabilizar os 4,7% dos valores restantes.



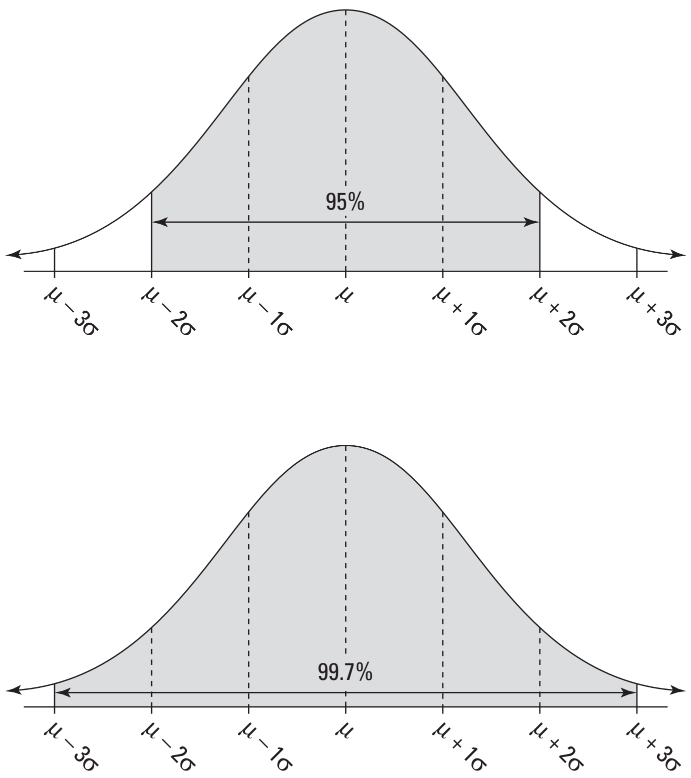


Figura 8-5: A regra empírica (68%, 95% e 99,7%).

FIQUE ATENTO

Preciso enfatizar a palavra aproximado na descrição anterior da regra empírica. Esses resultados são apenas aproximações (mas boas aproximações). Ainda neste capítulo (veja a seção "Transformando em Escore Padrão"), você verá como dar informações mais precisas com relação a porcentagem dos valores dentro de uma distribuição que estejam entre, acima ou abaixo de determinados valores. Entretanto, a regra empírica é uma regra muito importante na estatística, e o conceito de que "dois desvios padrões lhe dão 95% dos valores" é um que você vai encontrar com frequência ao longo de todo o livro.

Com o exemplo das lâmpadas fabricadas pela Light Out (veja a Figura 8-2), o desvio padrão é de 100 horas e a média é 1.000 horas. Usando a regra empírica, você pode discutir a posição relativa de certos marcos dentro dos dados. Por exemplo, de acordo com esse modelo, espera-se que cerca de 68% das lâmpadas durem de 900 a 1.100 horas (1.000 ± 100), cerca de 95% deveriam durar de 800 a 1.200 (1.000 ± 2 × 100) e 99,75 das lâmpadas deveriam durar de 700 a 1.300 horas.

DICA

Também é possível usar a simetria da distribuição normal em conjunto com a regra empírica para responder outras perguntas sobre a vida útil das lâmpadas. Por exemplo, qual a porcentagem de lâmpadas que deveriam durar 1.000 horas ou mais? A resposta é 50%, pois a média é 1.000 horas, e metade dos valores são maiores do que a média. Qual a porcentagem de lâmpadas da Lights Out deveria durar mais do que

1.200 horas (veja a Figura 8-2)? A resposta é 2,5%. Por quê? Pois 95% das lâmpadas têm vida útil de 800 a 1.200 horas e, dado o fato de que a porcentagem total sob a curva inteira tem que ser 100%, as áreas restantes das duas pontas deve somar 5%. As lâmpadas com vida útil acima de 1.200 horas estão localizadas apenas na ponta direita e, em virtude da simetria, você pode dividir 5% exatamente ao meio para obter 2,5%. Assim, uma lâmpada que dure mais do que 1.200 horas pode ser considerada um fenômeno da natureza, pois isso acontece apenas 2,5% da vezes (pelo menos com as lâmpadas da Lights Out). Com a Lights Up (veja a Figura 8-3), uma lâmpada que durasse essas mesmas 1200 horas ficaria no anonimato, pois 1.200 horas é muito mais do que 3 desvios padrões acima da média para as lâmpadas produzidas por aquela empresa (veja a Figura 8-4).

Moral da história: se você gosta de arriscar, compre as lâmpadas da Lights Out, pois você terá chances maiores de conseguir ou uma lâmpada que dure muito ou uma que dure pouco; em outras palavras, as lâmpadas da Lights Out têm mais variabilidade com relação a sua vida útil. Se você é do tipo conservador, compre as lâmpadas da Lights Up; essas são mais consistentes e possuem menos chances de causar surpresas.

FIQUE ATENTO

A regra empírica não pode ser aplicada às distribuições que não possuam uma forma de montanha em seu centro. Ainda assim, você pode aproximar ou determinar onde certos marcos se posicionam dentro dos dados, por meio de um histograma e/ou por meio dos percentis (veja os Capítulos 4 e 5, para mais informações sobre histogramas e percentis, respectivamente).

Transformando em Escore Padrão

Suponha que uma estudante de fisioterapia tenha feito um teste padronizado para obter seu certificado e tornar-se uma fisioterapeuta. Seus resultados indicaram que sua pontuação foi de 235. No entanto, tudo o que você sabe é que as pontuações para esse teste tiveram uma distribuição normal. Será que a pontuação dessa aluna foi boa, ruim ou ficou no meio do caminho? É impossível responder a essa pergunta sem saber a posição dessa estudante em relação aos outros que também fizeram a prova. Em outras palavras, é preciso determinar a posição relativa da pontuação dessa estudante na distribuição de todas as pontuações.

Analisando o desvio padrão

Existem muitas maneiras de determinar a posição relativa da pontuação dessa estudante – algumas melhores do que outras. Primeiro, podemos observar essa pontuação levando em consideração a pontuação total, que, para essa prova, foi de 300. Isso, no entanto, não compara a pontuação dela com a dos demais estudantes, apenas compara sua pontuação com a pontuação máxima. Portanto, você continua sem saber a posição relativa da pontuação dessa aluna fictícia. Outra maneira é a de comparar a pontuação dessa aluna com a pontuação média. Suponha que a média tenha sido 250. Mas isso também não nos dá muita informação. Até aqui, o que sabemos é que a pontuação dessa aluna está 15 pontos abaixo da média, pois 250 – 235 = 15. Mas o que uma diferença de 15 pontos representa nessa situação específica?

Como você pode ver, no início deste capítulo, na seção "Endireitando a Curva em Forma de Sino", em que discute a vida útil de lâmpadas fabricadas por diferentes empresas (veja a Figura 8-4), para entender a posição relativa de qualquer valor em uma distribuição, você tem que saber qual é o desvio padrão. Suponha que o desvio padrão para essa distribuição seja igual a 5 (como o ilustrado na Figura 8-6).

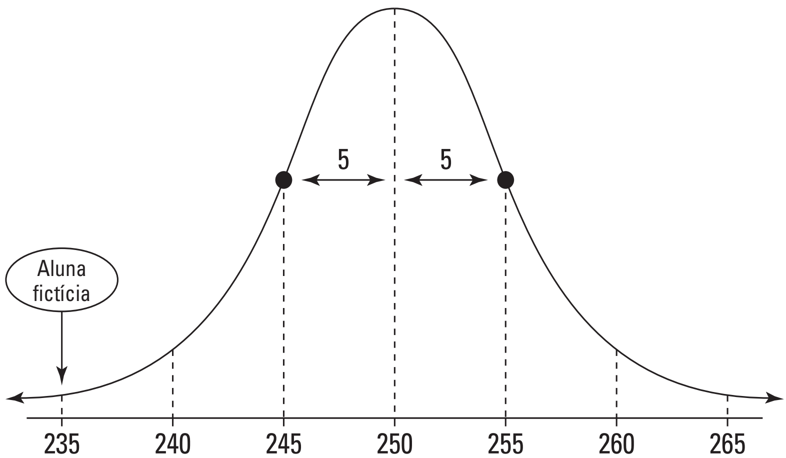


Figura 8-6: Pontuação de um teste padronizado com distribuição normal, média igual a 250 e desvio padrão igual a 5.

Uma distribuição com desvio padrão igual a 5 significa que as pontuações foram bastante próximas umas das outras e, portanto, 15 pontos abaixo da média faz, realmente, muita diferença. Nesse caso, a pontuação da aluna fictícia foi considerada bem abaixo da média, pois a diferença de 15 pontos representa 3 desvios padrões abaixo da média (cada desvio padrão é igual a 5 e $-15 \div 5 = -3$). Apenas uma minúscula fração de alunos pontuou menos do que ela (por causa da regra empírica, é possível saber que 99,7% dos outros estudantes pontuaram de 235 a 265 e que a porcentagem total para todas as possíveis pontuações é 100). Isso significa que a porcentagem dos que pontuaram fora da faixa de 235 a 265 é $100 - 99{,}7 = 0{,}3\%$. Como queremos saber a porcentagem de alunos que pontuaram menos do que nossa aluna fictícia, dividiremos 0,3% ao meio, ou seja, nesse caso, apenas 0,15 % ou 0,0015 pontuaram menos do que 235.

Agora, suponha que o desvio padrão seja um valor diferente, vamos dizer 15, mas a pontuação média continua a mesma (250). A Figura 8-7 mostra a aparência dessa distribuição.

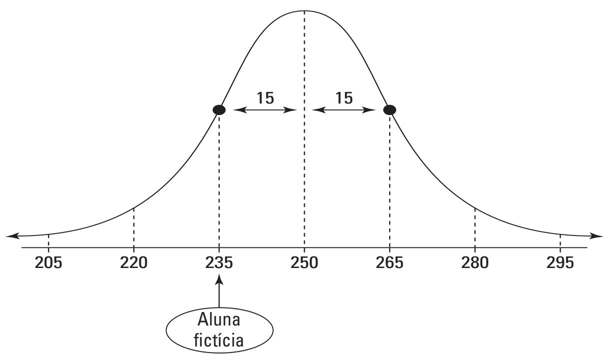


Figura 8-7: Pontuação de um teste padronizado com distribuição normal, média igual a 250 e desvio padrão igual a 15.

Um desvio padrão igual a 15 significa que as pontuações foram muito mais variáveis (ou espalhadas) do que na situação anterior. Nesse caso, o fato de que nossa aluna fictícia tenha ficado 15 abaixo da média não é tão ruim, pois esses 15 pontos representam apenas 1 desvio padrão abaixo da média (pois -15 ÷ 15 = -1). Nesse exemplo, de acordo com a regra empírica, 68% das pontuações ficaram entre 235 e 265, e metade das pontuações restantes de 32% representa os alunos que fizeram menos que 235 pontos. Portanto, nessa situação, 16% dos alunos pontuaram menos do que nossa aluna fictícia. Sua posição relativa ainda não é das melhores, mas, na segunda situação, sua posição relativa melhorou quando comparada à primeira situação. Observe que a pontuação da aluna não se alterou de um cenário para o outro; o que mudou foi a interpretação de sua pontuação em virtude da diferença nos desvios padrões.

A posição relativa de qualquer pontuação em uma distribuição depende muito do desvio padrão. As distâncias em unidades originais não significam muito sem que se saiba o desvio padrão.

Muitas vezes, na mídia, o desvio padrão não é mencionado. Nunca interprete qualquer resultado estatístico apenas o comparando à média, sem saber o desvio padrão. Os números podem estar mais longe da média do que eles aparentam.

Calculando o escore padrão

Para encontrar, mencionar e interpretar a posição relativa de qualquer valor em uma distribuição normal (tal como a pontuação da aluna de fisioterapia), é necessário transformar a variável estudada no que os estatísticos chamam de escore padrão. O escore padrão, ou escore z, é uma versão padronizada do escore original; ele representa o número de desvios padrões acima ou abaixo da média. A fórmula para calcular um escore padrão é a seguinte:

Escore padrão = (escore original – média) ÷ de $\frac{x-\mu}{\sigma}$ padrão. Ou, para utilizar uma forma mais resumida, escore padrão =

Para transformar um escore original em um escore padrão:

1. **Encontre a média e o desvio padrão da população com a qual você está lidando.**

   Por exemplo, você pode converter a pontuação da aluna do exemplo anterior de 235 para um escore padrão de acordo com cada um dos cenários discutidos na seção anterior. Para o primeiro cenário, a média é 250 e o desvio padrão é igual a 5, sendo assim, o primeiro passo já está pronto.

2. **Descubra a diferença entre o escore original e a média.**

   No caso do primeiro cenário, é possível descobrir a real distância da média ao subtrair 235 de 250 (235 – 250 = - 15).

3. **Divida o resultado obtido no segundo passo pelo desvio padrão.**

No caso do nosso exemplo, a distância é igual a -15. Transformar essa distância em números de desvios padrões significa dividi-la pelo valor do desvio padrão (-15 ÷ 5 = -3), para obter o escore padrão. No primeiro cenário (em que o desvio padrão era igual a 5), a pontuação da aluna estava a 3 desvios padrões abaixo da média.

No segundo cenário (em que o desvio padrão era igual a 15), o escore padrão da aluna é igual a (235 – 250) ÷ 15 = - 15 ÷ 15 = -1. Portanto, no segundo cenário, o seu escore é igual a 1 desvio padrão abaixo da média.

FIQUE ATENTO

Para evitar enganos durante a transformação para o escore padrão, certifique-se de fazer os passos 2 e 3 na ordem indicada.

Propriedades do escore padrão

As propriedades a seguir mostram-se muito úteis para a interpretação dos escores padrões.

- Quase todos os escores padrões (99,7% deles) se encontram entre valores de -3 a +3, graças à regra empírica.
- Um escore padrão negativo indica que o escore original estava abaixo da média.
- Um escore padrão positivo indica que o escore original estava acima da média.
- Um escore padrão igual a 0 indica que o escore original era a própria média.
- Escores que vêm de uma distribuição normal, quando padronizados, têm uma distribuição normal especial com média 0 e desvio padrão 1. Essa distribuição é chamada de distribuição normal padrão (veja a Figura 8-8).

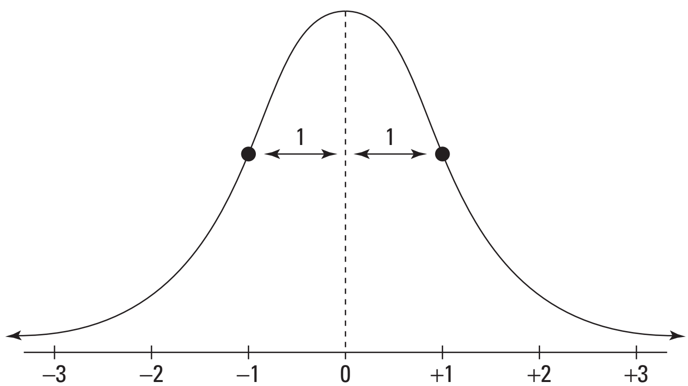


Figura 8-8: A distribuição normal padrão.

LEMBRE-SE

Os escores padrões têm uma interpretação universal, o que faz com que eles sejam tão bons. Se alguém lhe der o escore padrão, você pode interpretá-lo imediatamente. Por exemplo, um padrão escore igual a +2 significa que este escore está a 2 desvios padrões acima da média. Para interpretar um escore padrão, você não precisa saber qual é o escore original nem quais são os valores da média e do desvio padrão. O escore padrão lhe dá a posição relativa de um valor, que, na maioria dos casos, é o que lhe interessa.

A transformação dos valores originais em escores padrões não altera a posição relativa de nenhum desses valores dentro da distribuição; ela apenas altera as unidades (o mesmo que ocorre quando se transforma uma temperatura expressa em Fahrenheit para Celsius. A temperatura não se altera, o que se altera é a unidade utilizada para expressá-la). A subtração entre a média e o escore original centraliza tudo no 0. Se o escore original enquadrar-se diretamente na média, ele será convertido em um escore padrão igual a 0. A marcação aos dois lados da média ainda fica de acordo com o desvio padrão original (que pode ser de 5 em 5 de 15 em 15 e etc.). Para os escores padrões, a unidade é sempre de 1 em 1 para facilitar a interpretação. Isso significa que depois que você fizer a subtração, precisa dividir a diferença pelo valor do desvio padrão (operação parecida com a transformação de centímetro em metro, na qual você tem que dividir o valor em centímetro por 100 para obter o resultado em metros).

Comparando maças a laranjas usando o escore padrão

O escore padrão é comumente utilizado para comparar valores de diferentes distribuições, os quais seriam incomparáveis. Por exemplo, suponha que um aluno tenha se inscrito para disputar uma vaga em duas universidades diferentes (vamos chamá-las de Universidade dos Dados e Centro Universitário Estatístico) e que, para isso, ele precisa fazer uma prova de matemática em cada uma das universidades. As provas são totalmente diferentes (elas até mesmo possuem um número diferente de questões),e quando esse aluno recebe suas notas, ele quer compará-las e determinar em qual universidade sua pontuação obteve uma melhor posição relativa.

A Universidade dos Dados informou-lhe que sua pontuação foi 60 e que a distribuição de todas as pontuações é normal com média igual a 50 e desvio padrão igual a 5. O Centro Universitário Estatístico informou que sua pontuação foi 90 e que as pontuações tiveram uma distribuição normal com média igual a 80 e desvio padrão igual a 10. Em que prova esse aluno se saiu melhor? Não podemos simplesmente comparar 60 a 90, pois essas duas pontuações estão em escalas totalmente diferentes. Também não se pode afirmar que ele tenha tido o mesmo desempenho nas duas provas simplesmente pelo fato de que ele tenha ficado 10 pontos acima da média em cada uma das provas. Aqui, o desvio padrão é um fator muito importante. A maneira para se fazer uma comparação imparcial e honesta dessas pontuações é transformando-as em escores padrões, o que irá permitir que elas fiquem na mesma escala (onde a maior parte dos valores se enquadra, entre -3 e +3 com unidades de 1 em 1).

Recapitulando, a pontuação do aluno na Universidade dos Dados foi 60 com média igual a 50 e desvio padrão igual a 5. Seu escore padrão, portanto, é $(60 - 50) \div 5 = 10 \div 5 = +2$, o que significa que sua posição relativa no exame de admissão da Universidade dos Dados é de 2 desvios padrões acima da média. Já no Centro Universitário Estatístico, sua pontuação no teste de admissão foi 90 com média igual a 80 e desvio padrão igual a 10. Seu escore padrão, portanto, é (90 – 80) ÷ 10 = 10 ÷ 10 = +1, o que significa que sua posição relativa no exame de admissão para o Centro Universitário Estatístico é de apenas 1 desvio padrão acima da média. Relativamente falando, a sua nota na prova da Universidade dos Dados não foi tão alta quanto sua nota na prova do Centro Universitário Estatístico, mas, ainda assim, o aluno demonstrou melhor desempenho na prova de admissão para a Universidade dos Dados.

Não compare resultados de diferentes distribuições sem antes transformar tudo em escores padrões. O escore padrão permite-lhe realizar uma comparação justa e imparcial dentro de uma mesma escala.

Usando o Percentil para Calcular Resultados

Os percentis são utilizados em uma variedade de formas visando sempre a comparação e determinação da posição relativa. O peso, o comprimento e a circunferência da cabeça de bebês, por exemplo, são, geralmente, informadas e interpretadas em forma de percentis. Os percentis também são usados pela empresas para que estas saibam qual sua posição em vendas, lucros e satisfação do cliente com relação a seus concorrentes (para mais detalhes sobre percentis, veja o Capítulo 5).

Suponha que uma estudante de fisioterapia tenha feito uma prova para conseguir seu certificado e obteve uma pontuação igual a 235. Ela está tentando descobrir o que essa pontuação realmente significa. Sabe-se que as pontuações da prova apresentam uma distribuição normal com média igual a 250 e desvio padrão igual a 15, o que significa que seu escore padrão para essa prova é -1 (ou seja, um desvio padrão abaixo da média). Sua nota ficou abaixo da média, mas ainda é o suficiente para que ela consiga seu certificado? Todos os anos, o teste para a obtenção do certificado é diferente, portanto, a nota de corte para essa prova também se altera. No entanto, os responsáveis pela prova passam 60% dos estudantes que conseguiram as melhores notas e reprovam os 40% que obtiveram as menores notas. Conhecendo essa informação, será possível saber se nossa aluna passou no teste? E qual será a nota de corte para a prova desse ano? Os percentis é que vão ajudar nossa estudante a revelar o mistério de seu escore padrão e aliviar sua ansiedade.

Se os responsáveis pela prova sempre passam 60% dos estudantes que tiraram as maiores notas e reprovam os 40% que tiverem as menores notas, isso significa que a nota de corte para passar/reprovar nessa prova está no 40° percentil (lembre-se que o percentil refere-se às porcentagens abaixo). Em que percentil está a nota dessa estudante? Se voltarmos à Figura 8-7 e utilizarmos a regra empírica, sabemos que sua pontuação está a 1 desvio padrão abaixo da média. Assim, cerca de 68% dos resultados se encontram entre 235 e 265 (com um desvio padrão da média) e o resto (100% - 68% = 32%) se encontram fora dessa variação. Metade do que encontram-se fora da variação de ± 2 desvios padrões (ou seja, 32% ÷ 2 = 16%) pontuaram menos do que 235. Isso significa que a pontuação dessa estudante se encontra no percentil 16° (aproximadamente) e, portanto, ela não passou no teste. Para passar, ela precisaria ter ficado pelo menos no 40°.

A regra empírica pode apenas lhe ajudar a determinar os percentis; se você notar, a pontuação da aluna foi escolhida propositadamente por mim para que se enquadrasse exatamente em uma das marcações do gráfico. Mas, e se sua nota tivesse ficado em algum lugar entre as marcações do gráfico? Não tema, a tabela normal padrão está aqui. A Tabela 8-1 mostra os percentis correspondentes (a porcentagem abaixo) para qualquer escore padrão entre – 3,4 e + 3,4. Ela abrange cerca de 99,7% das situações com as quais você pode vir a se deparar. Observe que conforme o escore padrão aumenta, os percentis também aumentam. Note também que o escore padrão igual a 0 está no percentil 50°, o mesmo que a mediana (veja o Capítulo 5, para mais sobre médias e medianas).

Tabela 8-1 Escores Padrões e Percentis Correspondentes da Distribuição Normal Padrão

Escore Padrão	*Percentil*	*Escore Padrão*	*Percentil*	*Escore Padrão*	*Percentil*
-3,4	0,03%	-1,1	13,57%	+1,2	88,49%
-3,3	0,05%	-1,0	15,87%	+1,3	90,32%
-3,2	0,07%	-0,9	18,41%	+1,4	91,92%
-3,1	0,10%	-0,8	21,19%	+1,5	93,32%
-3,0	0,13%	-0,7	24,20%	+1,6	94,52%
-2,9	0,19%	-0,6	27,42%	+1,7	95,54%
-2,8	0,26%	-0,5	30,85%	+1,8	96,41%
-2,7	0,35%	-0,4	34, 46%	+1,9	97,13%
-2,6	0,47%	-0,3	38,21%	+2,0	97,73%
-2,5	0,62%	-0,2	42,07%	+2,1	98,21%
-2,4	0,82%	-0,1	46,02%	+2,2	98,61%
-2,3	1,07%	0,0	50,00%	+2,3	98,93%
-2,2	1,39%	+0,1	53,98%	+2,4	99,18%
-2,1	1,79%	+0,2	57,93%	+2,5	99,38%
-2,0	2,27%	+0,3	61,79%	+2,6	99,53%
-1,9	2,87%	+0,4	65,54%	+2,7	99,65%
-1,8	3,59%	+0,5	69,15%	+2,8	99,74%
-1,7	4,46%	+0,6	72,58%	+2,9	99,81%

Tabela 8-1 (continuação)

-1,6	5,48%	+0,7	75,80%	+3,0	99,87%
-1,5	6,68%	+0,8	78,81%	+3,1	99,90%
-1,4	8,08%	+0,9	81,59%	+3,2	99,93%
-1,3	9,68%	+1,0	84,13%	+3,3	99,95%
-1,2	11,51%	+1,1	86,43%	+3,4	99,97%

Para usar a Tabela 8-1 a fim de encontrar um percentil, você precisa primeiro converter o valor original em um escore padrão. O que é mais fácil do que ficar carregando uma tabela para toda distribuição normal possível com todas as possíveis média e desvios padrões, certo? É por isso que a distribuição normal padrão é tão especial; ela utiliza uma escala padronizada, em que uma única tabela se encaixa perfeitamente.

Para calcular um percentil quando os dados têm uma distribuição normal:

1. **Converta o valor original em um escore padrão descobrindo a diferença entre o valor original e a média e depois a divida pelo valor do desvio padrão. A fórmula utilizada para esse cálculo é: $\frac{x-\mu}{\sigma}$.**
2. **Use a tabela 8-1 e encontre o percentil correspondente ao escore padrão.**

Você já sabe que a pontuação da estudante de fisioterapia do nosso exemplo está no percentil 16°. Suponha que outro aluno que também fez o mesmo teste tenha conseguido 260 pontos. Será que ele passou? Para descobrir, vamos converter esse valor em um escore padrão e encontrar o percentil correspondente. Seu escore padrão é (260 – 250) ÷ 15 = 10 ÷ 15 = 0,67. Usando a Tabela 8-1, verificamos que sua pontuação ficou entre o percentil 72,58° e 72,80° (para sermos conservadores, vamos optar pelo percentil mais baixo). Em qualquer um dos casos, esse aluno passou no teste, pois seu desempenho foi melhor do que o de pelo menos 72% dos outros estudantes e, além disso, seu resultado ficou acima do percentil 40°, ultrapassando o percentil da nota de corte.

Não seria ótimo não ter que transformar as pontuações de todos os alunos em um escore padrão apenas para determinar se ele ou ela passou? Por que não apenas encontrar a nota de corte, em sua unidade original, e comparar todos os resultados a ela? Você sabe que a nota de corte está no percentil 40°. Usando a tabela 8-1, podemos ver que o escore padrão mais próximo que corresponde ao 40° percentil é de -0,3 (use a tabela de baixo para cima). Isso significa que a nota de corte está a 3 desvios padrões da média. Mas, quanto é a nota de corte em unidade padrão? Você pode usar a fórmula do escore padrão de trás para frente para encontrar o valor original.

Lembre-se de que a fórmula para transformar o valor original (x) em um escore padrão (chamado Z) é $Z = \frac{x - \mu}{\sigma}$. Com um pouco de álgebra, é possível reescrever esta equação para que você possa transformar um escore padrão (Z) de volta ao valor original (x). Esta fórmula ficará assim: $x = Z\sigma + \mu$.

Para transformar um escore padrão em um valor original:

1. **Encontre a média e o desvio padrão para a população com a qual você está lidando.**

2. **Multiplique o escore padrão (Z) pelo valor do desvio padrão.**

3. **Some o valor da média ao resultado encontrado no segundo passo.**

Voltando ao exemplo, sabe-se que a nota de corte da prova é -0,3, portanto $Z = -0,3$. Você também sabe que a média das pontuações é $\mu = 250$ e que o desvio padrão $\sigma = 15$. Para transformar o escore padrão no valor original, você faz $x = -0,3 \times 15 + 250 = -4,5 + 250 = 245,5$ (ou 246). Portanto, a nota de corte para essa prova era de 246. Todos que pontuaram abaixo de 246 foram reprovados e todos que pontuaram acima de 246 foram aprovados.

Mas, e se os valores não tiverem uma distribuição normal? Você ainda pode calcular os percentis, mas terá que fazer manualmente ou com o auxílio de algum programa de computador (como o Microsoft Excel).

Para encontrar o k° percentil quando os dados não se apresentam em uma distribuição normal:

1. **Ordene os valores do menor para o maior.**

2. **Designe o tamanho do conjunto de dados pela letra *n*. Multiplique *k* por cento vezes *n* e arredonde para o número inteiro mais próximo.**

3. **Conte os dados até alcançar o ponto identificado no passo número 2. Este será o *k*° em seu conjunto de dados.**

   Por exemplo, suponha que você tenha o seguinte conjunto de dados: 1, 6, 2, 5, 3, 9, 3, 5, 4, 5 e queira saber o 90° percentil.

   1º passo: Ordene os dados: 1, 2, 3, 3, 4, 5, 5, 5, 6, 9.

   2º passo: $n = 10$, $k = 90\%$, e k por cento vezes n é 0,90 vezes 10 = 9

   3º passo: Isso significa que o 9° número (do menor para o maior), que nesse caso é o número 6, é o 90° percentil. Cerca de 90% dos valores são menores do que 6 e 10% são maiores (veja o Capítulo 5, para discussões e exemplos de percentis).

Capítulo 9

Atenção: Os Resultados Amostrais Variam

Neste Capítulo

- Investigando a curva com forma de sino
- Usando a regra empírica para a maioria dos casos
- Calculando e interpretando os escores padrões

As estatísticas geralmente nos são apresentadas da seguinte forma: "um em cada dois casamentos acabam em divórcio", "quatro em cinco dentistas recomendam Trident" ou "a expectativa de vida média para uma mulher nascida no ano 2000 é de 80 anos". As pessoas ouvem essas estatísticas e acabam presumindo que esses resultados aplicam-se a elas. Uma pessoa comum pode supor que, por exemplo, suas chances de se divorciar são de 50% , que seu dentista provavelmente irá lhe recomendar o chiclete Trident e que ela e seu marido (se eles não tiverem se divorciado antes) tiverem uma filha no ano 2000, eles podem esperar que ela viva até 80 anos.

Mas essas estatísticas não deveriam estar acompanhadas por uma indicação de "para mais ou para menos", referindo-se ao fato de que os resultados variam? Com certeza! Mas isso acontece? Não com muita frequência. A verdade é que a menos que os pesquisadores sejam capazes de realizar um censo para conseguir seus resultados (ou seja, coletar dados de cada um dos membros de uma população), esses resultados vão variam de amostra para amostra e essa variabilidade pode ser maior do que você imagina! A questão é quanto se deve esperar que um resultado estatístico varie? Você espera (ou, talvez, até suponha automaticamente) que eles não variem muito e que você possa aplicá-los a quase todo mundo. Mas isso sempre acontece? Claro que não; a variabilidade em um resultado estatístico depende de uma série de fatores, fatores esses que serão discutidos neste capítulo.

Esperando que os Resultados Amostrais Variem

Um dia desses, eu estava assistindo um comercial na TV sobre uma bebida que substitui a refeição para auxiliar na perda de peso. O comercial mostrava uma história inspiradora de uma mulher que havia perdido 22 quilos em seis meses (e por mais de um ano tinha conseguindo manter o peso). Durante seu testemunho, uma mensagem aparecia por alguns segundos na parte de baixo da tela dizendo, "resultados atípicos".

Isso me levou à pergunta: "O que é típico?" Ou seja, quantos quilos se espera perder utilizando aquele produto por seis meses ou quantos meses são necessários para perder 22 quilos? Você sabe que, independente dos resultados obtidos, eles variam de pessoa para pessoa. Mas aquele comercial estava tentando lhe induzir a acreditar que você iria perder 22 quilos em seis meses (embora a mensagem na parte de baixo da tela lhe dissesse que não). Como seria bom se o fabricante desse produto lhe dissesse o quanto os resultados podem variar e como seria ótimo se o comercial apresentasse os resultados obtidos de uma amostra de pessoas e não apenas de uma única pessoa.

Testemunhos são atraentes, mas não têm significado estatístico!

Suponha que você esteja tentando estimar a proporção de pessoas nos Estados Unidos que aprovam o presidente. Se você entrevistar uma amostra aleatória de 1.000 cidadãos americanos, perguntando se eles aprovam o trabalho feito pelo presidente, você conseguirá um resultado amostral (por exemplo, 55% de aprovação). Você não deveria informar que 55% de toda a população americana aprovam o presidente, pois seus resultados basearam-se em uma amostra de apenas 1.000 pessoas.

Se você selecionar aleatoriamente outra amostra com 1.000 pessoas e fizer a mesma pergunta, provavelmente você irá obter resultados diferentes. Na realidade, dado o fato de que a população americana é imensa e não são todos que gostariam de dar sua opinião sobre o presidente, 100 amostras aleatória diferentes de 1.000 pessoas – todas tiradas da mesma população e questionadas sobre a mesma coisa – revelariam 100 diferentes resultados. Portanto, como você deve relatar seus resultados amostrais? Algumas medidas de quanto se espera que os resultados variem devem fazer parte do pacote.

Sempre espere que os resultados amostrais variem de amostra para amostra. Não acredite logo de cara em uma estatística e nem tente aplicá-la sem alguma indicação da variação esperada para os resultados.

Medindo a Variabilidade em Resultados Amostrais

Você deve estar se perguntando como avaliar o quanto uma amostra estatística vai variar sem ter que selecionar todas as amostras possíveis e observar suas estatísticas, e já deve estar querendo fazer um censo, não é mesmo? Felizmente, graças a alguns resultados estatísticos importantes (especialmente o teorema do limite central), é possível encontrar o quanto se deve esperar que as médias ou proporções amostrais variem sem ter que coletar todas as amostras possíveis (que alívio!). O teorema do limite central, de um modo geral, diz que a distribuição de todas as médias (ou proporções) amostrais são normais, desde que o tamanho das amostras seja grande o bastante. E o que é mais impressionante, o teorema do limite central não se importa com a aparência da distribuição da população original. Como isso pode acontecer? Você precisa ter uma amostra grande o suficiente e precisa saber algumas características da população original (como a média e o desvio padrão). E, então, uma mágica da teoria estatística começa a aparecer.

Erros padrões

A variabilidade em médias amostrais (ou proporções) é medida em termos de erros padrões. O erro padrão tem o mesmo conceito básico de um desvio padrão; ambos representam uma distância típica da média. Vejamos como o erro padrão se diferencia do desvio padrão. Os valores da população original desviam-se um dos outros graças a um fenômeno natural (as pessoas têm diferentes alturas, pesos,etc.), portanto temos o nome desvio padrão para medir sua variabilidade. As médias amostrais variam por causa do erro que ocorre por não sermos capazes de realizar um censo e termos que coletar amostras, portanto temos o nome erro padrão para medir a variabilidade das médias amostrais (veja o Capítulo 5,para mais saber sobre desvio padrão. Falarei mais sobre como interpretar os erros padrões nesta seção. Para detalhes sobre fórmulas do erro padrão, veja o Capítulo 10).

Aqui vai um exemplo: O U.S. Bureau of Labor Statistics tenta rastrear em que as pessoas gastam seu dinheiro a cada ano através da Consumer Expenditure Survey (Pesquisas dos Gastos do Consumidor). O bureau seleciona uma amostra de residências e pede a cada uma delas que informe seus gastos (o erro sistemático é uma questão a ser considerada aqui). O tamanho comum de suas amostras é de 7.500 pessoas. A tabela 9-1 mostra alguns dos resultados obtidos de pesquisa realizada em 2001. Esta tabela não apenas inclui o valor médio gasto pelas pessoas da amostra em vários itens (as médias amostrais para cada item), mas também inclui o erro padrão para cada uma dessas médias amostrais.

Tabela 9-1 Média de Gastos Domésticos Anuais de Residência Americanas em 2001

Despesa	*Média*	*Erro Padrão*
Alimentação (comendo em casa)	$3.085,52	$42,30
Alimentação (comendo fora)	$2.235,37	$ 38,35
Telefone	$914,41	$ 9,69
Combustível (para veículos)	$1.279,37	$12,88
Materiais de leitura	$141,00	$2,99

Você pode interpretar os resultados da Tabela 9-1 fazendo comparações relativas. Por exemplo, note que cerca de 42% de toda a média refere-se a gastos com alimentação fora de casa, dado que $2.235,27 ÷ ($3.085,52

+ $ 2.235,37) = $ 2.235,37 ÷ $5.320,89 = 0,42 ou 42%. Os erros padrões para a média de gastos com alimentação são maiores do que os de outras despesas da lista, porque as despesas com alimentação variam muito de residência para residência. No entanto, você deve estar se perguntando o porquê de os erros padrões para os gastos com alimentação não serem maiores do que os mostrados na Tabela 9-1. Lembre-se que o erro padrão lhe fala o quanto você deve esperar de variabilidade em uma média, caso você venha analisar outra amostra. Se o tamanho da amostra é grande, a média não irá variar muito. E você sabe que o governo nunca utiliza pequenas amostras!

Uma lista com os erros padrões para as médias amostrais não é algo que você verá com frequência na mídia. Entretanto, você pode (e deve, quando os resultados forem muito importantes para você) ir mais fundo e descobrir os erros padrões. A melhor coisa a fazer é buscar os artigos científicos e procurar pelos erros padrões nesses artigos.

Distribuições Amostrais

Uma lista de todos os valores que uma média amostral pode ter e a frequência com que tais valores ocorrem é chamada distribuição amostral da média amostral. Uma distribuição amostral, como qualquer outra distribuição, tem um formato, um centro e uma medida de variabilidade (nesse caso, o erro padrão) (veja o Capítulo 4, para mais informações sobre distribuições).

Segundo o teorema do limite central, se as amostras forem grandes o bastante, a distribuição de todas as médias amostrais possíveis terá uma distribuição normal, em forma de sino, com a mesma média da população original (veja o Capítulo 3, para mais informações sobre a distribuição normal). Isso ocorre porque as médias amostrais ficam agrupadas ao redor do valor médio geral, que é a média da população. Valores altos em uma amostra são compensados pelos baixos valores que também aparecem na amostra, perfazendo uma média. A variabilidade na distribuição amostral é medida em termos de erros padrões. Um dos benefícios de utilizar uma média para chegar a uma estimativa (ao invés do valor total ou de um único valor) é que a variabilidade nas médias amostrais diminui conforme o tamanho da amostra aumenta (características semelhantes também se aplicam à distribuição amostral da proporção da amostra, no caso dos dados categorizados [sim/não] coletados por meio de pesquisas de opinião e enquetes).

Usando a regra empírica para interpretar os erros padrões

Devido ao fato de que a distribuição amostral das médias amostrais (ou proporções amostrais) é normal (formato de montanha), você pode utilizar a regra empírica para ter uma ideia do quanto um dado resultado pode variar, desde que o tamanho da amostra seja grande o bastante (veja o Capítulo 8, para a cobertura completa a respeito da regra empírica).

Aplicada às médias ou proporções amostrais, a regra empírica diz que você deve esperar:

- Que cerca de 68% das médias amostrais encontrem-se dentro de 1 erro padrão da média da população.
- Que cerca de 95% das médias amostrais encontram-se dentro de 2 erros padrões da média da população.
- Que cerca de 99,7% das médias amostrais encontram-se dentro de 3 erros padrões da média da população.
- Que valores semelhantes para dados categorizados (sim/não): 68%, 95% ou 99,7% das proporções amostrais encontram-se dentro de 1, 2 ou 3 erros padrões, respectivamente, da proporção da população.

O que a regra empírica fala sobre o quanto você pode esperar que uma dada média amostral varie? Tenha em mente que 95% das médias amostrais devem se encontrar dentro de 2 erros padrões da média da população e é você quem deve estimar a média da população. Assim, se sua estimativa for realmente uma variação que inclua sua média amostral de aproximadamente, 2 erros padrões, sua estimativa estará correta 95% das vezes (o número de erros padrões somados ou subtraídos é chamado de margem de erro. Para saber mais sobre a margem de erro, veja o Capítulo 10).

Considere este exemplo: segundo o U.S. Bureau of Labor Statistics, em média, as residências familiares em 2001 continham 2,5 pessoas (das quais, 0,7 eram crianças com menos de 18 anos e 0,3 eram pessoas acima dos 65) e 1,9 veículos (desculpe-me, mas o erro padrão para esse dado não foi informado). Voltando à tabela 9-1, podemos notar que os gastos médios com telefone em 2001 para a amostra de 7.500 residências foi de $914,41 por residência. Quanto se espera que esses resultados variem de amostra para amostra (ou seja, se amostras diferentes, cada uma com 7.500 residências, tivessem sido selecionadas da mesma população)? O erro padrão para os gastos com telefone para essa amostra foi de $9,69. Isso significa que 95% dos gastos médios com telefone da amostra devem se encontrar dentro de 2 erros padrões, 2 × $9,69 (ou $19,38) para cima ou para baixo da real média populacional. Isso demonstra o quanto a média com os gastos de telefone podem variar quando o tamanho amostral é de 7.500 lares.

No exemplo anterior, não estamos dizendo que 95% de todas as residências da população estudada têm gastos com telefone igual a essa variação. Ao invés disso, estamos dando uma estimativa para a média de gastos com telefone em todas as residências da população. O gasto médio com telefone deveria ser, na verdade, um número único, mas, pelo fato de que não se pode chegar ao número real, nós o estimamos usando essa variação de valores.

Podemos também usar a regra empírica para fornecer uma estimativa aproximada da média de gastos com telefone de todas as residências nos Estados Unidos (e não apenas para a amostra de 7.500 residências). Utilizando a segunda propriedade da regra empírica, pode-se esperar que a média de gastos com telefone de todas as residências nos Estados

Unidos seja aproximadamente \$914,41, aproximadamente 2 × \$9,69 = \$19,38. Esse tipo de estimativa estará correta em 95% das amostras selecionadas (e espera-se que a amostra coletada na Consumer Expenditure Survey seja uma delas). Usando o jargão estatístico apropriado, isso significa que você pode estimar, com cerca de 95% de certeza, que a média de gastos com telefone por anos para todas as residências nos Estados Unidos fica em algum lugar entre \$914,41 - \$19,38 e \$914,41 + \$19,38, ou seja, entre \$895,03 e \$933,79. Caso você queria ter aproximadamente, 99,7% de certeza, ao invés de 95%, você deve somar e subtrair 3 erros padrões.

Esse tipo de resultado, envolvendo uma estatística mais ou menos de um determinado número de erros padrões, é chamado de *intervalo de confiança* (para mais sobre isso, veja o Capítulo 11). O valor somado ou subtraído é chamado de margem de erro.

Não é muito comum se ver o valor da margem de erro para uma média amostral envolvendo dados quantitativos (tais como renda familiar, preço de residências ou valores da bolsa de ações) sendo informado pela mídia, ainda que a margem de erro devesse sempre estar presente para que as pessoas pudessem avaliar a precisão dos resultados! Já com os dados coletados por meio de pesquisas de opinião e enquetes (dados categorizados informados como proporções) você vê, com frequência, a margem de erro, que está diretamente relacionada ao erro padrão (ver Capítulo 10). Por que essa discriminação? Não sei dizer com certeza.

Especificidades do teorema do limite central

Observe que os resultados que aplicam a regra empírica na seção anterior nos dão uma maneira aproximada de interpretar 1, 2 ou 3 erros padrões e dizem o que esperar em termos de média amostral (ou proporção amostral). Esses resultados ocorrem, na verdade, em virtude do teorema do limite central (TLC). Os estatísticos amam o TLC; sem ele, eles não teriam seus empregos. O TLC permite-lhes realmente dizer onde, aproximadamente, eles podem esperar que os resultados das amostras se encontrem, sem precisar ter que observar todas as amostras possíveis coletadas de uma determinada população. O teorema também lhes fornece uma fórmula para calcular os erros padrões, além de fornecer algumas informações mais específicas relacionadas a quais porcentagens das médias amostrais irão se encontrar entre qualquer número de erros padrões (e não apenas 1, 2 ou 3).

O teorema do limite central diz que para qualquer população com média μ e com desvio padrão σ:

- A distribuição de todas as médias amostrais, $\overline{x}$, é aproximadamente normal para tamanhos amostrais suficientemente grandes. Isso significa que é possível utilizar a distribuição normal para responder às perguntas ou tirar conclusões sobre a média amostral (veja o Capítulo 8 para maisinformações sobre a distribuição normal).

- Quanto maior o tamanho amostral (n), mais próxima de uma distribuição normal será a distribuição das médias amostrais. (a maioria dos estatísticos concorda que se n for pelo menos igual a 30, fará um bom trabalho na maioria dos casos).

  A média da distribuição das médias amostrais também é μ.

- O erro padrão das médias amostrais é $\frac{\sigma}{\sqrt{n}}$. E diminui conforme n aumenta.
- Se os dados originais possuem uma distribuição normal, a média amostral sempre irá ter uma distribuição normal exata, independente do tamanho amostral.

Se o desvio padrão da população, σ, for desconhecido (o que ocorrerá na maioria das vezes), você pode estimá-lo usando s, o desvio padrão da amostra, para calcular o erro padrão na fórmula anterior. Mais sobre esse assunto no Capítulo 12.

Observe que o TLC diz que, mesmo se os dados originais não estiverem distribuídos de modo normal, a distribuição da média amostral será normal, desde que o tamanho da amostra seja grande o suficiente. Isso, mais uma vez, ocorre devido ao efeito mediador.

Verificando as notas de matemática do ACT

Considere as notas da prova de matemática do ACT (exame realizado para a admissão em faculdades americanas) para os alunos e alunas do ensino médio nos Estados Unidos em 2002 (essa é uma situação em que a média populacional e o desvio padrão são conhecidos, pois todas as provas feitas em 2002 foram corrigidas e arquivadas). A nota média para a prova de matemática entre os estudantes do sexo masculino foi de 21,2 com um desvio padrão de 5,3 e os do sexo feminino fizeram uma média de 20,1 (com um desvio padrão de 4,8). Estudos anteriores demonstraram que as notas do ACT têm uma distribuição normal aproximada. A Figura 9-1 mostra as distribuições das notas para os estudantes dos dois sexos, segundo a informação dada anteriormente. Em cada caso, o tamanho de toda a população que prestou o teste (meninos e meninas que fizeram a prova) é de cerca de um milhão de estudantes.

Usando a regra empírica (veja o Capítulo 8), cerca de 95% dos meninos pontuaram de 10,6 a 31,8 e cerca de 95% das meninas pontuaram de 10,5 a 29,7 na prova de matemática do ACT. As notas de meninos e meninas são bastante compatíveis.

Suponha que você esteja interessado nas notas médias de amostras de 100 alunos retirados da população total de 500.000 meninos que fizeram a prova em 2002. Por quê? Talvez você tenha 100 alunos em sua sala de aula e queira saber como eles se saíram comparados com todas as outras possíveis salas com o mesmo tamanho. Como será a aparência da distribuição de todas as médias amostrais? Segundo o TLC, ela será uma distribuição normal com a mesma média (21,2) e o erro padrão será 5,3 dividido pela raiz quadrada de 100 (pois nessa amostra hipotética, n = 100). Assim, o erro padrão para essa amostra é $\frac{5.3}{\sqrt{100}}$ = 5,3 ÷ 10 = 0,53.

A Figura 9-2 mostra a aparência da distribuição amostral das médias amostrais para amostras de 100 meninos.

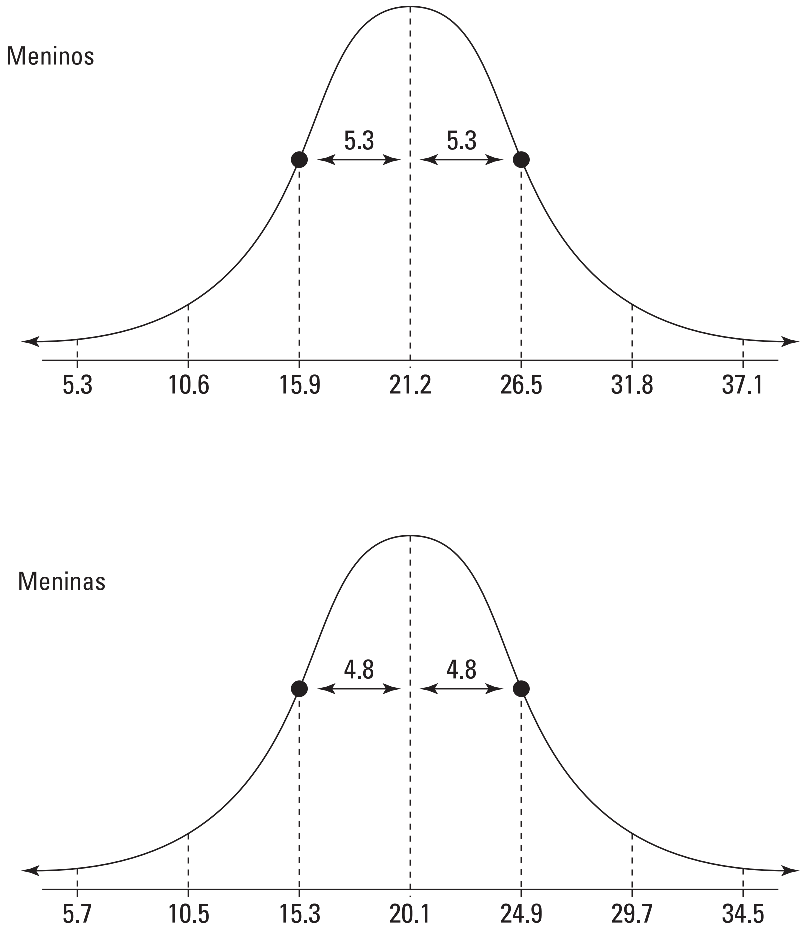


Figura 9-1: Notas da prova de matemática do ACT realizada por meninos e meninas do ensino médio em 2002 nos Estados Unidos.

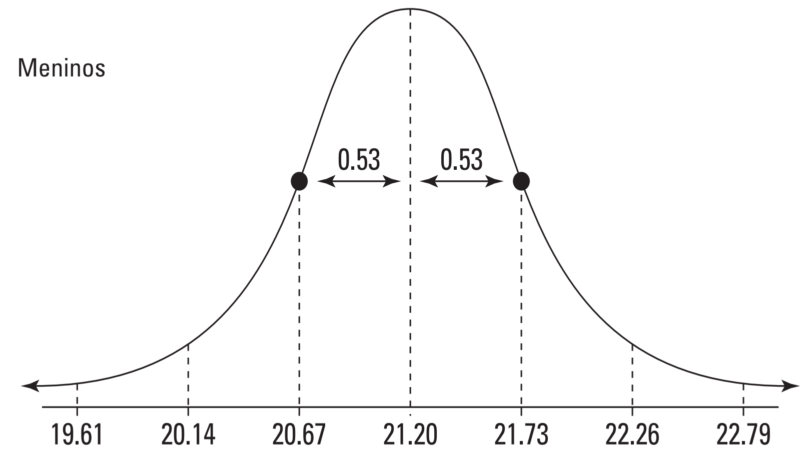


Figura 9-2: Média das notas tiradas na prova de matemática do ACT realizada por uma amostra de 100 meninos em 2002 nos Estados Unidos.

Observe na Figura 9-2 que o erro padrão da média amostral é bem menor do que o desvio padrão das notas originais mostrado na Figura 9-1. Isso ocorre, pois cada média amostral na Figura 9-2 contém informações de 100 alunos, comparadas a cada nota individual na Figura 9-1, que contém informações de apenas um único aluno. As médias amostrais não irão variar tanto quanto as notas individuais. É por isso que, para estimar a média populacional, a média amostral é muito melhor do que apenas uma nota individual (ou um testemunho).

A Figura 9-3 mostra o que acontece à distribuição da amostra da média amostral quando o tamanho da amostra aumenta para 1.000 alunos. O erro padrão se reduz para 5,3, dividido pela raiz quadrada de 1.000, ou seja, 5,3 ÷ 31,62 = 0,17. O erro padrão para a Figura 9-3 é menor do que o erro padrão na Figura 9-2, pois cada média amostral na Figura 9-3 baseia-se em 1.000 estudantes e contém ainda mais informações do que as médias amostrais mostradas na Figura 9-2 (cada uma baseada em 100 estudantes).

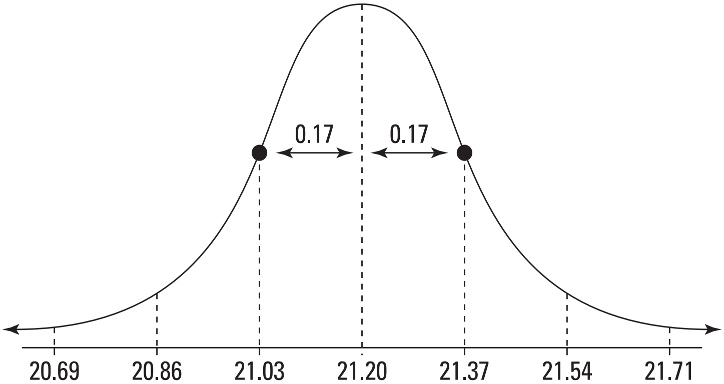


Figura 9-3: Média das notas da prova de matemática do ACT realizado por uma amostra de 1.000 meninos do ensino médio americano em 2002.

Você pode usar o teorema do limite central para responder a perguntas sobre os resultados amostrais em situações como a das notas da prova de matemática no ACT. Por exemplo, suponha que você queira saber as chances que uma amostra de 100 meninos terá de obter a nota média de 22 ou uma menor na prova de matemática do ACT. Usando a técnica mostrada no Capítulo 8, você transforma a nota 22 em um escore padrão, subtraindo-a da média populacional (21,2) e dividindo a diferença pelo erro padrão (ao invés de usar o desvio padrão). A fórmula para essa conversão é $\frac{\bar{x}-\mu}{\sigma \div \sqrt{n}}$ onde $\bar{x}$ é escore padrão da amostra (nesse caso, 22) e $\frac{\sigma}{\sqrt{n}}$ é o erro padrão. Note que σ é o desvio padrão da população (5,3). Nesse exemplo, onde o erro padrão foi calculado com sendo 5,3 ÷ √n = 0,53 (veja Figura 9-2), a média da sala igual a 22 se converte a um escore padrão de (22 – 21,2) ÷ 0,53 = 0,8 ÷ 0,53 = 1,51. Você vai querer saber a porcentagem das notas que se encontram à esquerda desse valor (em outras palavras, o percentil correspondente ao escore padrão de 1,51). Se voltarmos à tabela 8-1, no Capítulo 8, notaremos que essa porcentagem é de cerca de 93,32%.

Para o exemplo anterior, não se esqueça de, primeiramente, encontrar a diferença entre 22 e 21,2 e, depois, dividi-la por 0,53. Caso contrário, você encontrará o valor -18, que é errado.

LEMBRE-SE

Um grupo de médias sempre possui menos variabilidade do que um grupo de valores individuais. E as médias que se baseiam em tamanhos amostrais maiores têm ainda menos variabilidade do que as médias que se baseiam em tamanhos amostrais menores.

O teorema do limite central não se aplica apenas às médias amostrais. Você também pode usar o teorema do limite central para responder perguntas ou tirar conclusões a respeito das proporções populacionais, baseando-se em sua proporção amostral. As mesmas conclusões sobre a forma, o centro e a variabilidade nas médias amostrais aplicam-se às proporções amostrais. É claro que as fórmulas serão um pouco diferentes, mas os conceitos são todos iguais. Primeiramente, a proporção amostral é denotada pelo símbolo $\hat{p}$, que é igual ao número de indivíduos na amostra na categoria de interesse, dividido pelo tamanho amostral total (n).

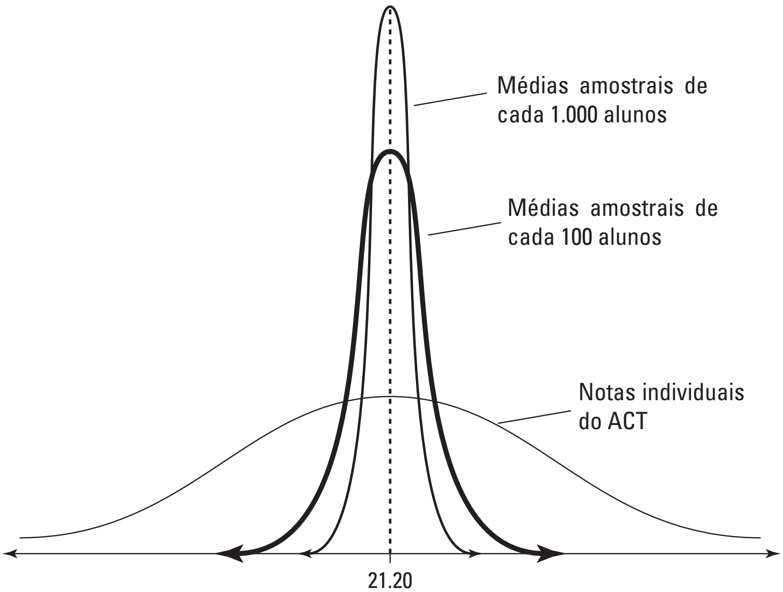


Figura 9-4: As distribuições amostrais para as notas da prova de matemática do ACT de meninos que realizaram a prova em 2002, mostrando as notas originais, as médias amostrais para amostras com 100 e com 1.000 meninos.

O teorema do limite central diz que para qualquer população de dados com p sendo a porcentagem total da população:

- A distribuição de todas as proporções amostrais possíveis ($\hat{p}$) é aproximadamente normal, desde que o tamanho da amostra seja grande o suficiente (leia o Capítulo 8, para mais informações acerca da distribuição normal).
- Quanto maior o tamanho da amostra (n), mais próxima de uma distribuição normal é a distribuição das proporções amostrais (isso significa que você pode usar a distribuição normal para responder às perguntas ou tirar conclusões sobre sua proporção amostral).
- A média da distribuição das proporções amostrais também é p.
- O erro padrão das proporções amostrais é $\sqrt{\frac{p(1-p)}{n}}$. E diminui conforme n aumenta.

Note que o erro padrão da proporção amostral realmente contém p, que é a proporção da população. Esse valor provavelmente será desconhecido; você pode estimá-lo, entretanto, com a proporção amostral, $\hat{p}$. Mais sobre esse assunto no Capítulo 12.

Qual proporção necessita da ajuda da matemática?

Você pode usar o teorema do limite central para responder as perguntas envolvendo proporções. Por exemplo, suponha que você queira saber qual a proporção de alunos no primeiro ano da faculdade que esteja precisando de ajuda em matemática. Uma pesquisa de alunos acompanha a prova do ACT todos os anos, e uma das perguntas feitas é se os alunos gostariam de ter ajuda para melhorar suas habilidades em matemática. Em 2002, 38% dos alunos que fizeram o ACT responderam sim a essa pergunta. Essa é uma situação em que a proporção da população, p, é conhecida ($p = 0{,}38$). Os dados originais nesse caso (assim como todos os dados categorizados) não possuem uma distribuição normal, pois apenas dois resultados são possíveis: sim ou não. A distribuição da população das respostas para a necessidade de ajuda com suas habilidades matemáticas está ilustrada na Figura 9-5, em forma de um gráfico de barras (veja o Capítulo 4, para mais informações sobre gráficos de barras).

Suponha que você fosse coletar amostras com 100 estudantes selecionados dessa população combinada com mais de 1 milhão de alunos (todos os alunos que fizeram o ACT em 2002) e encontrasse a proporção de quem indicou que precisaria de ajuda em matemática em cada caso. A distribuição de todas as proporções amostrais é mostrada na Figura 9-6. Ela é uma distribuição normal com média $p = 0{,}38$ e erro padrão $= \sqrt{\frac{0.38\,(1-0.38)}{100}} = \sqrt{0.00236} = = 0{,}049$ ou 4,9%, ou, aproximadamente, 5%. Usando o TLC, você pode dizer que algumas das proporções amostrais são maiores do que 0,38; outras são mais baixas, mas a maioria delas (cerca de 95%) encontra-se na área de 0,38 mais ou menos $2 \times 0{,}05 = 0{,}10$ ou 38% $\pm$ 10%. Esses resultados ainda variam um pouco, cerca de 10% acima ou abaixo da proporção populacional.

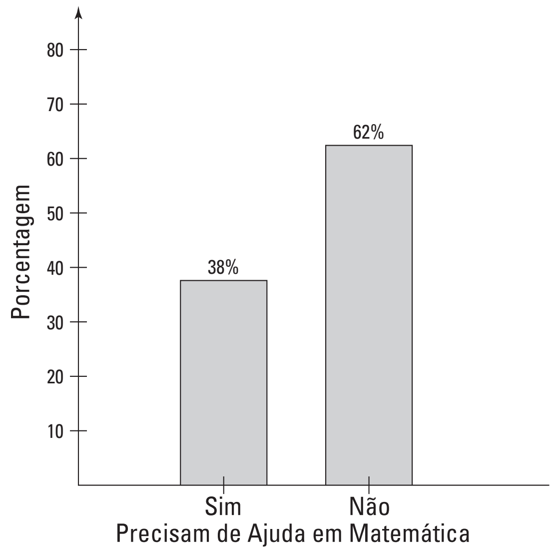


Figura 9-5: Porcentagens da população para todos os alunos que responderam à pergunta sobre a necessidade de ajuda em matemática para o ACT de 2002.

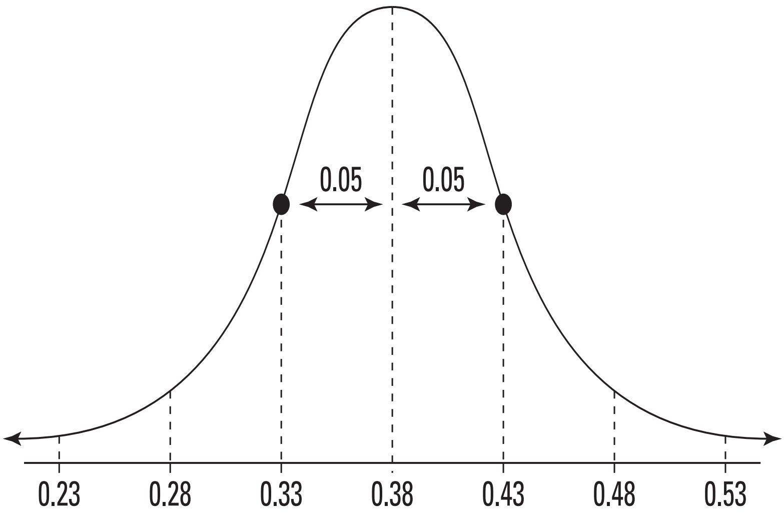


Figura 9-6: Proporção de alunos, para uma amostra de 100 alunos, que responderam "sim" à pergunta sobre a necessidade de ajuda em matemática feita no ACT de 2002.

Agora, pegue amostras com 1.000 selecionados da população original e encontre a proporção dos que responderam que necessitam de ajuda em matemática para cada amostra. A distribuição das proporções amostrais, nesse caso, irão se parecer muito com a Figura 9-7. Tudo se parecerá com a Figura 9-6, exceto pelo fato de que a distribuição ficará mais apertada; o erro padrão se reduzirá para $\sqrt{\frac{0.38(1-0.38)}{1,000}} = \sqrt{0.000236}$ = 0,015, ou 1,5%. Cerca de 95% dos resultados amostrais se encontrarão entre 0,38 – 2 (0,015) e 0,38 + 2 (0,015), ou seja, entre 0,35 e 0,41 (ou 35% e 41%). Em outras palavras, se você pegar várias amostras diferentes dessa população, todas com o mesmo tamanho (1.000) e encontrar a proporção amostral para cada amostra, sua proporção amostral não irá variar muito de amostra para amostra. Ao invés disso, todas elas ficariam bem próximas umas das outras. Isso por causa do tamanho da amostra.

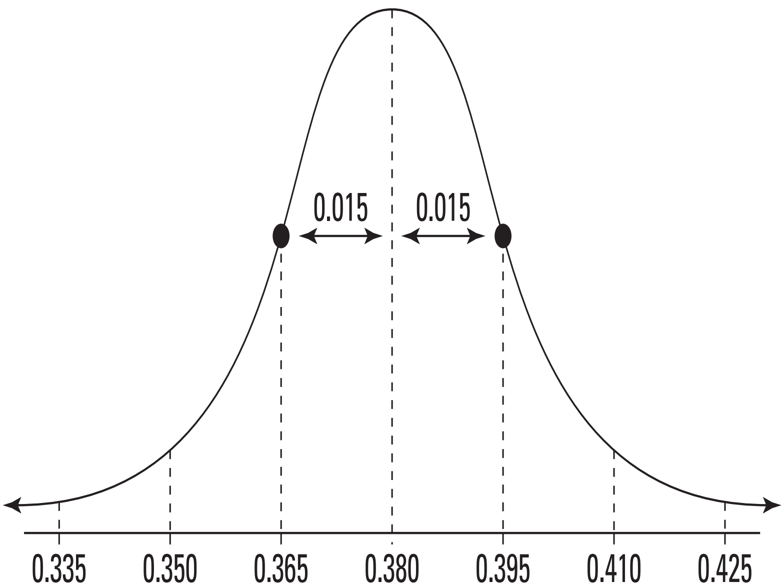


Figura 9-7: Proporção de alunos, em uma amostra de 1.000 alunos, que responderam "sim" à pergunta sobre a necessidade de ajuda em matemática no ACT de 2002.

Antes de tirar qualquer conclusão a partir de porcentagens amostrais, procure ter uma ideia de quanto os resultados podem variam por meio do erro padrão ou da margem de erro (que é, aproximadamente, 2 erros padrões; veja o Capítulo 10). Saber a quantidade de variabilidade esperada irá lhe ajudar a manter os resultados em perspectiva.

Mas qual é o tamanho suficiente para que o teorema do limite central possa lidar com os dados categorizados? A maior parte dos estatísticos concorda que $n \times p$ e $n \times (1 - p)$ deveriam ser maiores ou iguais a 5. Esse valor daria conta de quaisquer situações em que a proporção esteja muito próxima tanto de 1 quanto de 0 (ou seja, em situações extremas em que se considera quase todo mundo ou quase ninguém dentro do grupo de interesse). Em tais situações, você precisaria de uma amostra grande para garantir que todos os grupos estejam representados, até mesmo aqueles que não contenham muita gente. A maioria das enquetes e pesquisas de opinião facilmente coleta pessoas o bastante para levar em conta essas situações.

O TLC é uma ótima notícia para as pessoas que estão tentando interpretar os resultados amostrais. Desde que o tamanho da amostra seja grande o suficiente (e os dados sejam confiáveis e imparciais), a informação relatada estará muito próxima da verdade (mas, lembre-se do que eu disso, desde que os resultados sejam confiáveis e imparciais. Veja o Capítulo 2, para mais exemplos de como a estatística pode errar).

O teorema do limite central também permite responder outra importante questão com relação às médias e proporções amostrais. Por exemplo, se uma empresa promete entregar uma encomenda em até 2 dias e suas amostra de 30 encomendas levaram 2,4 dias, será que você tem evidências o suficiente para culpar a empresa por fazer uma propaganda enganosa? Ou será que essa foi apenas uma amostra atípica? Direciono esse tipo de pergunta ao Capítulo 14.

Se você estiver preocupado em ter que sempre saber a média populacional (μ) ou a proporção populacional (p) para poder usar o TLC, não tema! Você descobrirá os segredos que os estatísticos já conhecem há anos: caso você não saiba determinado valor, simplesmente estime-o e vá em frente. Mais sobre esse assunto no Capítulo 11.

Examinando os Fatores que Influenciam a Variabilidade em Resultados Amostrais

Dois fatores principais influenciam a variabilidade em uma média ou proporção amostral: o tamanho da amostra e a variabilidade na população original.

O Tamanho da Amostra

O tamanho amostral influencia a variabilidade dos resultados amostrais.

Suponha que você tenha um viveiro de peixes e queira saber o comprimento médio de todos os peixes do viveiro. Se você retirar várias amostras aleatórias com 100 peixes e várias amostras aleatórias com 1.000 peixes, sempre registrando a média amostral, quais médias amostrais irão variar mais, as com 100 ou as com 1.000 peixes? As amostras com 100 peixes irão variar mais, pois cada média amostral baseia-se em menos informações (em menos peixes). As proporções amostrais seriam influenciadas da mesma forma.

LEMBRE-SE

Amostras de tamanho pequeno resultam em médias amostrais (e proporções amostrais) com erros padrões grandes. Amostras de tamanho grande resultam em médias amostrais (e em proporções amostrais) com erros padrões menores. Em outras palavras, quanto mais dados forem coletados para uma única amostra, menor será a variação de uma amostra para a outra.

PAPO TÉCNICO

A variabilidade nas médias amostrais (ou nas proporções amostrais) é medida por meio de erros padrões. A variabilidade das médias amostrais é $\frac{\sigma}{\sqrt{n}}$ e a variabilidade das proporções amostrais é $\sqrt{\frac{p\,(1-p)}{n}}$. O denominador de cada uma dessas fórmulas é composto apenas pelo *n* (e nada mais). Portanto, conforme o tamanho amostral aumenta (o denominador), o erro padrão (a fração inteira) diminui. Quanto mais informações forem fornecidas pela amostra (por meio de um grande tamanho amostral) menor será a variabilidade das médias e proporções amostrais.

Variabilidade da População

Conforme a variabilidade na população aumenta, também aumenta a variabilidade na média e na proporção amostrais. Suponha que você tenha dois viveiros de peixe e queira descobrir o comprimento médio de todos os peixes de cada viveiro. O tamanho dos peixes no Viveiro Varia-Muito varia muito mais do que o tamanho dos peixes no Viveiro Varia-Menos. Você coleta uma amostra de 100 peixes de cada viveiro e encontra comprimento médio dos peixes em sua amostra. Se você coletar várias amostras com 100 peixes e registrar a média amostral para cada uma das amostras, quais médias amostrais irão variar mais, as do Viveiro Varia-Muito ou as do Viveiro Varia-Menos? As médias amostrais do Viveiro Varia-Muito irão variar mais, pois, para começar, o tamanho da população de peixes neste varia muito mais.

PAPO TÉCNICO

A variabilidade das proporções amostrais é influenciada da mesma forma pela variabilidade da população original. Por exemplo, suponha que você queira estimar a proporção de peixe no Viveiro Varia-Menos que esteja saudável (vamos chamar esse dado de p). Se quase todos os peixes do Viveiro Varia-Menos estiverem saudáveis (o que significa que o valor de*p* está próximo de 1), o desvio padrão da população, *p* (1 – p), será menor porque a maioria dos peixes apresenta o mesmo estado de saúde. Se, depois, você coletar muitas amostras dessa população homogênea (com relação ao estado de saúde) e encontrar a porcentagem que está saudável, você não deve esperar que essa porcentagem mude muito de uma amostra para a outra. Assim, o erro padrão de uma proporção amostral é menor quando o valor de*p* for próximo a 1. O mesmo ocorre quando a maioria dos peixes não é saudável (o valor de*p* está próximo a 0). Entretanto, se 50% dos peixes estiverem saudáveis e 50% não estiverem, você verá mais variabilidade em suas proporções amostrais de uma amostra para outra, pois a população apresenta uma maior variabilidade em seu estado de saúde. Na realidade, uma população em que o valor de$p$ é igual a 0,5 apresenta uma maior variabilidade, que acaba resultando em valores elevados para os erros padrões das proporções amostrais.

Quanto maior for a variabilidade da população original, maior será a variabilidade nos erros padrões das médias amostrais (ou das proporções amostrais). Observe, no entanto, que essa elevada variabilidade pode ser compensada pelo aumento do tamanho amostral, como discutido acima.

Lembre-se de que a variabilidade das médias amostrais é $\frac{\sigma}{\sqrt{n}}$, e a variabilidade nas proporções amostrais é $\sqrt{\frac{p(1-p)}{n}}$. O numerador de cada uma dessas fórmulas é, na verdade, o desvio padrão da população original em cada caso (σ para dados numéricos e $p\ [1 - p]$ para dados categorizados). Portanto, conforme o desvio padrão da população (o numerador) aumenta, o erro padrão (a fração inteira) também aumenta. Quanto mais variabilidade houver nas médias da população, maior será a variabilidade nas médias amostrais (ou nas proporções amostrais). Essa elevada variabilidade pode, no entanto, ser compensada pelo aumento do tamanho amostral, pois, conforme o valor de n (o denominador) aumenta, toda a fração que compreende o erro padrão diminui.

Qualquer um pode colocar alguns números em uma fórmula e demonstrar um resultado que eles acreditem (ou queiram que você acredite) ser verdadeiro. Mas, se esses resultados forem tendenciosos, sua precisão não será relevante (no entanto, as fórmulas não sabem disso e, portanto, você precisa estar atento). Certifique-se de ter visto como a amostra foi selecionada para determinado estudo e como os dados foram coletados, antes de examinar quaisquer medidas de quanto esses resultados podem variar (o Capítulo 17 cobre em detalhes essa questão).

Capítulo 10

Deixando Espaço para a Margem de Erro

Neste Capítulo

- Entendendo a margem de erro
- Calculando a margem de erro
- Investigando os efeitos do tamanho amostral
- Descobrindo o que a margem de erro não mede

As boas pesquisas de opinião e os bons estudos científicos sempre incluem uma medida que indica a precisão de seus resultados, para que assim os consumidores de informação que eles geram possam colocar esses resultados em perspectiva. Essa medida é chamada de margem de erro – e é uma medida que indica o quanto os resultados da amostra deveriam se aproximar do parâmetro da população que está sendo estudada.

Felizmente, muitos jornalistas também estão se dando conta da importância da margem de erro para o acesso à informação e, portanto, estão começando a aparecer na mídia cada vez mais reportagens que incluem a margem de erro. Mas o que a margem de erro realmente significa? E será que ela realmente conta a história toda?

Este capítulo irá abordar a margem de erro e o que ela pode e não pode fazer para ajudar em seu acesso à precisão da informação estatística. Esse capítulo também irá examinar a questão relacionada ao tamanho amostral; você se surpreenderá em saber que uma amostra pequena pode ser utilizada para se obter um bom perfil de um país – ou do mundo –, desde que a pesquisa seja executada corretamente.

Explorando a Importância Daquele Mais ou Menos

A margem de erro é um termo no qual você já deve ter ouvido falar antes, especialmente em contextos relacionados aos resultados de pesquisas de opinião. Por exemplo, você já deve ter ouvido alguém falar: “Esta pesquisa tem uma margem de erro de três pontos percentuais para mais ou para menos”. Talvez você também já tenha se perguntado o que deveria fazer com essa informação, e qual é sua real importância. A verdade é que os resultados das pesquisas por si só (sem nenhuma margem de erro) são apenas uma medida de como a amostra dos indivíduos selecionados sentem-se com relação à determinada questão; eles não refletem o como uma população inteira poderia ter se sentido se todos tivessem sido questionados (que trabalho isso daria, não?). A margem de erro ajuda-lhe a medir sua proximidade da verdade com relação a toda população que está sendo estudada, ainda que utilize apenas a informação reunida a partir de uma amostra dessa população.

Como discutido no Capítulo 3, uma amostra é um grupo representativo retirado da população objeto de seu estudo. Os resultados fundamentados em uma amostra não serão exatamente os mesmos que você encontraria se estudasse toda uma população, pois, quando você retira uma amostra, você não obtém informações a respeito de todos em uma dada população. Mas, se o estudo for realizado da maneira correta (veja o Capítulo 17, para saber mais sobre como planejar bons estudos), os resultados obtidos a partir de amostras deverão ficar muito próximos dos reais valores para toda uma população.

PAPO TÉCNICO

A margem de erro não significa que alguém cometeu um erro; tudo o que ela significa é que você não conseguiu obter uma amostra de todos os indivíduos de uma população e, assim, você espera que seus resultados amostrais “não englobem” tudo. Em outras palavras, você reconhece que seus resultados podem se alterar com amostras subsequentes e sua precisão apenas é válida dentro de certa variação, que é a margem de erro.

Considere um exemplo do tipo de pesquisa de opinião conduzida por uma das principais organizações em pesquisas, como o Instituto Gallup. Suponha que a última pesquisa realizada pelo Instituto Gallup utilizou uma amostra com 1.000 cidadãos americanos e os resultados mostraram que 520 pessoas (52%) acham que o presidente está fazendo um bom trabalho, comparadas aos 48% que não têm a mesma opinião. Suponha que o Instituo Gallup informe que essa pesquisa possui uma margem de erro de 3% para mais ou para menos. Agora, você sabe que a maioria das pessoas dessa amostra aprova o presidente, mas você sabe dizer se a maioria de todos os cidadãos americanos aprova o presidente? Neste caso, você não pode. E por que não?

Se 52% dos indivíduos da amostra aprovam o presidente, espera-se que a porcentagem de toda a população americana que aprova o trabalho do presidente seja igual a 52%, mais ou menos 3%. Assim, você pode dizer que entre 49% a 55% de todos os americanos aprovam o presidente. Isso é o máximo que você consegue com uma amostra composta por 1.000 pessoas. Mas note que 49%, a extremidade inferior dessa variação, representa a minoria, pois é menos do que 50%. Sendo assim, você realmente não pode afirmar que a maioria da população americana apoia o presidente, apenas se baseando nessa amostra. O que você pode dizer é que entre 49% a 55% de toda a população americana apoia o presidente, o que pode ou não representar a maioria.

Pense no tamanho amostral por um momento. Não é interessante o fato de que uma amostra composta por 1.000 americanos retirados de uma população total de mais de 280.000.000 possa levar a uma margem de erro de apenas 3% para mais ou para menos? Isso é inacreditável e mostra que para se obter uma boa ideia do que está acontecendo com uma população muito grande, basta uma amostra bastante minúscula de seu total. A estatística é, de fato, uma ferramenta muito poderosa para se descobrir a opinião das pessoas com relação a vários assuntos. Talvez seja por isso que tantas pessoas realizem pesquisas de opinião e, por isso, também, você acaba ficando tão entediado por ter que responder a tantas pesquisas.

Para se ter uma vaga ideia do valor da margem de erro para um dado tamanho amostral, basta dividir o número 1 pela raiz quadrada do tamanho amostral (n). Para o exemplo do Instituto Gallup, $n = 1.000$ e sua raiz quadrada é, aproximadamente, 31,62, portanto a margem de erro é igual a 1 dividido por 31,62, ou seja, cerca de 0,03, que é o equivalente a 3%. Ainda neste capítulo, você verá como chegar a um cálculo mais preciso para a obtenção do valor da margem de erro

Descobrindo a Margem de Erro: Uma Fórmula Geral

A margem de erro é o valor “para mais ou para menos” que fica anexado aos seus resultados amostrais a partir do momento em que você deixa de pensar na amostra propriamente dita e passa a pensar na população total que ela representa. Sendo assim, você sabe que a fórmula geral para o cálculo da margem de erro contém, em sua frente, um sinal de ±. Portanto, como você irá obter aquele valor mais ou menos (além da maneira demonstrada acima)? Esta seção irá lhe mostrar o como fazer isso.

Medindo a variabilidade da amostra

Os resultados amostrais variam, mas quanto? Segundo o teorema do limite central (veja o Capítulo 9), quando os tamanhos amostrais são grandes o bastante, a distribuição das proporções amostrais (ou as médias amostrais) apresenta-se como uma curva em forma de sino (ou distribuição normal – veja o Capítulo 8). Algumas das proporções amostrais (ou médias amostrais) irão superestimar o valor da população e outras irão subestimá-lo, mas a maioria ficará próxima do centro. E o que é o centro? Se você fizer a média de todos os resultados a partir de todas as amostras possíveis, essa média seria a proporção real da população, no caso de dados categorizados, ou a média real da população, no caso de dados numéricos. Geralmente, não é possível ter tempo nem dinheiro suficiente para observar todos os resultados das possíveis amostras e fazer sua média, mas saber algo a respeito de todas as outras possibilidades de amostras realmente lhe ajudará a obter o valor esperado para a variação de sua única proporção (ou média) amostral.

Os erros padrões são o alicerce da margem de erro. O erro padrão de uma estatística é, basicamente, igual ao desvio padrão dos dados dividido pela raiz quadrada de n (o tamanho amostral). Isso reflete o fato de que o tamanho amostral exerce uma grande influência sobre o quanto a estatística amostral irá variar de uma amostra para outra (leia o Capítulo 9 para mais informações a respeito de erros padrões).

O número de erros padrões que você deseja somar ou subtrair para chegar à margem de erro depende da margem de confiança que você pretende dar a seus resultados (isso se chama seu intervalo de confiança). Geralmente, se deseja mostrar 95% de precisão, sendo assim, a regra básica é somar ou subtrair aproximadamente 2 erros padrões (1,96, para ser exata) a fim de obter a margem de erro. Isso lhe permitirá levar em consideração cerca de 95% de todos os possíveis resultados que venham a ocorrer com as novas coletas de amostras. Para mostrar 99% de precisão, some e subtraia aproximadamente 3 erros padrões (2,58 para ser exata). Veja o Capítulo 12, para mais discussões acerca dos intervalos de confiança e o número de erros padrões.

PAPO TÉCNICO

Para ser exato com relação ao número de erros padrões que devem ser somados e subtraídos para o cálculo da margem de erro para qualquer nível de credibilidade, é necessário utilizar uma curva em forma de sino especial, conhecida como distribuição normal padrão (veja o Capítulo 8, para mais detalhes). Para qualquer nível de credibilidade, um valor correspondente na distribuição normal padrão (chamado de valor Z) representa o número de erros padrões a serem somados ou subtraídos, a fim de contabilizar o nível de credibilidade. Para 95% de credibilidade, o valor Z exato é de 1,96 (aproximadamente 2) e para 99% de credibilidade, o valor Z exato é 2,58 (aproximadamente 3). Alguns dos níveis de credibilidade mais utilizados, juntamente com seus valores Z correspondentes, estão na Tabela 10-1.

Tabela 10-1 Valores Z para Níveis de Credibilidade Selecionados (expressos em Porcentagem)

Porcentagem de Credibilidade	*Valor Z*
80	1,28
90	1,64
95	1,96
98	2,33
99	2,58

Calculando a margem de erro para uma proporção amostral

Quando uma pesquisa pede às pessoas que elas escolham uma alternativa, (como por exemplo, “Você aprova, não aprova ou prefere não opinar a respeito do trabalho feito pelo presidente?”), a estatística utilizada para relatar os resultados é a proporção de pessoas da amostra que se encaixaram em cada grupo, também conhecida como proporção amostral ou porcentagem amostral. A fórmula geral para o calculo da margem de erro de uma proporção amostral é $Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, em que $\hat{p}$ é a proporção amostral, n é o tamanho amostral e Z é o valor Z apropriado para o nível de credibilidade desejado (veja a Tabela 10-1).

Veja a seguir os passos para o cálculo da margem de erro para uma porcentagem amostral:

1. **Encontre a proporção amostral, $\hat{p}$, e o tamanho amostral, *n***
2. **Multiplique $\hat{p} \times (1 - \hat{p})$.**
3. **Divida o resultado obtido no passo anterior por n.**
4. **Obtenha a raiz quadrada do valor calculado**

   E você já tem o erro padrão

5. **Multiplique o resultado pelo valor Z apropriado ao nível de credibilidade desejado.**

   Veja a Tabela 10-1. O valor Z é 1,96 para um nível de credibilidade de 95% em seus resultados.

Se observarmos o exemplo que envolve a aprovação do presidente pelos cidadãos americanos, podemos encontrar a margem de erro. Primeiro, considere que o nível de credibilidade desejado seja de 95%, portanto Z = 1,96. O número de americanos na amostra que disseram aprovar o presidente foi de 520. Isso significa que a proporção amostral, é 520 ÷ 1.000 = 0,52 (o tamanho amostral,*n,* era 1.000). A margem de erro para essa pesquisa deve ser calculada da seguinte maneira:

$$Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96 \times \sqrt{\frac{(0.520)(1-0.520)}{1,000}} = 1.96 \times \sqrt{\frac{0.2496}{1,000}}$$

$$= 1.96 \times \sqrt{0.0002} = 1.96 \times 0.0158 = 0.0310 = 3.10\%.$$

A proporção amostral é a versão decimal da porcentagem amostral. Em outras palavras, se você tem uma porcentagem amostral igual a 5%, na fórmula, você deve utilizar a representação 0,05 e não o número 5. Para transformar uma porcentagem em decimal, basta dividi-la por 100. Depois que todos os cálculos estiverem prontos, você pode transformar o resultado final, que estará em formal decimal, em porcentagem novamente, para isso, basta multiplicá-lo por 100.

Relatando os resultados

Para relatar os resultados obtidos a partir dessa pesquisa, você poderia dizer: "de acordo com minha amostra, 52%, mais ou menos uma margem de erro de 3,1%, dos americanos aprovam o presidente." (você já está tão bom quanto o Instituto Gallup!).

Mas como uma empresa especializada em pesquisas de opinião relata seus resultados? Veja a seguir como basicamente isso é feito:

> *De acordo com a amostra total de adultos (nesta) pesquisa, temos 95% de certeza de que a margem de erro para nosso procedimento de amostragem e seus resultados não é mais do que ± 3,1 pontos percentuais.*

Isso se parece com aquelas longas listas de exceções que aparecem no final de um comercial de financiamento de veículos. Mas agora você já consegue entender o que aquelas letras minúscula querem dizer!

LEMBRE-SE

Nunca aceite os resultados de uma pesquisa ou de um estudo científico sem a margem de erro. A margem de erro é a única maneira que você possui para medir o quanto a amostra realmente reflete a população estudada. Os resultados amostrais variam e, caso uma amostra diferente tivesse sido selecionada, poderíamos ter obtido um resultado amostral diferente. Você precisa da margem de erro para ser capaz de dizer o quanto os resultados amostrais realmente refletem os valores relacionados à população original. Da próxima vez que você ouvir falar dos resultados de uma pesquisa de opinião ou enquete, observe com atenção para ver se a margem de erro foi mencionada. Alguns meios de comunicação estão melhorando nesse aspecto e começaram a mencionar as margens de erros de suas pesquisas, mas, e os estudos científicos?

Calculando a margem de erro para uma média amostral

Quando, em uma pesquisa, nos pedem um valor numérico (por exemplo, "Quantas pessoas moram em sua casa?"), a estatística utilizada para relatar os resultados é a média de todas as respostas dadas pelas pessoas da amostra, também conhecida como média amostral.

A fórmula geral para o cálculo da margem de erro para sua média amostral é $Z \times \frac{s}{\sqrt{n}}$, onde s é o desvio padrão amostral, n é o tamanho amostral e Z é o valor Z apropriado para o nível de credibilidade desejado (consultar a Tabela 10-1).

A seguir, veja os passos para o cálculo da margem de erro para uma média amostral:

1. **Encontre o desvio padrão da amostra, *s*, e o tamanho amostral, *n*.**

   Para mais informações sobre como calcular a média e o desvio padrão, veja o Capítulo 5.

2. **Divida o desvio padrão da amostra pela raiz quadrada do tamanho amostral**

   E, então, você terá o erro padrão.

3. **Multiplique pelo valor Z apropriado (consulte a Tabela 10-1).**

   O valor Z é 1,96, caso o nível de credibilidade desejado seja de 95%.

Por exemplo, suponha que você seja o gerente de uma sorveteria e esteja treinando os novos funcionários para que eles sejam capazes de encher as casquinhas grandes com a quantidade correta de sorvete (10 onças ou 300 ml cada). Você deve estimar o peso médio das casquinhas que eles fazem, incluindo a margem de erro. Peça a cada um de seus funcionários que verifique aleatoriamente o peso das casquinhas grandes feitas por eles e que registrem os pesos em uma caderneta. Para $n = 50$ casquinhas selecionadas como amostra, a média encontrada foi de 10,3 onças (305 ml), com desvio padrão da amostra, *s,* igual a 0,6 onças (18 ml). Qual é a margem de erro? Supondo que você queira um nível de credibilidade de 95%, a margem de erro poderia ser calculada da seguinte forma:

$$Z \times \frac{s}{\sqrt{n}} = 1.96 \times \frac{0.6}{\sqrt{50}} = 0.17$$

Portanto, para relatar esses resultados, você deveria dizer que, segundo uma amostra com 50 casquinhas, você estima que o peso médio de todas as casquinhas grandes feitas pelos funcionários novos seja de 10,3 onças (305 ml), com uma margem de erro de mais ou menos 0,17 onças (5 ml).

Observe que no exemplo das casquinhas de sorvete, as unidades são onças (ou mililitros), não porcentagens! Quando você tiver que lidar e relatar os resultados sobre os dados, sempre se recorde das unidades. Também certifique de que as estatísticas estejam sendo corretamente mencionadas, de acordo com suas unidades de medida e, caso não isso não esteja acontecendo, pergunte sobre as unidades.

Para evitar um erro de arredondamento em seus cálculos, mantenha pelo menos dois dígitos diferentes de zero depois do ponto decimal ao longo de cada passo utilizado para os cálculos. Erros de arredondamento tendem a se acumular, e você pode errar no final se arredondar muito no início.

Tendo certeza de que você está certo

Caso você queira ter mais de 95% de credibilidade acerca de seus resultados, será necessário somar e subtrair mais dois erros padrões. Por exemplo, para ter 99% de confiança, você deve somar e subtrair cerca de 3 erros padrões para obter sua margem de erro. Isso faz com que a margem de erro fique maior, o que, geralmente, não é algo muito bom. A maioria das pessoas não acredita que seja interessante somar ou subtrair mais um erro padrão para conseguir apenas 4% a mais de credibilidade (99% versus 95%) nos resultados obtidos. A única maneira de estar 100% confiante em seus resultados é fazendo uma margem de erro grande o bastante (por meio da soma e da subtração de muitos e muitos erros padrões) para abranger todas as possibilidades. Por exemplo, você talvez acabe dizendo algo como: "Tenho 100% de certeza de que a porcentagem de pessoas da população que gostam de sorvete é de 50%, mais ou menos 50% pontos percentuais". Em casos como esse, você teria 100% de confiança em seus resultados, mas o que eles significariam? Nada.

Não é possível estar completamente certo de que seus resultados amostrais realmente reflitam a população, mesmo que você inclua uma margem de erro (a menos que você faça algo maluco como incluir 100% de todas as possibilidades como no exemplo anterior). Mesmo se você tiver 95% de confiança em seus resultados, isso, na verdade, significa que se você repetir o processo de amostragem várias e várias vezes, em 5% das vezes, a amostra não irá representar bem a população, simplesmente pelo fato da casualidade (nada relacionado a problemas com o processo de amostragem ou qualquer outra coisa). Portanto, todos os resultados precisam ser vistos a partir desse princípio. Afinal de contas, a estatística implica em nunca se ter certeza de que o que você está dizendo realmente esteja correto!

Determinando o Impacto do Tamanho Amostral

As duas ideias mais importantes relacionadas ao tamanho amostral são as seguintes:

- Todas essas fórmulas funcionam muito bem desde que o tamanho amostral seja grande o bastante (sendo assim, que tamanho é "grande o bastante"? Isso é o que vamos ver nesta seção).
- O tamanho amostral e a margem de erro possuem uma relação inversa.

Esta seção irá clarear ambos os conceitos.

Que tamanho é "grande o bastante?"

Quase todas as pesquisas de opinião são feitas com centenas e milhares de pessoas e, em geral, esse é um número de pessoas grande o bastante para fazer com que a teoria que dá suporte a todas as fórmulas estatísticas funcione. Entretanto, os estatísticos têm testado várias regras gerais para certificarem-se de que os tamanhos amostrais sejam realmente grandes o bastante.

Para as proporções amostrais, é necessário ter certeza de que n x $\hat{p}$ seja pelo menos igual a 5, e n x $(1 - \hat{p})$ seja pelo menos igual a 5. No exemplo anterior, a respeito de uma enquete sobre o presidente, n = 1.000, $\hat{p}$ = 0,52 e 1 – 0,52 = 0,48. Portanto, n x $\hat{p}$ = 1.000 x 0,52 = 520 e n x $(1 - \hat{p})$ = 1.000 x 0,48 = 480. Ambos estão bem acima de 5 e, por isso, tudo está OK.

Para médias amostrais, só é necessário observar o tamanho amostral propriamente dito. Em geral, o tamanho amostral, *n,* deveria ficar aproximadamente acima de 30 para que a teoria estatística funcione. Agora, se o tamanho de sua amostra é igual a 29, não entre em pânico; 30 não é um número mágico, mas apenas uma regra geral.

Tamanho amostral e margem de erro

A relação entre a margem de erro e o tamanho amostral é simples: quanto maior for o tamanho amostral, menor será a margem de erro. Essa é uma relação inversa, pois os dois se movem em direções opostas. Se você pensar a respeito, verá que isso faz sentido, pois quanto mais informações você tiver, maior será a precisão de seus resultados (isso, se assumirmos que os dados foram coletados e manejados da maneira adequada).

Se você se interessa por matemática, no Capítulo 9, você encontrará mais explicações a respeito da relação inversa.

O muito nem sempre é o melhor!

No exemplo anterior, sobre uma enquete envolvendo a taxa de aprovação do presidente, os resultados de uma amostra composta por apenas 1.000 pessoas de um total de mais de 288.000.000 de cidadãos americanos conseguiu ficar dentro de 3% do que toda a população teria dito, caso todos tivessem sido entrevistados. Mas como isso funciona?

Usando a fórmula para a margem de erro para a proporção amostral, podemos observar como a margem de erro altera-se drasticamente em amostras de diferentes tamanhos.

Suponha que no exemplo da taxa de aprovação do presidente, n seja igual a 500 (lembre-se que $\hat{p}$ = 0,52 para esse exemplo). Sendo assim, a margem de erro para o nível de credibilidade de 95% é

$$Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 1.96 \times \sqrt{\frac{(0.520)(0.480)}{500}} = 1.96 \times 0.0223 = 0.0438,$$

que é o equivalente a 4,38%. Quando *n* é igual a 1.000 no mesmo exemplo, a margem de erro (para 95% de confiança) é $1.96 \times \sqrt{\frac{(0.520)(0.480)}{1,000}} = 1.96 \times 0.0158 = 0.0310,$ que é o equivalente a 3,10%. Se *n* fosse 1.500, a margem de erro (para o mesmo nível de confiança) se tornaria $1.96 \times \sqrt{\frac{(0.520)(0.480)}{1,500}} = 1.96 \times 0.0129 = 0.0253,$ ou 2.53%. E, por fim, quando *n* fosse igual a 2.000, a margem de erro seria

$$1.96 \times \sqrt{\frac{(0.520)(0.480)}{2,000}} = 1.96 \times 0.0112 = 0.0219, \text{ ou } 2.19\%.$$

Se observarmos esses diferentes resultados, poderemos perceber que depois de um certo ponto, em amostras cada vez maiores, você começa a ter um retorno diminuído. Cada pessoa a mais aumenta o custo da pesquisa e o aumento do tamanho da amostra de 1.500 para 2.000 pessoas reduz a margem de erro em apenas 0,34% (um terço de um por cento!). O custo extra e o trabalho para conseguir essa diminuição na margem de erro não valem a pena. O muito nem sempre é o melhor!

Mas, o que realmente vai lhe deixar surpreso não é o fato de que o muito nem sempre é o melhor, na verdade, o muito pode ser o pior! Explicarei isso na seção a seguir.

Limitando a Margem de Erro

A margem de erro é uma medida do quão próximo você espera que seus resultados representem toda a população que está sendo estudada (ou, pelo menos, que ela limite seus erros). Devido ao fato de que você está baseando suas conclusões em uma única amostra, é necessário levar em conta o quanto esses resultados amostrais podem variar, simplesmente por causa do acaso.

Outra visão que se tem da margem de erro é que ela representa a distância máxima esperada entre os resultados amostrais e os reais resultados da população (caso você tivesse sido capaz de obtê-los por meio de um censo). Mas é claro que se você tivesse certeza absoluta sobre uma determinada população, você não estaria fazendo uma pesquisa, não é?

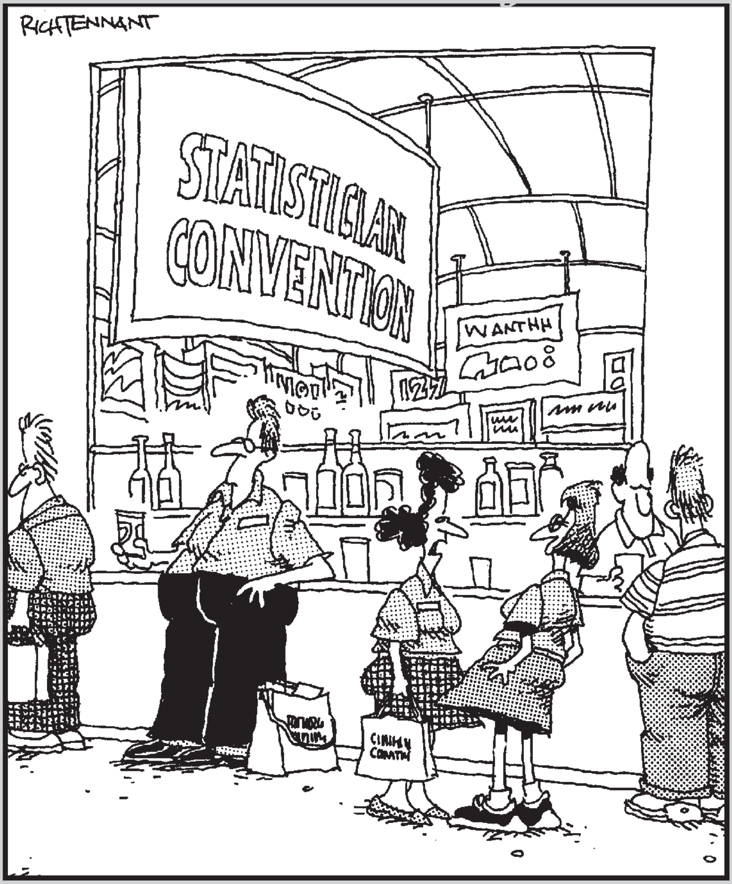
Tão importante quanto saber o que a margem de erro mede é saber o que a margem de erro não mede. A margem de erro não mede nada além da variação provocada pela casualidade. Ou seja, ela não mede nenhuma imparcialidade ou erro ocorrido durante a seleção dos participantes, da preparação e realização da pesquisa durante a coleta e o processo de inserção de dados, nem mesmo durante a análise dos dados e o esboço da conclusão final.

Amostras maiores nem sempre significam amostras melhores! Um bom slogan para ser lembrado, no momento em que você for analisar qualquer resultado estatístico, é, "o lixo que entra é igual ao lixo que sai". Independente da aparência científica da margem de erro, lembre-se de que a fórmula usada para seu cálculo não faz a menor ideia da qualidade dos dados em que a margem de erro se baseia. Se a proporção ou a média amostral for baseada em uma amostra parcial (amostra que favorece determinadas pessoas em detrimento de outras), em maus procedimentos de coleta de dados, má organização, perguntas tendenciosas ou erros sistemáticos durante o registro dos dados, então o cálculo da margem de erro será inútil, pois ela não terá sentido nenhum. Por exemplo, um total de 50.000 pessoas entrevistadas parece algo muito bom, mas se todos forem visitantes de um site na internet, a margem de erro para o resultado dessa pesquisa não significará nada, pois todos os cálculos se basearão em resultados parciais e artificiais! É claro que algumas pessoas vão em frente e relatam seus achados de qualquer maneira; sendo assim, você tem que descobrir o que foi usado na fórmula: boa informação ou lixo? Caso você descubra que foi lixo, já sabe o que fazer com a margem de erro: ignore-a (para mais informações sobre os erros que podem ocorrer durante uma pesquisa de opinião ou um estudo científico, veja os Capítulos 16 e 17, respectivamente).

O Instituto Gallup fala sobre o que é ou não medido pela margem de erro em um termo de responsabilidade utilizado pelo Instituto para o relato dos resultados de suas pesquisas. O Instituto nos fala que, além de erros de amostragem, as pesquisas podem apresentar outros erros ou tendências em virtude da maneira como as perguntas são feitas e de questões logísticas envolvidas na realização das pesquisas (tais como a perda de informação devido a números de telefones não atualizados). Isso significa que, mesmo tendo as melhores intenções e a mais meticulosa atenção aos detalhes e ao controle do processo, várias coisas podem acontecer. Nada é perfeito. Mas o que você precisa saber é que a margem de erro não mede a extensão desses outros tipos de erros.

Parte V
Estimando com Certeza

A 5ª Onda por Rich Tennant



"Tenha cuidado. Ele parece um bom partido, mas existe um histograma sobre seus casamentos e divórcios correndo por aí que não parece ser nada bom".

Nesta Parte...

Vira e mexe, alguém aparece com uma estatística. Essa pessoa, na verdade, não lhe conta a história toda. A estatística por si só perde a parte mais importante da história: o quanto se espera que a estatística varie. Todas as boas estimativas contêm não apenas uma estatística, mas também uma margem de erro. A combinação de uma estatística com uma margem de erro para mais ou para menos é conhecida como intervalo de confiança. Os intervalos de confiança vão além de uma única estatística; ao invés disso, eles fornecem informações importantes para a precisão da estimativa.

Esta parte dará uma visão geral e intuitiva a respeito dos intervalos de confiança: suas funções, fórmulas, cálculos, fatores influentes e interpretação. Você também verá rápidas referências e exemplos dos intervalos de confiança mais utilizados.

Capítulo 11

Estimativa: Interpretando e Avaliando Intervalos de Confiança

Neste Capítulo

- Vendo como funciona o processo de estimativa (ou pelo menos como *deveria* funcionar)
- Obtendo a fórmula geral para um intervalo de confiança
- Interpretando os resultados dos intervalos de confiança
- Detectando os resultados enganosos

A maioria dos estatísticos costuma estimar algumas características acerca de uma população de interesse, tais como a média da renda familiar, a porcentagem de pessoas que compra os presentes de Natal pela internet ou a média de sorvete consumida anualmente nos Estados Unidos (é melhor deixar essa estatística para lá). Tais características da população são chamadas de parâmetros. Normalmente, as pessoas querem estimar (chutar) o valor de um parâmetro através da coleta de uma amostra da população e pelo uso da estatística da amostra. A questão aqui é: Como definir uma "boa estimativa"?

O melhor chute seria não ter chute algum – ou seja, ir às ruas e voltar com o parâmetro propriamente dito. No entanto, não é possível determinar o valor exato de um parâmetro sem conduzir um censo de toda a população – tarefa desencorajadora e cara na maioria dos casos. Os estatísticos, no entanto, não se abalam com o desafio e sempre dizem que "ser um estatístico significa nunca ter que dizer que você está certo; você apenas precisa chegar perto." É claro que os estatísticos querem que seus resultados fiquem o mais próximo possível da verdade, mas dentro de um limite de tempo e orçamento, o "próximo" pode ficar mais fácil do que você possa imaginar. Desde que o processo seja feito corretamente (e, em se tratando da mídia, isso quase nunca acontece!), uma estimativa pode chegar bem próxima de um parâmetro. Este capítulo fornecerá uma visão geral dos intervalos de confiança (o tipo de estimativa usado e recomendado pelos estatísticos), o motivo pelo qual eles devem ser usados (em oposição à apresentação de apenas um único número), como interpretar um intervalo de confiança e como detectar as estimativas enganosas.

Nem Todas as Estimativas São Criadas da Mesma Forma

Leia qualquer revista ou jornal ou assista a qualquer noticiário, e você irá notar um grande número de estatísticas, muitas das quais são estimativas de uma quantidade ou outra. Você deve se perguntar como eles conseguem esses números: em alguns casos, esses números são realmente pesquisados; em outros, eles são apenas tiros no escuro. Aqui vão alguns exemplos de estimativas com as quais me deparei em apenas uma única edição de uma revista líder em assuntos de negócios. Eles foram tirados de uma variedade de fontes:

- 26 milhões de pessoas jogam golfe pelo menos uma vez ao ano.
- 6,7% dos lares nos Estados Unidos foram comprados sem entrada.
- Embora esteja cada vez mais difícil se conseguir alguns empregos, algumas áreas realmente estão recrutando profissionais: durante os próximos oito anos, 13.000 enfermeiros anestesistas serão necessários. Os salários iniciais são de $80.000 a $95.000.
- O número médio de tacos usados por um jogador de beisebol da Liga Principal é de 80 por temporada.
- 7,4 milhões de residentes nos Estados Unidos fizeram viagens de cruzeiro em suas férias em 2002. Dessas pessoas, cerca de 4% visitaram a equipe médica do navio.
- A Lamborghini Murcielago pode ir de 0 a 100 km em 3,7 segundos, com velocidade máxima de aproximadamente 330 km/h.

Algumas estimativas são mais fáceis de serem obtidas do que outras. Aqui vão algumas observações que eu pude fazer sobre essas estimativas.

- Como você sabe que 26 milhões de pessoas jogam golfe pelo menos uma vez ao ano? Na realidade, essa informação pode não ser tão difícil de ser obtida, pois os jogadores de golfe devem sempre assinar em qualquer lugar que eles joguem em um campo de golfe. Sendo assim, uma pesquisa nas planilhas dos campos de golfe poderia nos dar uma boa estimativa de quantas pessoas jogam pelo menos uma vez ao ano (a tarefa mais difícil seria a de não contar duas vezes as pessoas que já haviam sido contadas nas planilhas anteriores).
- Uma pesquisa também seria capaz de estimar a porcentagem de turistas em um cruzeiro que precisam de atendimento médico ou a porcentagem de casas que foram compradas sem entrada. Se a pesquisa for feita da maneira correta (veja o Capítulo 16), esses dados são, provavelmente, bastante precisos.
- Como você estima a quantidade de enfermeiros anestesistas que serão necessários nos próximos oito anos? Você pode começar

observando quantos estarão se aposentando dentro desse período; mas isso não irá contabilizar o crescimento. Uma previsão para daqui um ou dois anos seria mais fácil, mas seria muito difícil de acertar uma previsão para daqui oito anos.

- O número médio de tacos usados por um jogador de beisebol na Liga Principal pode ser encontrado por meio de pesquisas feitas com os próprios jogadores, com os responsáveis pelos equipamentos ou com as empresas que fornecem os tacos.
- A determinação da velocidade de um carro é o mais difícil, mas poderia ser conduzida por meio de um teste com o uso de cronômetros e com vários carros (ao invés de apenas um) da mesma marca e modelo.

Nem todas as estatísticas são criadas da mesma forma. Para determinarmos se uma estatística é confiável, não lhe dê tanta importância logo de cara. Pense se ela realmente faz sentido e como você faria para formular uma estimativa. Se a estatística realmente for importante para você, descubra que processo foi utilizado para se chegar à estimativa.

Ligando uma Estatística a um Parâmetro

O U.S Census Bureau estima a renda familiar mediana nos Estados Unidos e a divide por estado em seu demonstrativo anual realizado a partir da Current Population Survey. Por que estimar a mediana (o número do meio) e não a média da renda familiar? (Veja o Capítulo 5, para mais discussões sobre médias e medianas). Devido ao fato de que a renda familiar tende a ser distorcida, com muita gente na parte inferior e poucos na parte superior.

Para estimar a renda familiar mediana, o Census Bureau coleta uma amostra aleatória de cerca de 28.000 famílias e faz perguntas (a renda familiar está entre essas perguntas). Baseando-se em uma amostra de 28.000 lares, o Bureau calcula a renda familiar mediana para essa amostra: para o ano de 2000, a renda familiar mediana foi de $42, 228.

O Census Bureau usa a renda familiar mediana da amostra (uma estatística) para estimar a renda familiar mediana para todo Estados Unidos (o parâmetro). Devido ao fato de que o Bureau sabe que sua amostra pode não refletir a população inteira de maneira totalmente precisa, ele inclui uma margem de erro (veja o Capítulo 10) junto com os resultados. Esse mais ou menos (±) que vem junto com as estimativas nos ajuda a colocar os resultados em perspectiva. Quando se sabe a margem de erro, é possível se ter uma ideia de quanto erro pode estar embutido na estimativa, simplesmente pelo fato de que essa estimativa foi baseada em uma amostra da população e não na população inteira.

Devido ao fato de que o Bureau não estuda toda a população e sabe que sua amostra pode não representar perfeitamente a população inteira, ele calcula a margem de erro para a mediana do grupo da amostra e a inclui como parte da estimativa. Para 2000, a margem de erro da renda familiar mediana da amostra era de \$258. Assim, o U.S. Census Bureau estima que em 2000, a renda familiar mediana era de \$42, 228 mais ou menos \$258 ou \$42, 228 ± \$258. Isso representa o intervalo de confiança para a renda familiar mediana nos Estados Unidos (veja a seguir, a seção sobre intervalos de confiança).

Observe que a margem de erro para o exemplo acima é bastante pequena; isso ocorre por causa do tamanho da amostra utilizada. Veja o Capítulo 10, para mais informações acerca da relação entre o tamanho amostral e o tamanho da margem de erro.

Fazendo a Sua Melhor Estimativa

A melhor maneira de estimar o parâmetro (uma característica de uma população inteira) é por meio de uma estatística com margem de erro para mais ou para menos que se baseie em uma amostra grande. Dessa forma, seus resultados apresentam uma estimativa baseada em sua amostra, juntamente com alguns indicadores do quanto essa estimativa poderia variar de uma amostra para outra.

Uma estatística com margem de erro para mais ou para menos é chamada de intervalo de confiança:

- A palavra intervalo é usada porque seu resultado se torna um intervalo. Por exemplo, se você disser que a porcentagem de crianças que gostam de beisebol é de 40% mais ou menos 3,5%, isso significa que a porcentagem de crianças que gostam de beisebol está em algum lugar entre 40% - 3,5% = 36,5% e 40% + 3,5% = 43,5%. Assim, a ponta inferior de seu intervalo é a sua estatística menos a margem de erro e a ponta superior é sua estatística mais a margem de erro.

- A palavra confiança é usada porque você possui certa confiança no processo pelo qual você chegou ao intervalo. Isso se chama nível de confiança (ou credibilidade).

Você irá encontrar fórmulas e exemplos dos intervalos de confiança mais utilizados no Capítulo 13.

Interpretando os Resultados com Confiança

Suponha que você seja um biólogo e esteja tentando pegar um peixe usando uma rede de mão, e o tamanho de sua rede representa a largura de seu intervalo de confiança (a largura é duas vezes a margem de erro, devido à soma e à subtração).

Suponha que seu nível de confiança é 95%. O que isso realmente significa? Isso significa que se você mergulhar essa rede várias e várias vezes na água, você irá pegar um peixe em 95% das vezes. Pegar um peixe, nessa situação, significa que seu intervalo de confiança estava correto e continha o verdadeiro parâmetro (nesse caso, o parâmetro é representado pelo próprio peixe).

Mas, isso significa que em qualquer tentativa você tem 95% de chances de pegar um peixe? Não. Está ficando confuso? Certamente. Aqui vai a explicação: em uma única tentativa, digamos que você feche os olhos antes de mergulhar a rede na água. Nesse ponto, suas chances de pegar um peixe são de 95%. Mas, então, vá em frente e mergulhe a rede na água com os olhos ainda fechados. Depois que isso é feito, você acaba tendo apenas duas possibilidades, você pode ou não pegar um peixe; a probabilidade não está envolvida nesse caso.

Da mesma forma, depois que os dados são coletados e o intervalo de confiança é calculado, você pode ou não ter capturado o verdadeiro parâmetro populacional. Assim, você não está dizendo que tem 95% de certeza de que o parâmetro está no intervalo, pois você pode ou não tê-lo capturado. O que você sabe é que, com o passar do tempo, você pode ter 95% de confiança no processo utilizado para a coleta de dados e formação do intervalo de confiança. Você sabe que esse processo irá resultar em intervalos que capturam a média em 95% das vezes. Nas outras 5% das vezes, os dados coletados na amostra, simplesmente devido ao acaso, possuem valores ou muito altos ou muito baixos e, por isso, não representam a população. Nesses casos, você não capturou o parâmetro.

Sendo assim, você sabe que com o tamanho e a composição de sua rede, você irá pegar um peixe em 95% das vezes. Em uma tentativa apenas, no entanto, você pode ou não pegar um peixe.

O Nível de confiança, o tamanho amostral e a variabilidade da população influenciam o tamanho da margem de erro e o comprimento do intervalo de confiança. Mas a margem de erro e o comprimento do intervalo de confiança não significam nada se os dados usados no estudo forem tendenciosos e/ou não confiáveis. O melhor conselho é o de observar a maneira como os dados foram coletados antes de aceitar uma margem de erro como sendo a verdade (veja o Capítulo 10).

Identificando os Intervalos de Confiança Enganosos

Quando os dados são provenientes de experimentos e pesquisas de opinião bem planejadas (veja os Capítulo 16 e 17), que se baseiam em grandes amostras aleatórias (veja o Capítulo 9), você pode se sentir seguro com relação à qualidade da informação. Quando uma margem de erro for pequena, relativamente falando, poderia se pensar que os intervalos de confiança fornecem estimativas precisas e confiáveis para seus parâmetros. Mas nem sempre é assim que acontece.

Nem todas as estimativas são tão precisas e confiáveis como alguns querem que você acredite. Por exemplo, uma enquete realizada por um site na internet, que se baseie em 20.000 acessos, poderá apresentar uma margem de erro pequena, de acordo com a fórmula, mas essa margem de erro não significará nada se a enquete apenas for acessível às pessoas que costumam visitar o site. Em outras palavras, a amostra nem chega perto de ser uma amostra aleatória (em que todos os membros da população possuem a mesma chance de serem selecionados para a pesquisa). Apesar disso, esses resultados são informados juntamente com suas margens de erro, para parecer que o estudo foi verdadeiramente científico. Tome cuidado com esses resultados artificiais! (Leia o Capítulo 10, para obter mais informações sobre os limites da margem de erro).

Antes de tomar qualquer decisão baseada na estimativa de alguém, faça o seguinte:

- Investigue o como a estatística foi criada; ela deveria ter sido o resultado de um processo científico que resulta em dados confiáveis, parciais e precisos (veja os Capítulos 2 e 3).
- Procure a margem de erro. Caso ela não tenha sido informada, procure-a na fonte original.
- Lembre-se de que se a estatística não for confiável e for tendenciosa, a margem de erro não servirá para nada (veja o Capítulo 16 para evitar a tendenciosidade nos dados em pesquisas e veja o Capítulo 17 para os critérios utilizados para obtenção de bons dados em experimentos).

Capítulo 12

Calculando Intervalos de Confiança Precisos

Neste Capítulo

- Esperando um determinado nível de confiança para seus resultados
- Encontrando um processo geral para o cálculo do intervalo de confiança
- Examinando os fatores que influencia o tamanho do intervalo de confiança

Um intervalo de confiança é uma palavra elegante para uma estatística que vem acompanhada de uma margem de erro (veja o Capítulo 11, para uma visão geral dos intervalos de confiança; veja o Capítulo 10, para informações sobre a margem de erro). Devido ao fato de a maioria das estatísticas ser calculada visando estimar as características de uma dada população (chamadas de parâmetros) retiradas de uma amostra, toda a estatística deveria incluir uma margem de erro como medida para seu nível de precisão. Afinal de contas (assim como o que diz o Capítulo 9), quando se está lidando com amostras, os resultados variam.

Neste Capítulo, você irá descobrir como calcular seu próprio intervalo de confiança. Você também irá estudar os pontos mais importantes dos intervalos de confiança: o que os faz estreitos ou largos, o que lhe faz ter mais ou menos confiança em seus resultados e o que eles medem e não medem. Com essas informações, você saberá o que procurar quando se deparar com resultados estatísticos e entenderá como avaliar a verdadeira precisão desses resultados.

Calculando um Intervalo de Confiança

Um intervalo de confiança é composto por uma estatística, mais ou menos uma margem de erro (veja o Capítulo 10). Por exemplo, suponha que você queira saber a porcentagem de veículos nos Estados Unidos que são do tipo caminhonetes (esse é o parâmetro, para esse caso). É impossível observar todos os veículos nos Estados Unidos, portanto você coleta uma amostra aleatória com 1.000 veículos em um trecho de uma rodovia em diferentes momentos do dia. Você descobre que 7% dos veículos de sua amostra são do tipo caminhonete. Agora, você não pode dizer que exatamente 7% de todos os veículos que transitam nas rodovias dos Estados Unidos sejam caminhonetes, pois você sabe que esse resultado baseia-se apenas em 1.000 veículos coletados para sua amostra. Ao mesmo tempo em que você espera que esses 7% estejam próximos da verdadeira porcentagem, você não pode ter certeza absoluta, pois está se baseando em uma amostra de veículos e não em todos os veículos dos Estados Unidos.

Então, o que fazer? Você deve somar e subtrair uma margem de erro para indicar o quanto de erro você espera que seus resultados tenham (veja o Capítulo 10 para saber mais sobre margem de erro). O erro não é causado por nada de errado que você possa ter feito, ele simplesmente ocorre pelo fato da utilização de uma amostra (o estudo de uma porção da população) e não de um censo (o estudo da população inteira).

O tamanho de seu intervalo de confiança é duas vezes a margem de erro. Por exemplo, suponha que a margem de erro seja de 5%. Um intervalo de confiança de 7% mais ou menos 5%, varia de 7% - 5% = 2% até 7% + 5% = 12%. Isso significa que seu comprimento é 12% - 2% = 10%. Uma maneira simples de calcular isso é dizer que o comprimento do intervalo de confiança é duas vezes a margem de erro. Nesse caso, o tamanho do intervalo de confiança é 2 × 5% = 10%.

LEMBRE-SE

O comprimento de um intervalo de confiança é a distância da ponta inferior do intervalo (estatística menos a margem de erro) até a ponta superior do intervalo (estatística mais a margem de erro). Você pode sempre calcular rapidamente o comprimento do intervalo de confiança multiplicando a margem de erro por 2.

A seguir, veja os passos para estimar um parâmetro com o intervalo de confiança, juntamente com referências de onde você pode encontrar mais informações detalhadas sobre como realizar cada passo.

1. **Escolha seu nível de confiança e seu tamanho amostral (veja o capítulo 9)**

2. **Selecione uma amostra aleatória de indivíduos da população (veja o Capítulo 3)**

3. **Colete dados confiáveis e relevantes dos indivíduos da amostra.**

   Veja o Capítulo 16 para informações sobre dados de pesquisas de opinião e o Capítulo 17 sobre dados de experimentos.

4. **Resuma os dados em estatística, geralmente uma média ou proporção (veja o Capítulo 5)**

5. **Calcule a margem de erro (veja o Capítulo 10)**

6. **Some e subtraia a margem de erro da estatística para chegar a sua estimativa final do parâmetro.**

Esse resultado será o intervalo de confiança para o parâmetro.

Adolescentes contra o tabaco sem fumaça

Um estudo em andamento conduzido pela University of Michigan tem monitorado o comportamento de adolescentes em relação a várias questões, incluindo o risco do tabaco sem fumaça (também conhecido como tabaco para mascar). O estudo mostra que os adolescentes de hoje percebem que o tabaco sem fumaça causa muitos riscos comparados com os adolescentes de 15 anos atrás. Os resultados são os seguintes:

- Em uma amostra colida em 2001, e composta por 2.100 estudantes do ensino médio, 45,4% percebiam que o tabaco sem fumaça causa muitos prejuízos. A margem de risco era de mais ou menos 2%. Sendo assim, um intervalo de confiança de 95% para a porcentagem de todos os estudantes do ensino médio que percebiam o perigo causado pelo tabaco sem fumaça é de 45,4% ± 2%.
- Em uma amostra composta por 3.000 estudantes do ensino médio em 1986, o intervalo de confiança para todos os estudantes que acreditavam que o tabaco sem fumaça era nocivo à saúde foi de 25,8% ± 1,6%.

Escolhendo um Nível de Confiança

Observe que o exemplo pertinente ao comportamento dos adolescentes em relação ao tabaco (veja o quadro “Adolescentes contra o tabaco sem fumaça”) inclui a frase “intervalo de confiança de 95%”. Todos os intervalos de confiança (e toda a margem de erro) possuem um nível de confiança associado a ele. Nesse exemplo, o nível de confiança foi de 95%. O nível de confiança ajuda-lhe a contabilizar os outros possíveis resultados amostrais que você poderia ter conseguido, quando estivesse fazendo a estimativa para um parâmetro por meio do uso dos dados coletados a partir de uma única amostra. Caso você queira contabilizar 95% dos outros resultados possíveis, seu nível de confiança seria de 95%.

A variabilidade nos resultados amostrais é contabilizada em termos de números de erros padrões. O erro padrão é semelhante ao desvio padrão de um conjunto de dados, a única diferença é que o erro padrão é aplicado apenas às médias e porcentagens amostrais que poderiam ser obtidas caso diferentes amostras fossem coletadas (veja o Capítulo 10, para informações acerca dos erros padrões). Todo nível de confiança possui um número de erros padrões correspondentes, que devem ser somados e subtraídos. Esse número de erros padrões é chamado de valor Z (pois corresponde à distribuição normal padrão). Consulte a Tabela 10-1 no Capítulo 10.

Qual é o nível de confiança normalmente utilizado pelos pesquisadores? Eu já vi níveis de confiança que variam de 80% a 99%, mas o mais comum é 95%. De fato, os estatísticos têm um ditado que diz: “Por que os estatísticos gostam de seu trabalho? Porque eles têm que estar certos apenas em 95% das vezes.” (piadinha sem graça, mas interessante, não?)

Estar 95% confiante significa que se você coletar muitas, mas muitas amostras e calcular o intervalo de confiança para todas, 95% dessas amostras terão intervalos de confiança que abrangerão o alvo e realmente irão conter o parâmetro verdadeiro. Para se alcançar um nível de confiança de 95%, a regra empírica diz que você precisa somar e subtrair “cerca” de 2 erros padrões. O teorema do limite central permite-lhe ser mais exato com relação a esse número, portanto, “cerca de 2”, na verdade, é 1,96. Veja a Tabela 10-1, no Capítulo 10, para os níveis de confiança selecionados, juntamente com seus valores Z correspondentes.

Caso você deseje ter mais do que 95% de confiança em seus resultados, será preciso somar e subtrair mais do que dois erros padrões. Por exemplo, para ter 99% de confiança, você deve somar e subtrair cerca de 3 erros padrões, a fim de obter sua margem de erro. Quanto maior for o nível de confiança, maiores serão o valor Z, a margem de erros e o intervalo de confiança (supondo que tudo continue igual). Tudo tem seu preço.

Observe que eu disse: “supondo que tudo continue igual”. É possível compensar um aumento na margem de erro com o aumento do tamanho amostral. Veja a seção “Ampliando o Tamanho Amostral”, para saber mais sobre isso.

Estreitando a Largura

O objetivo final, quando se está fazendo uma estimativa usando um intervalo de confiança, é ter um intervalo de confiança o mais estreito possível. Isso significa que você estará focando o parâmetro. Ter que somar e subtrair valores altos apenas faz com que seus resultados fiquem muito menos precisos. Por exemplo, suponha que você esteja tentando estimar a porcentagem de caminhonetes nas rodovias entre meio dia e seis da tarde e consegue um intervalo de confiança que diz que a porcentagem de caminhonetes é 50%, para mais ou para menos 50%. Uau, esse resultado realmente estreitou seu intervalo! (Não). Você, na verdade, acabou com o propósito de se conseguir uma boa estimativa.

Nesse caso, o intervalo de confiança está muito largo. Seria melhor que você dissesse algo como: o intervalo de confiança de 95% para a porcentagem de caminhonetes que usam as rodovias entre o meio dia e a seis da tarde é de 50%, mais ou menos 3%. Para isso, você necessitaria de uma amostra muito maior, mas valeria mais a pena.

Portanto, se uma margem de erro pequena é boa, uma margem de erro menor seria melhor? Nem sempre. Para se conseguir um intervalo de confiança extremamente estreito, é necessário que se realize um estudo tão mais difícil – e mais caro – que o aumento do custo não justificaria a pequena diferença na precisão. A maioria das pessoas se sente a vontade com uma margem de erro de 2% ou 3% quando a estimativa é uma porcentagem (como a porcentagem de mulheres, republicanas ou fumantes).

Um intervalo de confiança estreito é algo bom.

Mas o que fazer para garantir que seu intervalo de confiança seja o mais estreito possível? Definitivamente, você deve pensar sobre isso antes de coletar os dados; depois que os dados já estiverem coletados, a largura do intervalo de confiança já estará determinada.

Três fatores afetam a largura do intervalo de confiança:

- O nível de confiança (como o discutido na seção anterior)
- O tamanho da amostra
- A variabilidade encontrada na população

A fórmula da margem de erro associada à média amostral é $Z \times \frac{s}{\sqrt{n}}$, em que:

- Z é o valor retirado da distribuição normal padrão correspondente ao nível de confiança (veja a Tabela 10-1, no Capítulo 10).
- n é o tamanho da amostra (veja o Capítulo 9).
- $\frac{s}{\sqrt{n}}$ é o erro padrão da média amostral. (Veja o Capítulo 10, para mais informações sobre os erros padrões).

O intervalo de confiança para a média seria, então, $\bar{x}$ mais ou menos a margem de erro. O Capítulo 13 lhe dará as fórmulas para os intervalos com os quais você provavelmente poderá se deparar.

Cada um desses três fatores (nível de confiança, tamanho amostral e variabilidade da população) influencia a largura do intervalo de confiança. Você já viu os efeitos do nível de confiança. Na seção a seguir, você irá investigar como o tamanho amostral e a variabilidade da população influenciam a largura de um intervalo de confiança.

Observe que a estatística amostral (por exemplo, 7% dos veículos na amostra são do tipo caminhonete) não se relaciona à largura do intervalo de confiança. Ao invés dela, é a margem de erro e os três fatores citados acima que são os únicos responsáveis pela determinação da largura de um intervalo de confiança.

Ampliando o Tamanho Amostral

A relação entre a margem de erro e o tamanho amostral é simples: quanto maior for o tamanho da amostra, menor é a margem de erro. Isso confirma que o que você espera é verdadeiro: quanto mais informação você tiver, mais precisos serão seus resultados (isso quando supomos que a informação seja boa e confiável. Veja o Capítulo 2, para saber como as estatísticas erram).

Se observarmos a fórmula para a margem de erro de uma média amostral, notaremos que ela possui o n como denominador da fração (isso é o que ocorre na maioria das fórmulas para o cálculo da margem de erro): $Z \times \frac{s}{\sqrt{n}}$. Conforme n aumenta, o denominador dessa fração aumenta, fazendo com que a fração total fique menor. Isso é o que faz com que a margem de erro seja menor e resulte em intervalos de confiança mais estreitos.

Quando um elevado nível de confiança for necessário, você deverá aumentar o valor Z e, por isso, a margem de erro irá resultar em intervalo de confiança mais amplo, o que não é algo desejável. No entanto, você pode compensar esse amplo intervalo de confiança aumentando o tamanho da amostra, reduzindo a margem de erro e estreitando o intervalo de confiança. O aumento do tamanho amostral ainda lhe permitirá ter o nível de confiança desejado, mas também assegura que a largura do intervalo de confiança seja pequena (seu objetivo final). Você pode determinar essa informação antes mesmo de começar o estudo: se você souber a margem de erro a qual quer chegar, você pode determinar o tamanho de sua amostra da maneira mais apropriada (veja o Capítulo 9).

Quando sua estatística se tornar uma porcentagem (como a porcentagem de pessoas que preferem usar chinelo no verão), uma maneira aproximada de calcular a margem de erro é dividir o número 1 pela raiz quadrada de *n* (o tamanho amostral). Você pode experimentar diferentes valores para*n,* a fim de ver o como a margem de erro será afetada.

Qual é o tamanho amostral aproximado necessário para se obter um intervalo de confiança estreito com relação às enquetes? Usando a fórmula do parágrafo anterior, você consegue fazer algumas comparações rápidas. Uma pesquisa com 100 pessoas irá ter uma margem de erro de cerca de $\frac{1}{\sqrt{100}} = 0.10$ = 0,10 ou 10% para mais ou para menos (o que significa que a largura do intervalo de confiança é de 20%, valor que pode ser considerado muito grande). No entanto, se sua pesquisa tiver 1.000 pessoas, sua margem de erro cai drasticamente para 3% para mais ou para menos; portanto, a largura agora é igual a apenas 6%. Uma pesquisa com 2.500 pessoas resulta em uma margem de erro de 2% para mais ou para menos (portanto, a largura é reduzida para 4%). Essa é uma amostra bastante pequena para ser tão precisa, se pensarmos no tamanho da população (a população americana, por exemplo, de mais de 280 milhões de pessoas).

Tenha em mente, no entanto, que você não deve ir tão longe com o tamanho de sua amostra, pois, a partir de um certo ponto, você começa a ter um retorno muito pequeno. Por exemplo, se aumentarmos a amostra de 2.500 para 5.000, obteremos um estreitamento de aproximadamente 2 × 1,4 = 2,8%. Todas as vezes em que você entrevista uma pessoa a mais, o custo de sua pesquisa aumenta e, portanto, adicionar mais 2.500 pessoas a uma pesquisa apenas para estreitar o intervalo de confiança em um pouco mais de 1% pode não valer a pena.

A real precisão depende da qualidade dos dados, além do tamanho amostral. Uma amostra grande composta por dados tendenciosos (veja o Capítulo 2) pode ter um intervalo de confiança estreito que não tenha valor nenhum. É como competir em um campeonato de arco e flecha e acertar a maioria de suas flechas no alvo, mas, no final, descobrir que você esteve o tempo todo atirando no alvo de outra pessoa; isso é para você ter uma ideia do quanto estará errado. Dentro do campo da estatística, no entanto, você não consegue medir a tendenciosidade, você pode apenas tentar minimizá-la.

Quanto maior for o tamanho amostral, menor será a margem de erro e a largura de seu intervalo de confiança, desde que tudo permaneça igual e os dados sejam de boa qualidade.

Considerando a Variabilidade da População

Um dos fatores que influenciam a variabilidade nos resultados amostrais é o fato de que a própria população contém variabilidade. Se todos os valores da população fossem exatamente os mesmo, imagine como o mundo seria um lugar entediante (na realidade, os estatísticos não existiriam se não fosse a variabilidade). Por exemplo, em uma população de casas em uma cidade grande como Columbus, em Ohio, você nota uma grande variedade não apenas nos tipos de casas, mas, também, nos tamanhos e nos preços. E a variabilidade nos preços das casas em Columbus, Ohio, pode ser maior do que a variabilidade de preços de casas em construção de um determinado conjunto residencial em desenvolvimento em Columbus.

A variabilidade é medida pelo desvio padrão. Geralmente, o desvio padrão de uma população (σ) não é conhecido, portanto você deve estimá-lo com de *s,* o desvio padrão da amostra (veja o Capítulo 4). Observe que o *s* aparece no numerador do erro padrão na fórmula da margem de erro para a média amostral: $Z \times \frac{s}{\sqrt{n}}$. Logo, como o desvio padrão (o numerador) aumenta, o erro padrão (toda a fração) também aumenta, resultando em uma margem de erro maior e em um intervalo de confiança mais amplo.

Quanto maior é a variabilidade na população original, maior será a margem de erro e mais amplo será o intervalo de confiança. Esses valores podem, no entanto, ser compensados pelo aumento do tamanho amostral.

Capítulo 13

Intervalos de Confiança mais Utilizados: Fórmulas e Exemplos

Neste Capítulo

- Dividindo os intervalos de confiança em fórmulas
- Calculando com confiança

Sempre que você quiser determinar a média de uma população, mas não conseguir encontrá-la exatamente, devido a questões de restrição de tempo/dinheiro (geralmente, algo muito comum), a melhor coisa a se fazer é coletar uma amostra da população, achar seu meio e usá-lo para estimar a média de toda a população. Depois (e veja os Capítulos 11 e 12, para mais detalhes), você deve incluir alguma medida para definir o quão preciso você espera que seus resultados amostrais sejam; afinal de contas, você sabe que seus resultados mudariam pelo menos um pouco se outra amostra fosse observada. Portanto, juntamente com sua média amostral, inclua uma margem de erro (medida que expressa o quanto se espera que os resultados variem de uma amostra para a outra) e a sua média amostral mais ou menos a margem de erro se combinarão para formar o intervalo de confiança para a média populacional.

Mas chegar a um intervalo de confiança pode ser um pouco confuso. Sendo assim, neste capítulo, enfatizarei as fórmulas para quatro dos intervalos de confiança mais utilizados (ICs), explicarei os cálculos e darei-lhe alguns exemplos.

Calculando o Intervalo de Confiança para a Média Populacional

Quando a característica que estiver sendo estudada for numérica (tais como o QI, o preço, a altura, a quantidade ou o peso de algo), a maioria das pessoas prefere relatar o valor médio para a população, pois a média é um número de resumo da população que nos diz onde fica o centro da população. A média populacional é estimada através da média amostral mais ou menos uma margem de erro. O resultado obtido chama-se intervalo de confiança para a média populacional.

A fórmula para um IC para uma média populacional é $\bar{x} \pm Z \times \frac{s}{\sqrt{n}}$, em que $\bar{x}$ é a média amostral, s é o desvio padrão amostral, n é o tamanho amostral e Z é o valor correspondente ao nível de confiança desejado, segundo a distribuição normal padrão (veja, no Capítulo 3, as fórmulas para $\bar{x}$ e s; veja a Tabela 10-1, no Capítulo 10, para os valores Z correspondentes a alguns níveis de confiança).

Para calcular um IC para uma média populacional, faça o seguinte:

1. **Determine o nível de confiança e encontre o valor Z correspondente.**

   Veja a Tabela 10-1 no Capítulo 10.

2. **Encontre a média amostral ($\bar{x}$), o desvio padrão (*s*) e o tamanho amostral (*n*).**

   Veja o Capítulo 3.

3. **Multiplique Z por *s* e divida o resultado pela raiz quadrada de *n*.**

   Essa é a margem de erro.

4. **Some e subtraia a margem de erro da $\bar{x}$ para obter o IC.**

   A ponta inferior do IC é a $\bar{x}$ menos a margem de erro e a ponto superior do IC é a $\bar{x}$ mais a margem de erro.

Por exemplo, suponha que você trabalhe no Departamento de Recursos Naturais e queira estimar, com 95% de confiança, o tamanho médio de dourados em um viveiro de reprodução.

Devido ao fato de que você deseja um nível de confiança de 95%, seu Z é igual a 1,96.

Suponha que você colete aleatoriamente uma amostra de 100 dourados e determine que o tamanho médio dos peixes é de 7,5 polegadas (19,05 cm) e o desvio padrão (*s*) é igual a 2,3 polegadas (5,84 cm) (veja, no Capítulo 4, como calcular a média e o desvio padrão). Isso significa que $\bar{x} = 7{,}5$, $s = 2{,}3$ e $n = 100$.

Multiplique 1,96 por 2,3 e divida o resultado pela raiz quadrado de 100 (10). A margem de erro, então, será $\pm\ 1{,}96 \times (2{,}3 \div 10) = 1{,}96 \times 0{,}23 = 0{,}45$ polegadas (aproximadamente 1,15 cm).

Seu intervalo de confiança de 95% para o tamanho médio dos dourados no viveiro de reprodução é 7,5 polegadas (19,05 cm) mais ou menos 0,45 polegadas (1,15 cm). (A ponta inferior do intervalo é 7,5 – 0,45 = 7,05 polegadas, ou 17,90 cm; a ponta superior do intervalo é 7,5 + 0,45 = 7,95 polegadas, 20,19 cm). Você pode dizer então, com 95% de confiança, que o tamanho médio dos dourados em todo o viveiro de reprodução é de 7,05 (17,90 cm) a 7,95 polegadas (20,19 cm), de acordo com sua amostra.

Quando seu tamanho amostral é pequeno (menor do que 30), será necessária uma pequena modificação em seus cálculos. Esse assunto será discutido no Capítulo 15.

Determinando o Intervalo de Confiança para a Proporção populacional

Quando a característica que está sendo estudada for categórica (por exemplo, a opinião sobre determinada questão [apoio, oposição ou neutra], sexo, partido político ou tipo de comportamento [usa ou não o cinto de segurança enquanto está dirigindo]), a maioria das pessoas prefere relatá-la por meio da proporção (ou porcentagem) de pessoas da população que se enquadra em uma determinada categoria de interesse. Por exemplo, a porcentagem de pessoas a favor de que haja apenas quatro dias úteis por semana, a porcentagem de Republicanos que votaram na última eleição ou a proporção de motoristas que não usam o cinto de segurança. Em cada um desses casos, o objetivo é estimar a proporção da população usando uma proporção amostral mais ou menos uma margem de erro que resultará no intervalo de confiança para a proporção populacional.

A fórmula para o IC para uma proporção populacional é $\hat{p} \pm Z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, em que $\hat{p}$ é a proporção amostral, n é o tamanho amostral e Z é o valor correspondente ao nível de confiança desejado, segundo a distribuição normal padrão (veja o Capítulo 3 para fórmulas e cálculos para $\hat{p}$; veja, no Capítulo, a Tabela 10-1 para os valores de Z correspondentes a diferentes níveis de confiança).

Para calcular o IC para a proporção populacional:

1. **Determine o nível de confiança e encontre o valor Z correspondente.**

   Veja a Tabela 10-1, no Capítulo 10

2. **Encontre a proporção amostral ($\hat{p}$) dividindo o número de pessoas na amostra que possui a característica a ser estudada pelo tamanho amostral (*n*).**

   Observe: $\hat{p}$ deve ser um valor decimal entre 0 e 1.

3. **Multiplique $\hat{p}$ por (1- $\hat{p}$) e, então, divida a diferença pelo valor de *n*.**

4. **Calcule a raiz quadrada para o resultado obtido no passo anterior.**

5. **Multiplique o resultado obtido em 4 pelo valor Z.**

   Essa será sua margem de erro

6. **Some e subtraia a margem de erro do valor de $\hat{p}$ para obter o IC. A ponta inferior do IC será $\hat{p}$ menos a margem de erro e a ponta superior será $\hat{p}$ mais a margem de erro.**

Por exemplo, suponha que você queira estimar a porcentagem de vezes que você teve que parar em um sinal vermelho em um determinado cruzamento.

Devido ao fato de que o nível de confiança desejado é de 95%, seu valor Z é igual a 1,96.

Você selecionou aleatoriamente 100 passagens por esse cruzamento e descobriu que, em 53 vezes, você teve que parar no sinal vermelho, portanto $\hat{p} = 53 \div 100 = 0{,}53$.

Multiplique 0,53 por 1- 0,53 e divida a diferença por 100 para obter $0{,}2491 \div 100 = 0{,}002491$.

Ache a raiz quadrada de $0{,}002491 = 0{,}0499$.

A margem de erro é, portanto, $\pm\, 1{,}96 \times (0{,}0499) = 0{,}0978$.

Seu intervalo de confiança para a porcentagem de vezes que você irá parar no sinal vermelho daquele cruzamento em particular é igual a 0,53 (ou 53%) mais ou menos 0,0978 (arredondado para 0,10 ou 10%). (A ponta inferior do intervalo é $0{,}53 - 0{,}10 = 0{,}43$ ou 43%; a ponta superior é igual a $0{,}53 + 0{,}10 = 0{,}63$ ou 63%). Em outras palavras, a porcentagem de vezes em que você pode ter que parar no sinal vermelho daquele cruzamento fica em algum ponto entre 43% e 63%, segundo a sua amostra.

DICA

Enquanto você estiver fazendo qualquer cálculo que envolva porcentagens amostrais, use a forma decimal. Depois que todos os cálculos estiverem prontos, transforme os decimais em porcentagens multiplicando-os por 100. Para evitar erros cometidos por arredondamento, mantenha pelo menos 2 casas decimais durante todo o cálculo.

Desenvolvendo um Intervalo de Confiança para a Diferença de Duas Médias

O objetivo de muitas pesquisas é o de comparar duas populações, como homens versus mulheres, rendas familiares altas versus rendas familiares baixas e Republicanos versus Democratas. Quando a característica a ser estudada for numérica (por exemplo, a altura, o peso ou a renda), o objeto de interesse é a diferença entre as médias das duas populações. Por exemplo, imagine que você queira comparar a diferença na média de idade de Republicanos versus Democratas ou a diferença na renda média de homens e mulheres. A diferença de duas médias populacionais pode ser estimada através da diferença das médias amostrais retiradas de uma amostra de cada população mais ou menos uma margem de erro. O resultado desse procedimento é conhecido como o intervalo de confiança para a diferença de duas médias populacionais.

A fórmula para o IC para a diferença entre duas médias populacionais é $(\bar{x}-\bar{y}) \pm Z \times \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$, em que $\bar{x}$, s_1 e n_1 são a média, o desvio padrão e o tamanho amostral da primeira amostra e $\bar{y}$, s_2 e n_2 são a média, o desvio padrão e o tamanho amostral da segunda amostra. Z é o valor correspondente ao nível de confiança desejado, segundo a distribuição normal padrão (veja, no Capítulo 3, as fórmulas e cálculos para as médias e desvios padrões; veja, também, no Capitulo 10, a Tabela 10-1 para os valores Z correspondentes a alguns níveis de confiança).

Para calcular um IC para a diferença entre duas médias populacionais, faça o seguinte:

1. **Determine o nível de confiança e encontre o valor Z correspondente.**

   Veja, no Capítulo 10, a Tabela 10-1.

2. **Encontre a média ($\bar{x}$), o desvio padrão (s_1) e o tamanho amostral (n_1) para a primeira amostra e a média ($\bar{y}$), o desvio padrão (s_2) e o tamanho amostral (n_2) para a segunda amostra.**

   Veja o Capítulo 3.

3. **Encontre a diferença, $(\bar{x}-\bar{y})$, entre as médias amostrais.**

4. **Eleve s_1 ao quadrado e divida-o por n_1; eleve s_2 ao quadrado e divida-o por n_2. some os resultados e ache a raiz quadrada.**

5. **Multiplique o resultado obtido no passo anterior pelo valor Z.**

   Essa será sua margem de erro.

6. **Some e subtraia a margem de erro da diferença $(\bar{x}-\bar{y})$ para obter o IC.**

   A ponta inferior do IC é $(\bar{x}-\bar{y})$ menos a margem de erro, enquanto que a ponta superior do IC é $(\bar{x}-\bar{y})$ mais a margem de erro.

Suponha que você queria estimar com 95% de certeza a diferença entre o tamanho médio dos sabugos de duas variedades diferentes de milho verde (deixando com que eles se desenvolvam durante o mesmo número de dias e sob as mesmas condições). Vamos denominar essas variedades de Milho estatísticum e Verde estatísticum.

Devido ao fato de que o nível de confiança desejado é de 95%, o valor Z a ser usado será de 1,96.

Suponha que você selecione aleatoriamente 100 espigas do milho da variedade Milho estatísticum com tamanho médio de 8,5 polegadas (21,58 cm), com o desvio padrão de 2,3 polegadas (5,84 cm) e, também, selecione aleatoriamente 110 espigas da variedade Verde estatísticum, com tamanho médio de 7,5 polegadas (19,05 cm), com um desvio padrão de 2,8 polegadas (7,11 cm). Isso significa que $\bar{x} = 8{,}5$, $s_1 = 2{,}3$ e $n_1 = 100$; $\bar{y} = 7{,}5$, $s_2 = 2{,}8$ e $n_2 = 110$.

A diferença entre as médias amostrais, $(\bar{x} - \bar{y})$, do terceiro passo, é $8{,}5 - 7{,}5 = +1$ polegada (2,54 cm). Isso significa que a média para a variedade Milho estatísticum menos a média da variedade Verde estatísticum é positiva, fazendo a variedade Milho estatísticum ser a maior entre as duas, com relação a essa amostra. Mas será que essa diferença é o bastante para generalizar uma população inteira? É isso que esse intervalo de confiança irá lhe ajudar a decidir.

Eleve s_1 (2,3) ao quadrado e você terá o número 5,29; divida-o por 100 e chegue ao resultado 0, 0529. Eleve s_2 (2,80) ao quadrado e divida-o por 110: $7{,}84 \div 110 = 0{,}0713$. A soma é $0{,}0529 + 0{,}0713 = 0{,}1242$; a raiz quadrada desse resultado é 0, 3524.

Multiplique 1,96 por 0, 3524, para chegar a 0,69 polegadas (1,75 cm).

Seu intervalo de confiança de 95% para a diferença entre os tamanhos médios dessas duas variedades de milho é 1 polegadas (2,54 cm) mais ou menos 0,69 polegadas (1,75 cm). (A ponta inferior do intervalo é $1 - 0{,}69 = 0{,}31$ polegadas (0,78 cm); a ponta superior é $1 + 0{,}69 = 1{,}69$ polegadas (4,29 cm). (Note que todos os valores nesse intervalo são positivos. Isso significa que a variedade Milho estatísticum deverá sempre ser, na média, maior do que do que a variedade Verde estatísticum, segundo essa amostra).

Observe que você poderia ter encontrado um valor negativo para $(\bar{x} - \bar{y})$ Por exemplo, se você tivesse alterado a ordem das variedades, você teria obtido o valor -1 para a diferença entre as médias. Sem problema, apenas se lembre qual é cada grupo. A diferença positiva das médias indica que o primeiro grupo tem um valor maior do que o segundo; uma diferença negativa das médias indica que o primeiro grupo tem um valor menor do que o segundo. Caso você queira evitar valores negativos, sempre considere como o primeiro grupo aquele que tiver o maior valor – e, assim, todas suas diferenças serão positivas.

Caso seu tamanho amostral seja menor do que 30, veja, no Capítulo 15, as modificações necessárias para a execução desse cálculo.

Calculando o Intervalo de Confiança para a Diferença de Duas Proporções

Quando a característica, como a opinião acerca de determinado assunto (apoia/não apoia), de dois grupos que estão sendo comparados for categórica, as pessoas relatam as diferenças entre as duas proporções populacionais – por exemplo, a diferença entre a proporção de mulheres e a proporção de homens que são a favor de uma semana com apenas quatro dias úteis. Você pode estimar a diferença entre duas proporções da população através da coleta de amostras de cada uma das populações e da diferença entre as duas populações, com a margem de erro para mais ou menos. O resultado obtido desse procedimento é conhecido como intervalo de confiança para a diferença de duas proporções populacionais.

A fórmula para um intervalo de confiança para a diferença entre duas proporções populacionais é $\left(\hat{p}_1 - \hat{p}_2\right) \pm Z\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$, em que $\hat{p}_1$ e n_1 são a proporção amostral e o tamanho amostral da primeira amostra e $\hat{p}_2$ e n_2 são a proporção amostral e o tamanho amostral da segunda amostra. Z é o valor correspondente ao nível de confiança desejado, segundo a distribuição normal padrão (veja o Capítulo 3, para as proporções amostrais e, no Capítulo 10, a Tabela 10-1).

Para calcular o IC para a diferença entre duas proporções populacionais, faça o seguinte:

1. **Determine o nível de confiança e encontre o valor Z correspondente.**

   Veja, no Capítulo 10, a Tabela 10-1.

2. **Encontre a proporção amostral $\hat{p}_1$ para a primeira amostra através da divisão do número total dos indivíduos da primeira amostra que se enquadram na categoria de interesse pelo valor do tamanho amostral, n_1. Da mesma forma, encontre $\hat{p}_2$ para a segunda amostra.**

3. **Faça a diferença entre as proporções amostrais ($\hat{p}_1 - \hat{p}_2$).**

4. **Multiplique $\hat{p}_1$ por ($1 - \hat{p}_1$) e divida o resultado por n_1. Multiplique $\hat{p}_2$ por ($1 - \hat{p}_2$) e divida o resultado por n_2. Some os resultados e encontre a raiz quadrada.**

5. **Multiplique o resultado do passo anterior pelo valor Z.**

   Essa será sua margem de erro.

6. **Some e subtraia o resultado da diferença ($\hat{p}_1 - \hat{p}_2$) da margem de erro, calculada no passo número 5, para chegar ao IC.**

A ponta inferior do IC é ($\hat{p}_1 - \hat{p}_2$) menos a margem de erro e a ponta superior do IC é ($\hat{p}_1 - \hat{p}_2$) mais a margem de erro.

Enquanto estiver realizando cálculos que envolvam porcentagens, sempre use os números em sua forma decimal. Depois que os cálculos estiverem prontos, você pode converter os resultados em porcentagem, basta que você os multiplique por 100. Para evitar os erros causados por arredondamento, mantenha, pelo menos, duas casas decimais.

Suponha que você trabalhe para a Câmara de Comércio de Las Vegas e queira estimar com 95% de confiança a diferença entre a proporção de mulheres que já viram um sósia do Elvis e a porcentagem de homens que já viram um sósia do Elvis, a fim de ajudar-lhe a decidir como fazer o marketing de seus produtos.

Devido ao fato de que você deseja 95% de confiança, seu valo Z é 1,96.

Imagine que você tenha selecionado aleatoriamente uma amostra composta por 100 mulheres, em que 53 já assistiram a um show de um sósia do Elvis, portanto $\hat{p}_1$ é $53 \div 100 = 0{,}53$. Imagine também que você tenha selecionado aleatoriamente uma amostra de 110 homens, em que 37 já assistiram a um show de um sósia do Elvis, sendo assim, $\hat{p}_2$ é $37 \div 100 = 0{,}34$.

A diferença entre as proporções amostrais (mulheres – homens) é $0{,}53 - 0{,}34 = 0{,}19$.

Multiplicando 0,53 por $(1 - 0{,}53)$ e dividindo o resultado por 100, chegamos a $0{,} 2491 \div 100 = 0{,} 0025$. Depois, multiplique 0,34 por $(1 - 0{,}34)$ e divida o resultado por 110 para obter $0{,} 2244 \div 110 = 0{,} 0020$. Some esse dois resultados, $0{,} 0025 + 0{,} 0020 = 0{,} 0045$, e encontre sua raiz quadrada, 0, 0671.

$1{,}96 \times 0{,} 0671$ lhe dá o resultado 0,13, ou 13%, sua margem de erro.

Seu intervalo de confiança de 95% para a diferença entre a porcentagem de homens e mulheres que já assistiram a um show de um sósia do Elvis é 0,19, ou 19% (número que você obteve no 3° passo), mais ou menos 13%. A ponta inferior do intervalo é $0{,}19 - 0{,}13 = 0{,}06$ ou 6%; o ponto superior do intervalo é $0{,}19 + 0{,}13 = 0{,}32$, ou 32%. Logo, você pode afirmar com 95% de certeza que a porcentagem de mulheres que já assistiram a um sósia do Elvis é mais alta do que a porcentagem de homens e que a diferença entre essas porcentagens encontra-se em algum ponto entre 6% e 32%, segundo sua amostra. Agora, será que os homens realmente iriam admitir ter ido a um show de um sósia do Elvis? Isso pode gerar certa tendenciosidade nos resultados (a última vez em que estive em Las Vegas, eu realmente achei que tivesse visto o Elvis, ele estava dirigindo um táxi para o aeroporto...).

DICA

Observe que você pode obter um número negativo para $(\hat{p}_1 - \hat{p}_2)$. Por exemplo, se você alterasse a ordem de mulheres e homens, teria encontrado o valor – 0,19 para essa diferença. Uma diferença positiva significa que o primeiro grupo possui um valor maior do que o segundo grupo; uma diferença negativa significa que o primeiro grupo possui um valor menor do que o segundo. Você pode evitar as diferenças negativas se sempre considerar o grupo com o valor maior como sendo o primeiro.

Parte VI
Testando uma Hipótese

A 5ª Onda por Rich Tennant

Nesta Parte...

Muitas estatísticas formam as bases de algumas afirmações como: "Quatro em cinco dentistas recomendam esse chiclete" ou "Nossas fraldas absorvem 25% a mais do que a marca líder". Como você pode dizer que essas afirmações são verdadeiras? Os pesquisadores (que sabem o que estão fazendo) usam o que é chamado de teste de hipótese.

Nesta parte, você irá investigar os fundamentos dos testes de hipótese, determinado como estabelecê-los, conduzi-los e interpretar seus resultados (tudo o que é necessário saber para que você tente fazer uma afirmação sobre toda uma população, baseando-se apenas em uma amostra). Você também terá referências rápidas e exemplos para os testes de hipótese mais comuns.

Capítulo 14

Hipóteses, Testes e Conclusões

Neste Capítulo

- Testando os argumentos produzidos
- Usando a estatística como evidência
- Pesando as evidências e tomando decisões
- Descobrindo que você pode estar enganado

A todo o momento, ouvimos argumentos envolvendo estatísticas; a mídia não faz economia com relação a isso:

- Vinte e cinco por cento de todas as mulheres nos Estados Unidos têm varizes (puxa, há coisas que é melhor não saber!)
- O uso do êxtase entre adolescente diminuiu pela primeira vez nos últimos anos. O declínio avaliado durante um ano variou entre aproximadamente um décimo até um terço, dependendo da série escolar em que estavam os adolescentes.
- Um bebê de seis meses dorme em média de 14 a 15 horas em um período de 24 horas.(isso é verdade!)
- Uma mistura para torta leva apenas 5 minutos para ser preparada.

Muitos argumentos envolvem números que parecem sair do nada. Alguns argumentos fazem comparações entre um produto ou um grupo e outro. Talvez você se pergunte se tais argumentos são válidos, e realmente deveria se perguntar. Nem todos esses argumentos podem mudar sua vida (afinal de contas, qual o problema em usar um sabonete que não seja 99,99% puro?), mas alguns podem – por exemplo, qual tratamento para o câncer funciona melhor, quais utilitários são os mais seguros ou se determinados medicamentos devem ser aprovados para consumo. Enquanto muitos argumentos estão fundamentados em pesquisas científicas (e estatísticas), outros não. Neste capítulo, você irá descobrir como usar as estatísticas para determinar se um argumento é realmente válido e ficará por dentro do processo que os pesquisadores deveriam utilizar para validar todos os argumentos que eles fazem.

Respondendo aos Argumentos: Algumas Regras

Na era da informação (e de muito dinheiro), uma grande questão envolve a capacidade de embasar seus argumentos. As empresas que dizem que seus produtos são melhores do que a da marca líder devem ser capazes de provar o que dizem ou irão enfrentar processos judiciais. Os medicamentos que são aprovados pelo FDA (Food and Drug Administration) têm que mostrar evidências sólidas de que funcionam sem causar efeitos colaterais durante o tratamento. Os fabricantes devem garantir que seus produtos estejam sendo produzidos de acordo com especificações, de modo a evitar a necessidade de recalls, queixas dos consumidores e prejuízos.

A pesquisa também pode resultar em argumentos que podem significar a diferença entre a vida e a morte, tais como o melhor tratamento para o câncer, os efeitos colaterais mais comuns de um tipo de cirurgia, a taxa de sobrevivência para determinado tratamento e se um novo medicamento experimental pode ou não aumentar a expectativa de vida. A pesquisa que vai responder a essas perguntas precisa ser segura a tal ponto que se possa tomar a decisão mais apropriada (pelo menos a mais estatisticamente informativa). Caso contrário, os pesquisadores podem perder sua reputação, credibilidade e seu patrocínio (e, às vezes, sentem-se pressionados a produzir resultados, o que também pode levar a outros problemas).

Conhecendo suas opções

Como um consumidor na era da informação, todas as vezes que você ouvir um argumento como, por exemplo: “nosso sorvete foi a primeira escolha de 80% dos degustadores”; você tem basicamente três opções:

- Acreditar imediatamente (ou, pelo contrário, rejeitá-lo no mesmo instante)
- Conduzir seu próprio teste para verificar ou descartar o argumento
- Ir mais a fundo a fim de obter mais informações para que você possa tomar sua própria decisão.

Acreditar nos resultados sem questioná-los (ou rejeitá-los prontamente) não é inteligente; as únicas vezes em que você deve fazer isso é quando a fonte já lhe tenha deixado uma boa (ou péssima) imagem ou quando os resultados simplesmente não lhe interessam (afinal de contas, você não pode sair por aí checando todos os argumentos que cruzam seu caminho).

Afastando-se dos testemunhos

A segunda opção de reação a um argumento, a abordagem do teste-você-mesmo, é tomada por organizações, tais como o Instituto Gallup, que conduz suas próprias pesquisas; o Insurance Institute for Highway Safety (Instituto para a Segurança em Autoestradas), que realiza testes de colisão e faz relatos sobre a segurança dos veículos; Consumer Reports, que testa e relata a qualidade e os valores dos produtos; e o Good Housekeeping Institute (órgão semelhante ao Inmetro), que testa os produtos antes de dar o selo de aprovação.

A abordagem do teste-você-mesmo pode ser eficiente se feita da maneira correta, com resultados que se baseiam em estudos bem projetados e que coletam dados precisos e não tendenciosos (veja os capítulos 16 e 17, para mais informações sobre o planejamento de estudos).

Essa abordagem pode ser feita no trabalho. Por exemplo, um concorrente pode fazer declarações sobre seus produtos, que você não considera verdadeiras e que, por isso, deveriam ser testadas. Ou você acha que seu produto funciona melhor do que um produto da concorrência e, por isso, decide colocar os produtos à prova. Muitas indústrias também têm seu próprio controle de qualidade (veja o Capítulo 19), portanto elas testam seus produtos para ver se estão realmente dentro das especificações.

Ainda que essa opção seja viável para os grupos que tenha recursos e conhecimento para conduzir seu próprio estudo, a fim de testar um argumento, essa abordagem pode levar a resultados equivocados se executada de maneira inadequada.

Uma maneira usada pela mídia para testar os argumentos de produtos é enviar pessoas às ruas para verificar os produtos por si mesmos. Esse é um método ultrapassado e não científico (apesar de divertido) de testar uma hipótese. Por exemplo, suponha que um programa de TV determinou o que o mundo tem que saber: "Uma mistura para torta de determinada marca realmente leva apenas 5 minutos para ficar pronta?". Talvez ela leve mais, talvez leve menos. Estatisticamente falando, a variável de interesse nesse caso é numérica – tempo de preparo – e a população são todas as tortas feitas por aquela determinada marca. O parâmetro de interesse é o tempo médio de preparação para todas as tortas feitas com aquela mistura. (um parâmetro é um único número que resume a população e, normalmente, é sobre o argumento). O argumento aqui é que o tempo médio de preparo da torta é de apenas 5 minutos. A missão: testar esse argumento. Quantas tortas serão usadas? Chute – apenas uma!

Haveria câmeras percorrendo o palco e os apresentadores ficariam batendo papo a respeito do quanto é divertido fazer a torta, o como ela parece deliciosa sem tirar o olho do tempo para prepará-la (afinal de contas, logo eles teriam que chamar um intervalo comercial). No final, eles diriam que o tempo de preparo foi de 5,5 minutos, muito próximo do argumento, mas não exatamente igual. E eles acabariam com um comentário dizendo que uma ótima opção para a cobertura da torta seria uma barra de chocolate Snickers (o que, por sinal, é mesmo).

Se esses programas de TV tivessem um residente em estatística que pudesse fornecer a divisão estatística dos resultados (o que causaria um estouro do orçamento, eu sei), eu me candidataria para a vaga. A principal ideia que eu gostaria de passar à plateia é de que os resultados de uma amostra variam (de pessoa para pessoa, de torta para torta) – veja o Capítulo 9, para saber mais a esse respeito. A estatística entra em campo para permitir que essa variabilidade seja medida e entendida. O importante é: para se conseguir resultados confiáveis e conclusivos sobre qualquer argumento, são necessários dados (ou seja, mais do que uma única observação). A maioria das pessoas não se dá conta de que para testar um argumento da maneira correta, é necessário mais do que um tamanho amostral composto por 1 (ou, até mesmo, 2 ou 3) indivíduo, devido ao fato de que os resultados amostrais variam.

Você não pode (ou pelo menos, não deveria) construir nenhum tipo de conclusão duradoura a partir de um testemunho, que é o que uma amostra com 1 indivíduo realmente é. Em estatística, uma amostra composta por um indivíduo não faz sentido algum. Não é possível medir a variabilidade com apenas um valor (veja o Capítulo 5, para a fórmula do desvio padrão para entender o que eu estou querendo dizer). Esse é o problema com a maioria dos programas de TV que mostram as pessoas testando argumentos com o uso de apenas um ou dois produtos; eles não estão realizando um teste científico e ainda estão enviando a mensagem errada sobre como testar uma hipótese. Agora, embora tirar conclusões sobre as tortas de cinco minutos sem dados suficientes não pareça algo que mude o mundo, pense em quantas vezes a experiência de apenas uma única pessoa influenciou-lhe a tomada de alguma decisão em sua vida.

Fique atento com os resultados de estudos que se baseiam em tamanhos amostrais extremamente pequenos, especialmente aqueles que se baseiam em amostras formadas por somente um indivíduo. Por exemplo, se um estudo envia uma pessoa para testar uma embalagem de carne, examinar um brinquedo de criança ou testar a precisão de uma única farmácia em preencher uma prescrição em um determinado dia, afaste-se dele. Isso serve para criar histórias divertidas e pode revelar problemas que devem ser investigados com mais cuidado. No entanto, esses resultados sozinhos não são científicos, e você não deveria tirar conclusões baseadas neles.

Cavando a fundo

Cavar a fundo, a fim de obter mais informação sobre os argumentos que lhe interessam, é o que você deve fazer. É necessário que você encontre mais informações para questionar e tomar uma decisão bem embasada.

A grande diferença entre os testes estatísticos de uma hipótese e os testes que encontramos por aí é que o bom teste usa dados coletados de maneira científica, sem tendenciosidade e baseiam-se em amostras aleatórias grandes o suficiente para se obter informações precisas. (Veja mais a esse respeito no Capítulo 2). A maioria das pesquisas científicas, incluindo pesquisas médicas, farmacêuticas, de engenharia e governamentais, baseia-se no uso de estudos estatísticos para testar, verificar ou descartar argumentos de vários tipos. Como consumidores de grande parte dessa informação, que, muitas vezes, nos é dada em flashes, necessitamos ser capazes de saber o que procurar, a fim de avaliar o estudo, entender os resultados e tomar nossas próprias decisões acerca do argumento que está sendo feito.

Talvez, você, como consumidor, deva estar se perguntando se está protegido contra os argumentos disparados por todas essas pesquisas. O governo dos Estados Unidos regula e monitora uma grande parte das pesquisas e produções que são realizadas (por exemplo, O FDA regula as pesquisas e distribuição de medicamentos, o USDA monitora a produção de alimentos, e assim por diante). Mas, algumas áreas, tais como a dos suplementos alimentares (vitaminas, suplementos minerais e herbáceos), não são regulamentadas com tanto rigor.

Como consumidores de todos os resultados que nos são atirados, necessitamos estar armados com informações para tomar boas decisões. O primeiro passo é entrar em contato com o pesquisador (ou jornalista) para saber se existem estudos científicos para embasar o seu argumento. Caso a resposta seja sim, pergunte se você pode ver as descrições e os resultados desses estudos e, então, avalie criticamente a informação que lhe foi dada.

Fazendo um Teste de Hipótese

Um teste de hipótese é um procedimento estatístico desenvolvido para testar um argumento. Normalmente, o argumento é feito sobre um parâmetro populacional (um número que caracteriza toda a população). Devido ao fato de que os parâmetros tendem a ser quantidades desconhecidas, todos querem fazer um argumento sobre quais são seus valores. Por exemplo, o argumento de que 25% (ou 0,25) de todas as mulheres têm varizes é um argumento sobre a proporção (o parâmetro) de todas as mulheres (a população) que têm varizes (a variável, ter o não ter varizes).

Você acha que alguém realmente sabe com certeza que a porcentagem de mulheres que têm varizes é exatamente 25%? É claro que não; eles apenas fazem um argumento, não afirmam de fato. Tenha cuidado com declarações como essa.

Definindo o seu objeto de teste

Para sermos mais específicos, o argumento sobre as varizes é que o parâmetro, a proporção da população (p), é igual a 0,25 (esse argumento é chamado de hipótese nula). Se sairmos para testar esse argumento, estaremos questionando o argumento e conseguiremos nossa própria hipótese (chamada de hipótese de pesquisa ou hipótese alternativa). Você pode estabelecer uma hipótese, por exemplo, de que a real proporção de mulheres que têm varizes é menor do que 0,25, segundo suas observações. Ou você ainda pode estabelecer a hipótese de que, em virtude da popularidade dos sapatos de salto alto, a proporção pode ser ainda maior do que 0,25. Ou, se você apenas estiver questionando se a verdadeira proporção é mesmo 0,25, sua hipótese alternativa é:"Não, não é 0,25".

Além de testar as hipóteses para variáveis categoriais (ter ou não ter varizes é uma variável categorial), você também pode testar hipóteses sobre variáveis numéricas, tais como o tempo médio que as pessoas gastam para ir ao trabalho em Los Angeles ou sua renda familiar média. Nesses casos, o parâmetro a ser estudado é a média populacional (denotada como μ). Mais uma vez, o argumento será de que esse parâmetro é igual a determinado valor versus um valor alternativo.

As hipóteses também podem ser testadas para mais de um único parâmetro. Por exemplo, você talvez queira comparar as rendas familiares médias ou o tempo que as pessoas gastam para ir ao trabalho em duas ou mais cidades. Ou, talvez, você queira ver se há uma relação entre a renda familiar e o tempo gasto para chegar ao trabalho. Todas essas perguntas podem ser respondidas por meio do teste de hipóteses; ainda que os detalhes diferenciem-se em cada situação, a ideia geral é a mesma. Neste capítulo, veremos um caso composto por apenas uma amostra para médias e proporções (amostras grandes); no Capítulo 15, veremos as particularidades dos testes de hipótese mais comuns.

Configurando a hipótese

Todos os testes de hipótese contêm duas hipóteses. A primeira hipótese é chamada de hipótese nula, denotada por H_0. A hipótese nula sempre afirma que o parâmetro da população é igual ao valor do argumento. Por exemplo, se o argumento é de que o tempo médio de preparo de uma mistura para torta de determinada marca é de cinco minutos, a equação estatística para a hipótese nula é a seguinte: H_0: $\mu = 5$.

Qual é a alternativa?

Antes mesmo de conduzir um teste de hipótese, é necessário estabelecer duas hipóteses possíveis – a hipótese nula é uma delas. Mas, se a hipótese nula não for verdadeira, qual será a alternativa? Na verdade, existem três possibilidades para a segunda hipótese (ou hipótese alternativa), denotada como H_a. São elas, juntamente com a equação no contexto do exemplo:

- O parâmetro da população não é igual ao valor do argumento (H_a: $\mu \neq 5$).
- O parâmetro da população é maior do que o valor do argumento (H_a: $\mu > 5$).
- O parâmetro da população é menor do que o valor do argumento (H_a: $\mu < 5$).

A hipótese alternativa que você deve escolher ao estabelecer seu teste de hipótese depende do que você pretende concluir, caso tenha evidências o suficiente para desmentir a hipótese nula (o argumento).

Por exemplo, se você quiser testar se uma empresa está correta ao afirmar que o tempo médio de preparo de sua torta é de 5 minutos e, além disso, quiser saber se o tempo médio é maior ou menor do que o mencionado, você deve usar a alternativa de não igualdade. Suas hipóteses para esse teste seriam H_0: $\mu = 5$ versus H_a: $\mu \neq 5$.

Se você apenas quer ver se o tempo médio de preparo é maior do que o que a empresa afirma (ou seja, a empresa está fazendo uma propaganda enganosa), use a alternativa de superioridade e suas duas hipóteses serão H_0: $\mu = 5$ versus H_a: $\mu > 5$.

Por fim, digamos que você trabalhe para a empresa que comercializa a torta e acha que a torta pode ser feita em menos de 5 minutos (e poderia ser vendida assim pela empresa). A alternativa de inferioridade é a apropriada para você e suas duas hipóteses seriam H_0: $\mu = 5$ versus H_a: $\mu < 5$.

Qual hipótese é qual?

Como saber qual hipótese pôr em H_0 e qual pôr em H_a? Normalmente, a hipótese nula diz que nada de novo irá acontecer, os resultados anteriores são os mesmos agora como foram os mesmos de antes, ou os grupos possuem a mesma média (sua diferença é igual a zero). E, geralmente, você supõe que os argumentos das pessoas são verdadeiros até que se prove o contrário.

DICA

Os testes de hipótese funcionam como um tribunal, em que H_0 é semelhante ao veredicto de inocência e H_a é o veredicto de culpabilidade. Em um tribunal, sempre se parte do princípio de que o réu não é culpado até que se prove o contrário. Se o júri diz que as evidências não estão claras, eles rejeitam H_0, não culpado, em favor de H_a, culpado.

Em geral, quando a hipótese é testada, você estabelece H_0 e H_a para que você acredite que H_0 é verdadeiro, a menos que suas evidências (seus dados e suas estatísticas) provem o contrário. E caso você tenha evidências suficientes contra H_0, você o rejeita em favor de H_a. O ônus da prova está com o pesquisador, que terá que mostrar evidências científicas contra H_0 para, então, rejeitá-lo (é por isso que H_a é chamado de hipótese de pesquisa, pois é a hipótese de interesse do pesquisador). Se H_0 é rejeitada em favor de H_a, o pesquisador pode dizer que encontrou resultados estatísticos significativos; ou seja, seus resultados desmentem o argumento anterior e algo novo começa a surgir.

FIQUE ATENTO

Em muitos casos, as pessoas estabelecem testes de hipótese, pois estão em campo para provar que H_0 é falso e sustentar a hipótese alternativa (a mentalidade aqui é, por que fazer uma pesquisa apena para mostrar que algo permaneceu o mesmo?) Os resultados que você ouve na imprensa são geralmente os capazes de mostrar que H_0 não é verdadeiro. Afinal, é isso que gera uma notícia. Em muitos casos, isso é algo bom, pois fabricantes e cientistas têm que fazer o máximo para evitar uma publicidade negativa, como o recall de um produto, um processo judicial ou uma investigação governamental. Tudo porque se um de seus argumentos (H_0) for rejeitado por alguém que conduziu um teste de hipótese independente, cientistas e fabricantes serão julgados culpados por fazer propaganda enganosa, o que não é nada bom.

Reunindo provas: a amostra

Depois de estabelecer a hipótese, o próximo passo é coletar as evidências e verificar se elas corroboram ou não o argumento feito em H_0. Lembre-se que o argumento refere-se a uma população, mas testar uma população inteira não é possível; o melhor que você tem a fazer, então, é coletar uma amostra. Assim como em qualquer outra situação em que seja necessário coletar estatísticas, a qualidade dos dados é de extrema importância (veja o Capítulo 2, para muitos exemplos de estatísticas que não deram certo).

Bons dados começam com uma boa amostra. As duas principais questões a serem consideradas quando você for selecionar sua amostra é a prevenção da parcialidade e a obtenção da precisão. Para evitar a parcialidade, colete uma amostra aleatória (o que significa que todos na população de interesse devem ter as mesmas chances de serem escolhidos) e escolha um tamanho amostral grande o suficiente para que, desta forma, os resultados sejam precisos (veja o Capítulo 3).

Compilando as evidências: a estatística

Depois de ter selecionado a amostra, você deve começar a realizar os cálculos apropriados. Sua hipótese nula faz uma declaração sobre o valor de um parâmetro populacional (por exemplo, a proporção de todas as mulheres que têm varizes ou a distância percorrida, em quilômetros, por litro por uma caminhonete). No jargão estatístico, os dados que você coleta medem a variável de interesse, e as estatísticas que você calcula incluirão a estatística amostral que mais se aproxima do parâmetro populacional. Isso quer dizer que, se você estiver testando o argumento que diz respeito à proporção de mulheres que têm varizes, é preciso calcular a proporção de mulheres com varizes em sua amostra. Se você estiver testando um argumento que diz respeito ao número de quilômetros que uma caminhonete faz com um litro de combustível, sua estatística deverá ser a média de quilômetros por litro das caminhonetes em sua amostra (veja o Capítulo 5 para todas as informações necessárias para o cálculo das estatísticas).

Padronizando as evidências: a estatística de teste

Depois de obter sua estatística amostral, você pode pensar que já fez a parte de análise e está pronto para chegar a uma conclusão – mas, na verdade, você não está. O problema é que você não tem como colocar seus resultados em perspectiva se apenas os observar em suas unidades.

Isso porque você sabe que seus resultados fundamentam-se em uma única amostra e que os resultados amostrais irão variar. Tal variação precisa ser levada em consideração ou suas conclusões poderão estar completamente equivocadas. Quanto os resultados amostrais variam? A variação amostral é medida em erros padrões; veja o Capítulo 9, para saber mais a esse respeito.

Suponha que o argumento seja de que a porcentagem de todas as mulheres com varizes seja de 25% e sua amostra de 100 mulheres apresentou 20% de mulheres com varizes. O erro padrão para sua porcentagem amostral é de 4 por cento (segundo as fórmulas apresentadas no Capítulo 9) e seus resultados podem variar até duas vezes esse valor, ou aproximadamente 8 por cento, segundo a regra empírica (veja o Capítulo 10). Portanto, uma diferença de 5 por cento entre o argumento e seu resultado amostral (25% - 20% = 5%) não é muito grande, nesses termos. Isso representa uma distancia menor do que 2 erros padrões do argumento. Portanto, você deve aceitar o argumento, H_o, pois seus dados não foram capazes de desmenti-lo.

No entanto, suponha que sua porcentagem amostral tenha se baseado em uma amostra composta por 1.000 mulheres e não 100. Isso diminui a quantia pela qual se espera que os resultados variem, pois, nesse caso, você possui mais informações. O erro padrão agora é de 0, 012 ou 1,2% e a margem de erro é duas vezes esse valor, ou seja, 2,4% para mais ou para menos. Nessa situação, a diferença de 5% entre os resultados amostrais (20%) e o argumento (25%) é uma diferença mais significativa; ela representa mais de 2 erros padrões com relação ao argumento. Os resultados baseados em uma amostra com 1.000 pessoas não deveriam variar tanto assim com relação ao argumento. Então, o que podemos concluir? Que o argumento (H_o) é falso, pois seus dados não o sustentaram.

O número de erros padrões em que se encontra uma estatística (seja acima ou abaixo da média) é chamado de escore padrão (veja o Capítulo 8). A fim de interpretar sua estatística, você precisa convertê-la de sua unidade original em escores padrões. Quando encontrar o escore padrão, subtraia a média e divida o resultado pelo erro padrão. No caso dos testes de hipótese, você irá usar o valor de H_o como sendo a média (isso porque você supõe que H_o é verdadeiro até ter provas do contrário). Essa versão padronizada de sua estatística é chamada de estatística de teste, e é o principal componente de um teste de hipótese (o Capítulo 15 contém as fórmulas para os testes de hipótese mais comuns).

O procedimento geral para a conversão de uma estatística em uma estatística de teste (escore padrão) no caso de médias/proporções:

1. **Subtraia de sua estatística o valor do argumento (representado por H_o).**

2. **Divida pelo erro padrão da estatística (veja os Capítulo 9 e 10).**

Sua estatística de teste representa a distância entre os seus resultados amostrais reais e o valor do argumento feito a respeito da população, em termos de números de erros padrões. No caso de uma única média ou proporção populacional, você sabe que essas distâncias padronizadas deveriam assumir uma distribuição normal padrão se o tamanho de sua amostra fosse grande o suficiente (veja o Capítulo 8 e 9). Portanto, para interpretar sua estatística de teste, nesses casos, você precisa ver se ela se encontra na distribuição normal padrão (distribuição Z).

Embora você nunca espere que uma estatística amostral seja exatamente igual ao valor da população, você espera que ela seja próxima, caso H_o seja verdadeiro. Isso significa que, caso você veja que a distância entre o argumento e a estatística amostral seja pequena, em termos de erros padrões, sua amostra não estará longe do argumento e seus dados estarão lhe dizendo para confiar em H_o. No entanto, conforme a distância começa a ficar cada vez maior, seus dados estarão mostrando cada vez menos apoio a H_o. Em determinado ponto, você deverá rejeitar H_o, baseando-se em suas evidências, e optar por H_a. Mas em que ponto isso acontece? Veja a resposta na próxima seção.

Pesando as Evidências e Tomando Decisões: os P-Valores

Para testar se um argumento é verdadeiro, você irá observar a estatística de teste retirada a partir de uma amostra e verificar se ela comprova esse argumento. Mas como você determinará isso? Para a maioria dos casos, você deve olhar onde sua estatística de teste acaba na distribuição amostral padrão (distribuição Z) – veja o Capítulo 9. A distribuição Z tem média igual a 0 e o desvio padrão igual a 1. Se sua estatística de teste for próxima a 0, ou, pelo menos, ficar dentro da variação em que a maioria dos resultados se encontra, então você pode dizer que sim, que o argumento (H_o) provavelmente é verdadeiro, e os resultado amostrais comprovam isso. Se sua estatística de teste estiver fora das extremidades da distribuição normal padrão, então você pode dizer não, pois as chances de meus resultados amostrais acabarem tão longe da distribuição são pequenas demais; portanto, meus resultados não comprovam o argumento (H_o).

Mas que distância pode ser considerada "longe demais" de 0? Desde que você tenha uma amostra grande o bastante, você sabe que sua estatística de teste se encontrará em algum lugar dentro de uma distribuição normal padrão, segundo o que diz o teorema do limite central (veja o Capítulo 10). Se a hipótese nula for verdadeira, a maioria (cerca de 95%) das amostras irá resultar em uma estatística de teste que se encontra aproximadamente há 2 erros padrões do argumento. Se H_a for a alternativa de não igualdade, qualquer teste que fique fora dessa variação resultará na rejeição de H_o (veja a Figura 14-1).

PAPO TÉCNICO

Observe que se a hipótese alternativa for a de inferioridade, você rejeitará o H_o apenas se a estatística de teste se enquadrar na extremidade esquerda da distribuição. Da mesma forma, se H_a for a alternativa de superioridade, você rejeitará H_o somente se a estatística de teste se encontrar na extremidade direita.

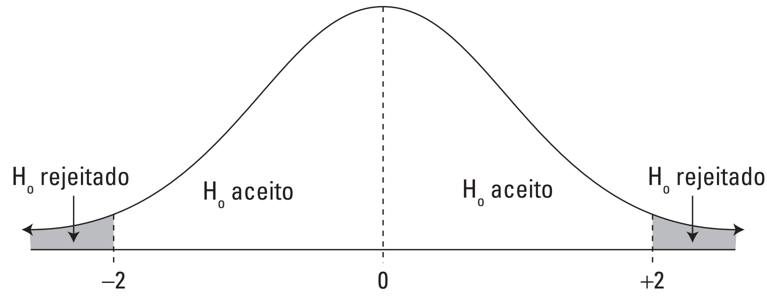


Figura 14-1: A estatística de teste e sua decisão.

Fundamentos do P-valor

Você pode ser ainda mais específico acerca de suas conclusões se observar exatamente a distância em que se encontra a estatística de teste dentro da distribuição normal padrão. Dessa forma, todos saberão onde estão os resultados e o que isso significa em termos de força para as evidências contra o argumento. Para fazer isso, é necessário consultar a estatística de teste na distribuição normal padrão (distribuição Z) e descobrir a probabilidade de que ela esteja naquele valor ou além dele (mas na mesma direção), usando a Tabela 8-1 (veja o Capítulo 8). O *p*-valor mede a probabilidade de que você tenha tirado seus resultados amostrais de uma hipótese nula verdadeira. Quanto mais longe sua estatística de teste estiver com relação às extremidades da distribuição normal padrão, menor será o *p*-valor e mais evidências você terá contra a veracidade da hipótese nula.

Todos os *p*-valores são probabilidades entre 0 e 1.

Para encontrar o *p*-valor para a sua estatística de teste (médias/proporções de amostras grandes):

1. **Consulte a localização de sua estatística de teste dentro da distribuição normal padrão (veja a Tabela 8-1, no Capítulo 8).**

2. **Encontre a porcentagem de chances de que ela esteja naquele valor ou além dele na mesma direção:**

   a. Se H_a contiver uma alternativa de inferioridade, encontre o percentil na Tabela 8-1, no Capítulo 8, que corresponda a sua estatística de teste.

   b. Se H_a contiver uma alternativa de superioridade, encontre o percentil da Tabela 8-1, no Capítulo 8, que corresponda a sua estatística de teste e, depois, subtraia 100% (nesse caso, você vai precisar da porcentagem à direita de sua estatística de teste e os percentis lhe dão a porcentagem apenas do lado esquerdo. Veja o Capítulo 5).

3. **Multiplique essa porcentagem por dois se (e apenas se) H_a for a alternativa de não igualdade.**

   Isso fará com que você abranja as possibilidades de superioridade e inferioridade.

4. **Transforme a porcentagem em uma probabilidade dividindo-a por 100 ou movendo uma cada decimal para a esquerda.**

Para interpretar o *p*-valor:

- Para *p*-valores pequenos (geralmente menores do que 0,05), rejeite H_o. Seus dados não o confirmarão e suas evidências estarão além de uma dúvida aceitável.
- Para *p*-valores grandes (geralmente maiores do que 0,05), você não deve rejeitar H_o. Você não conseguiu evidências o suficiente contra ele.
- Se seu *p*-valor estiver próximo ou na linha limite entre a aceitação e a rejeição, seus resultados são marginais (eles podem ir para qualquer lado).

De modo geral, os estatísticos ficam com o H_o, a menos que as evidências deem margem para dúvidas, assim como em um tribunal. O que, provavelmente, reflete o ponto de corte? Pode ser algo bastante arbitrário (o termo "p-valor pequeno" pode significar valores diferentes a cada pessoa). Para a maioria dos estatísticos, se o *p*-valor for menor do que 0,05, para os dados por eles coletados, o H_o irá ser rejeitado e H_a será o escolhido. Algumas pessoas podem ter pontos de corte ainda mais rigorosos, tais como 0,01, exigindo, assim, mais evidências para a rejeição do H_o. Cada leitor toma sua própria decisão. É por esse motivo que os pesquisadores precisam mencionar também os *p*-valores, e não apenas as suas decisões. Dessa forma, as pessoas poderão tirar suas próprias conclusões, baseando-se em seus pontos de corte subjetivos. Por exemplo, se o seu *p*-valor foi 0, 026 para o teste de hipótese $H_o: p = 0{,}25$ versus $H_a: p < 0{,}25$ no caso das mulheres com varizes, um leitor, com o limite de corte subjetivo de 0,05 concluiria que H_o é falso, pois o *p*-valor é menor do que 0,05. Por outro lado, um leitor com um limite de corte subjetivo igual a 0,01 não teria evidências o suficiente (baseadas em sua amostra) para rejeitar H_o, pois um *p*-valor de 0,026 é maior do que 0,01.

Atenção: As interpretações variam!

Algumas pessoas realmente gostam de estabelecer uma probabilidade de corte antes de fazer um teste de hipótese; isso é o que chamamos de nível alfa (α). Os valores mais comuns para α são 0,05 ou 0,01. Veja a seguir como interpretam seus resultados nesse caso:

- Se o *p*-valor for maior ou igual a α, aceite H_o.
- Se o *p*-valor for menor do que α, rejeite H_o
- Os *p*-valores que ficarem no limite (muito próximos à α) são considerados como resultados marginais.

Outros preferem não estabelecer um ponto de corte predeterminado; apenas mencionam o *p*-valor e interpretam seus resultados ao examinar seu tamanho. Geralmente:

- Se o *p*-valor for menor do que 0,01 (muito pequeno), os resultados são considerados altamente significativos estatisticamente – H_o é rejeitado.
- Se o *p*-valor ficar entre 0,05 e 0,01 (mas não próximo de 0,05), os resultados são considerados significativos estatisticamente – H_o é rejeitado.
- Se o *p*-valor for próximo a 0,05, os resultados são considerados marginalmente significativos – a decisão pode tomar os dois rumos.
- Se o *p*-valor for maior do que 0,05 (mas não próximo a 0,05), os resultados não são considerados significativos – H_o é aceito.

FIQUE ATENTO

Quando alguém lhe disser que um resultado é estatisticamente significativo, peça o *p*-valor e tome sua própria decisão. Os pontos de cortes e as decisões resultantes variam de pesquisador para pesquisador.

Sabendo Que Você Pode Estar Errado: Erros de Teste

Depois de tomada a decisão com relação a aceitar ou rejeitar H_o, o próximo passo é viver com suas consequências, no que se refere a como as pessoas irão reagir à sua decisão.

- Se você concluir que um argumento não é verdadeiro, quando, na verdade, ele é, isso resultará em um processo judicial, em multa, em alterações desnecessárias no produto, ou no boicote dos consumidores que não deveria ter acontecido?
- Se você concluir que um argumento é verdadeiro, quando, na verdade, não é, o que irá acontecer? Os produtos continuarão a ser produzidos da mesma forma que agora? Nenhuma ação será tomada, nem leis serão criadas, por que você mostrou que nada estava errado?

LEMBRE-SE

Todas as decisões dos testes de hipóteses têm um impacto, se não fosse assim, para que serviriam os testes?

Portanto, uma consequência pode ser o resultado de uma decisão: você poderia estar errado! O slogan do Arquivo-X cai muito bem aqui: "A verdade está lá fora". Mas, o fato é que você não sabe qual é a verdade; por isso fez um teste de hipótese.

Alarme Falso: erro tipo-1

Imagine que uma empresa alega que o tempo médio de entrega de encomendas é de 2 dias, e que um grupo de consumidores testa essa hipótese e conclui que a alegação é falsa: eles acreditam que o tempo médio de entrega é, na verdade, mais do que 2 dias. Temos aqui um grande problema. Caso o grupo possa provar com suas estatísticas, ele faria um bem se informasse o público a respeito dessa propaganda enganosa. Mas, e se o grupo estiver errado? Ainda que o estudo baseie-se em um bom planejamento, colete bons dados e faça a análise corretamente, o grupo ainda pode estar errado.

Por quê? Porque suas conclusões foram baseadas em uma amostra de encomendas e não em toda a população de encomendas. O Capítulo 9 fala: os resultados amostrais variam de amostra para amostra. Se sua estatística de teste se situa na extremidade da distribuição normal padrão, isso significa que esses resultados são atípicos, se a alegação for verdadeira, pois, nesse caso, esperamos que os resultados encontrem-se muito próximos do centro da distribuição normal padrão (distribuição Z) . O simples fato de os resultados de uma amostra serem atípicos não significa que eles sejam impossíveis. Um *p*-valor de 0,04 significa que as chances de que sua estatística de teste esteja fora da extremidade da distribuição normal padrão, mesmo se a alegação for verdadeira, é de 4% (menos do que 5%). Ou seja, você rejeitaria H_o, nesse caso, pelo fato de as chances serem muito pequenas. Mas uma chance é sempre uma chance!

Talvez, sua amostra, ainda que coletada aleatoriamente, seja uma daquelas amostras atípicas, cujo resultado acaba ficando distante do centro em uma distribuição normal padrão. Portanto, H_o poderia ser verdadeiro, porém, seus resultados o levaram a uma conclusão diferente. Mas, com que frequência isso acontece? Em cinco por cento das vezes (ou, seja qual for sua probabilidade de corte para a rejeição do H_o).

O erro que cometemos ao rejeitar o H_o quando, na verdade, não deveríamos, é chamado de erro tipo 1. Eu particularmente não gosto desse nome, pois ele me parece muito indefinido. Prefiro chamá-lo de alarme falso. No caso das encomendas, se o grupo de consumidores cometeu um erro do tipo 1 quando rejeitou a alegação da empresa, este grupo gerou um alarme falso. Qual é o resultado disso? Uma empresa de entregas de encomendas muito brava, eu garanto!

Faltando uma detecção: erro tipo-2

Por outro lado, suponha que a empresa realmente não entregue as encomendas de acordo com o que alega. Quem vai dizer que a amostra daquele grupo de consumidores irá detectar isso? Se o tempo real para a entrega é de 2,1 dias, ao invés de 2 dias, a diferença seria muito difícil de ser identificada. Se o tempo real para a entrega é de 3 dias, uma amostra bastante pequena seria capaz de mostrar que algo está errado e a questão se encontraria naqueles valores intermediários como 2,5 dias. Se H_o realmente é falso, você deve descobrir a verdade e rejeitá-lo. O erro que cometemos por não rejeitar H_o quando deveríamos é chamado de erro tipo 2, que eu prefiro chamar de erro de detecção.

O tamanho amostral é a chave para sermos capazes de detectar situações em que H_o é falso, e de evitar o erro tipo-2. Quanto mais informação você tem, menos variabilidade de resultados (veja o Capítulo 8) se terá, e mais capacidade você terá para focar e detectar os problemas existentes em um argumento.

A capacidade de detectar quando o H_o é realmente falso chama-se poder de um teste. Poder é um assunto bastante complexo, mas o importante é você saber que quanto maior o tamanho amostral, mais poderoso é o teste. Um teste poderoso tem poucas chances de cometer um erro tipo-2.

Considere quaisquer resultados estatísticos significativos como sendo um grão de areia, ainda que o estudo tenha sido muito bem-feito. Seja qual for a decisão tomada, essa decisão poderia estar errada. Se o estudo é estabelecido da maneira correta, no entanto, (veja o Capítulo 16, para pesquisas de opinião e 17 para experimentos) as chances de que isso ocorra são um pouco menores.

Os estatísticos recomendam duas medidas de prevenção para minimizar as chances de cometer os erro do tipo 1 e 2:

- Estabelece uma probabilidade de corte baixa para a rejeição do H_o (como 5% ou 1%), para reduzir as chances de um alarme falso. (minimizando o erro tipo-1).

- Selecione uma amostra de tamanho grande, para garantir que quaisquer diferenças ou partidas que realmente existam não sejam negligenciadas (minimizando o erro tipo-2).

Tirando conclusões sobre as conclusões deles

Mesmo que você nunca tenha conduzido um teste de hipótese sozinho, o fato de apenas saber como eles deveriam ser pode aguçar seu senso crítico. Depois de concluído o teste, o próximo passo para os pesquisadores é publicar os resultados e oferecer comunicados de impressa para a mídia indicando suas descobertas. Esse é outro ponto que você precisa ter muito cuidado. Ao mesmo tempo em que existem muitos bons cientistas, que relatam seus achados de maneira cuidadosa, enfatizando as limitações de seus dados, existem outros, no entanto, que acabam tendo um pouco mais de liberdade em suas conclusões (se essa é sua real intenção ou não, não é o que estamos discutindo aqui).

O Teste de Hipótese: Entendendo como Funciona

Todos os testes de hipótese contêm uma série de passos e procedimentos. Esta seção lhe dará uma ideia geral de tudo o que está envolvido. Veja o Capítulo 15, para mais detalhes sobre os testes de hipótese mais utilizados, incluindo os testes que verificam uma alegação a respeito de um único parâmetro populacional, assim como os que comparam duas populações.

Revisando os passos para um teste de hipótese (uma média/ proporção, grandes amostras)

Forneço aqui um resumo dos cálculos envolvidos na execução de um teste de hipótese (as fórmulas necessárias para encontrar as estatísticas de teste para os testes de hipótese mais comuns podem ser encontradas no Capítulo 15).

1. **Estabeleça as hipóteses nulas e alternativas**
   a. A hipótese nula, H_o, diz que o parâmetro populacional é igual ao valor do argumento.
   b. Existem três hipóteses alternativas possíveis, escolha a que for mais relevante no caso em que os dados não sustentem H_o.
      i. H_a: O parâmetro populacional não é igual (≠) ao número do argumento.
      ii. H_a: o parâmetro populacional é menor do que (<) o número do argumento
      iii. H_a: o parâmetro populacional é maior do que (>) o número do argumento.
2. **Selecione uma amostra aleatória dos indivíduos da população e calcule a estatística amostral.**

   Isso lhe dará a melhor estimativa do parâmetro populacional (veja o Capítulo 4).
3. **Transforme a estatística amostral em uma estatística de teste, convertendo-a em um escore padrão (todas as fórmulas para a estatística de teste estão no Capítulo 15).**
   a. Faça a diferença entre a estatística amostral e o número da hipótese nula. Essa será a distância entre a alegação e seus resultados.
   b. Divida essa distância pelo valor do erro padrão de sua estatística (veja o Capítulo 10, para mais informações sobre o erro padrão). Esse procedimento transforma a distância em unidades padrões.
4. **Encontre o *p*-valor para sua estatística de teste**
   a. Encontre a porcentagem de chances de estar nesse valor ou além dele na mesma direção:
      i. Se H_a contiver uma alternativa de inferioridade, encontre o percentil da Tabela 8-1, no Capítulo 8, que corresponda a sua estatística de teste.

ii. Se H_a contiver uma alternativa de superioridade, encontre o percentil da Tabela 8-1 (veja o Capítulo 8) que corresponda a sua estatística de teste e, depois, faça a diferença entre 100% e esse percentil (isso lhe dá a porcentagem referente ao lado direito de sua estatística de teste).

b. Multiplique essa porcentagem por dois se (e apenas se) H_a for uma alternativa de não-igualdade.

c. Converta a porcentagem em uma probabilidade, dividindo-a por 100 ou movendo duas casas decimais para a esquerda. Esse será seu *p*-valor.

5. **Avalie seu *p*-valor e tome sua decisão**

a. P-valores menores mostram mais evidências contra H_o. Conclua que H é falso (ou seja, rejeite-o)

b. P-valores maiores mostram mais evidências a favor de H_o. Conclua que você não pode rejeitá-lo. Sua amostra sustenta o argumento.

Qual é o limite de corte entre sustentar ou não H_o? A maioria das pessoas acredita que 0,05 seja um bom limite de corte para aceitar ou rejeitar H_o; *p*-valores menores do que 0,05 colocam em dúvida a veracidade de H_o. O seu limite de corte é chamado de nível alfa (α).

Nos casos em que duas populações estão sendo comparadas, a maioria dos pesquisadores se interessa em comparar os grupos de acordo com algum parâmetro, tais como o peso médio dos homens versus o das mulheres ou a proporção de mulheres que são contra alguma questão em comparação com a proporção de homens. Nesses casos, as hipóteses são estabelecidas de maneira que você possa observar a diferença entre as médias e as proporções, e a hipótese nula é de que a diferença entre as duas populações é igual a zero (os grupos teriam a mesma média e proporção). O Capítulo 15 nos dá as fórmulas e exemplos para esses tipos de teste de hipótese, tanto para os casos de amostras grandes quando para os casos de amostras pequenas.

Trabalhando com outros testes de hipótese

Muitos tipos de testes de hipótese são realizados no mundo científico. Os mais comuns estão no Capítulo 15 (juntamente com suas fórmulas fáceis de usar, explicações passo a passo e exemplos). Portanto, existem muitos tipos de testes e seus resultados chegam diariamente até você – muitos por meio de flashes de notícias, de comunicados de impressa, de telejornais da noite e da internet.

Embora os testes de hipótese usados pelos pesquisadores possam ser bastante variados, as ideias principais (como o *p*-valor e como interpretar os resultados) são as mesmas.

O elemento mais importante em todos os testes de hipótese é o *p*-valor. Todos os *p*-valores têm a mesma interpretação, independente do teste realizado. Sendo assim, onde quer que você veja um *p*-valor, você saberá que um *p*-valor pequeno significa que o pesquisador encontrou um resultado "estatisticamente significativo", que resulta na rejeição da hipótese nula.

Você sabe que, independente do teste de hipótese usado, qualquer conclusão a que você chegue é o resultado de um processo de coleta e análise de dados que, mesmo sendo muito bem conduzido, pode, por casualidade, apresentar dados não representativos ou uma verdade muito difícil de ser detectada e, portanto, levar a uma decisão equivocada. Mas é isso que torna a estatística algo interessante – você nunca saberá se o que está fazendo está correto, mas sempre saberá que o que está fazendo é o certo; isso faz algum sentido?

Lidando com amostras pequenas: a distribuição t

Para médias e proporções, nos casos em que o tamanho amostral é pequeno (e quando digo pequeno, quero dizer menor do que 30), você acaba tendo menos informações nas quais poderia embasar suas conclusões. Outro inconveniente é que você não poderá se valer da distribuição normal padrão (a distribuição Z) para comparar sua estatística de teste, pois o teorema do limite central não abrange esse tipo de situação (o teorema do limite central exige tamanhos amostrais grandes o bastante para que os resultados sejam mediados pela curva em forma de sino; veja o Capítulo 8, para mais informações a esse respeito). Você também já sabe que deve descartar os resultados que se baseiam em amostras muito pequenas (especialmente aquelas compostas por 1 indivíduo). Sendo assim, o que fazer com aquelas situações em que o tamanho da amostra não é pequeno o bastante para ser descartado, mas, também, não é grande o suficiente para usar a distribuição normal padrão a fim de analisar suas evidências? A saída é usar uma distribuição diferente, chamada de distribuição *t* (provavelmente, você já ouviu falar do termo *teste-t*, em se tratando de testes de hipóteses. É daí que vem esse termo).

A distribuição *t* é, basicamente, uma versão mais curta e grossa da distribuição normal padrão (distribuição Z). A ideia aqui é que você deve pagar um preço por possuir menos informações, e esse preço é uma distribuição com extremidades mais grossas. Marcar um gol (alcançado aquela marca de variação mágica de 5%, onde H_o é rejeitado), usando uma amostra menor, significará ter que ir mais longe, colocando-se à prova para conseguir evidências mais fortes do que as normalmente necessárias, caso você tivesse uma amostra maior. A Figura 14-2 faz uma comparação entre a distribuição normal padrão (distribuição Z) e a distribuição t.

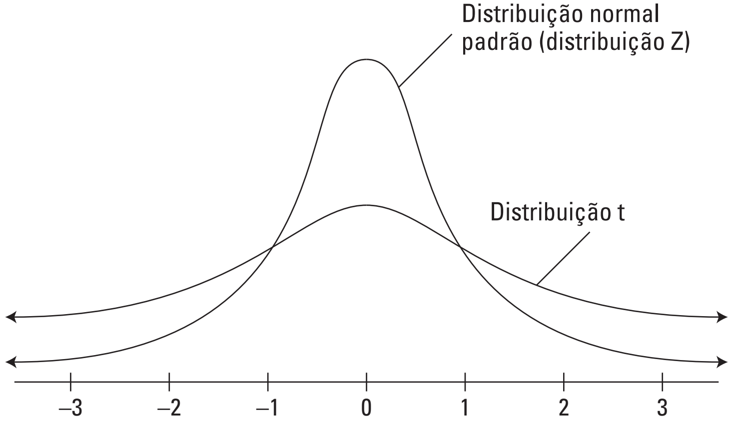


Figura 14-2: Comparação entre a distribuição normal padrão (distribuição Z) e a distribuição t.

Cada tamanho amostral possui sua própria distribuição t. É por isso que o preço a ser pago por se ter uma amostra de tamanho muito pequeno, como 5, é mais alto do que o preço por se ter uma amostra de tamanho um pouco maior, como 10 ou 20. Amostras de tamanhos menores possuem distribuições t mais curtas e mais espessas do que amostras maiores. E assim como você já deve ter imaginado, quando maior for o tamanho amostral, mais a distribuição t se parecerá com uma distribuição normal padrão (distribuição Z); e o ponto onde ambas começam a se tornar muito semelhantes é próximo a um tamanho amostral igual a 30. A Figura 14-3 mostra as diferentes aparências das distribuições t para os diferentes tamanhos amostrais e como todas se comparam à distribuição normal padrão (distribuição Z).

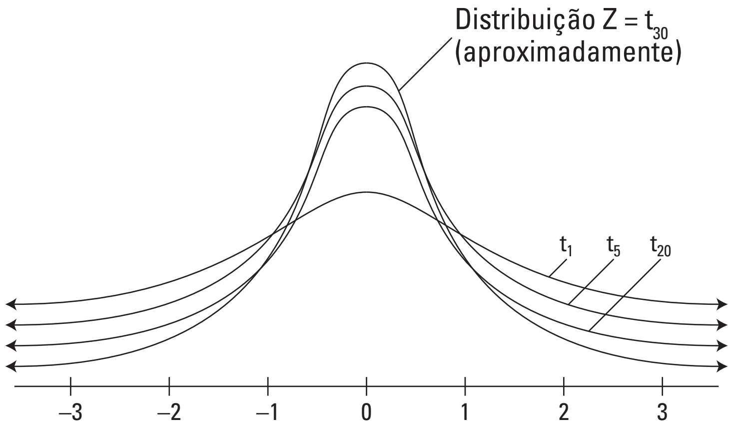


Figura 14-3: diferentes distribuições t para diferentes tamanhos amostrais.

PAPO TÉCNICO

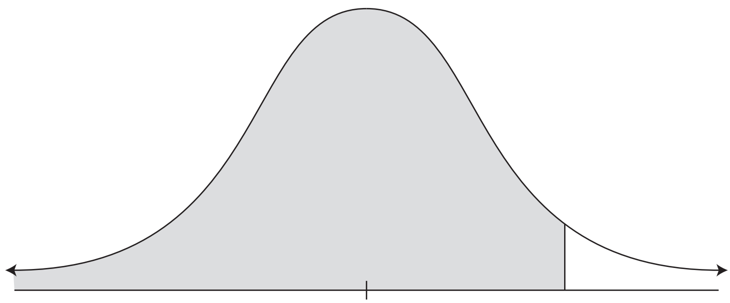
Cada distribuição *t* distingue-se por algo que os estatísticos convencionaram chamar de graus de liberdade. (o motivo pelo qual eles usam esse nome vai além do objetivo deste livro). Quando você está testando uma média populacional e o tamanho amostral é*n,* o grau de liberdade para a distribuição t correspondente é $n - 1$. Assim, por exemplo, se seu tamanho amostral é 10, você deverá utilizar uma distribuição t com 10 – 1, ou seja, 9 graus de liberdade, denotados como

t_9, ao invés de uma distribuição Z, para analisar sua estatística de teste (para qualquer teste que use a distribuição t, os graus de liberdade serão dados em fórmulas envolvendo o tamanho amostral. Veja o Capítulo 15, para mais detalhes).

A distribuição t lhe faz pagar um preço por não possuir uma amostra grande o suficiente. Mas qual é o preço? Um *p*-valor maior que o que você poderia ter conseguido por meio da distribuição normal padrão para a mesma estatística de teste. Isso por causa das extremidades mais espessas da distribuição t; uma estatística de teste distante do centro de uma distribuição Z tem pouca área ao seu redor. Mas a mesma estatística de teste próxima às extremidades mais espessas da distribuição t possui mais gordura (área) ao seu redor, e é exatamente essa área que o *p*-valor representa. Um *p*-valor maior indica menos chances de se rejeitar H_o. Menos dados geram um maior ônus à prova, assim os *p*-valores realmente funcionam da maneira que você esperaria que eles funcionassem, depois que você descobrir o que deve esperar que eles façam!

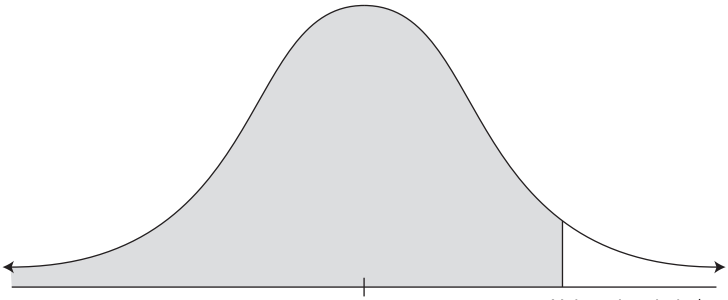
Devido ao fato de que cada tamanho amostral teria que ter sua própria distribuição t, com sua própria tabela t, para que pudéssemos encontrar os *p*-valores correspondente, os estatísticos criaram uma tabela abreviada que pode ser usada para intuir seus resultados (veja a Tabela 14-2). Os computadores também podem lhe dar os *p*-valores exatos para cada tamanho amostral.

Tabela 14-2 Distribuição t



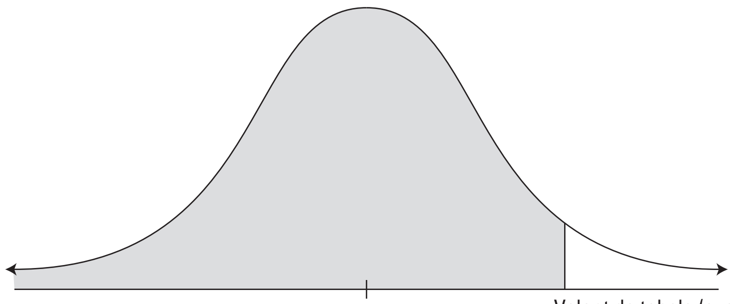
Graus de liberdade	*90° Percentil*	*95° Percentil*	*97,5° Percentil*	*98° Percentil*	*99° Percentil*
1	3, 078	6, 314	12, 706	31, 821	63, 657
2	1, 886	2, 920	4, 303	6, 965	9, 925
3	1, 638	2, 353	3, 182	4, 541	5, 841
4	1, 533	2, 132	2, 776	3, 747	4, 604
5	1, 476	2, 015	2, 571	3, 365	4, 032

Tabela 14-2 (continuação)



Graus de liberdade	*90° Percentil*	*95° Percentil*	*97,5° Percentil*	*98° Percentil*	*99° Percentil*
6	1, 440	1, 943	2, 447	3, 143	3, 707
7	1, 415	1, 895	2, 365	2, 998	3, 499
8	1, 397	1, 860	2, 306	2, 896	3, 355
9	1, 383	1, 833	2, 262	2, 821	3, 250
10	1, 372	1, 812	2, 228	2, 764	3, 169
11	1, 363	1, 796	2, 201	2, 718	3, 106
12	1, 356	1, 782	2, 179	2, 681	3, 055
13	1, 350	1, 771	2, 160	2, 650	3, 012
14	1, 345	1, 761	2, 145	2, 624	2, 977
15	1, 341	1, 753	2, 131	2, 602	2, 947
16	1, 337	1, 746	2, 120	2, 583	2, 921
17	1, 333	1, 740	2, 110	2, 567	2, 898
18	1, 330	1, 734	2, 101	2, 552	2, 878
19	1, 328	1, 729	2, 093	2, 539	2, 861
20	1, 325	1, 725	2, 986	2, 528	2, 845
21	1, 323	1, 721	2, 080	2, 518	2, 831
22	1, 321	1, 717	2, 074	2, 508	2, 819
23	1, 319	1, 714	2, 069	2, 500	2, 807
24	1, 318	1, 711	2, 064	2, 492	2, 797

Tabela 14-2 (cont.)



Graus de liberdade	*90° Percentil*	*95° Percentil*	*97,5° Percentil*	*98° Percentil*	*99° Percentil*
25	1, 316	1, 708	2, 060	2, 485	2, 787
26	1, 315	1, 706	2, 056	2, 479	2, 779
27	1, 314	1, 703	2, 052	2, 473	2, 771
28	1, 313	1, 701	2, 048	2, 467	2, 763
29	1, 311	1, 699	2, 045	2, 462	2, 756
30	1, 310	1, 697	2, 042	2, 457	2, 750
40	1, 303	1, 684	2, 021	2, 423	2, 704
60	1, 296	1, 671	2, 000	2, 390	2, 660
Valores Z	1, 282	1, 645	1, 960	2, 326	2, 576

Suponha que seu tamanho amostral seja 10, sua estatística de teste (tratada aqui como valor t) seja 2,5 e sua hipótese alternativa, H_a, seja uma alternativa de superioridade. Como o tamanho amostral é igual a 10, você deve usar a distribuição t com 10 – 1, ou seja, 9 graus de liberdade para o cálculo de seu *p*-valor. Isso significa que você deve consultar a linha na tabela t (Tabela 14-2) referente à coluna do grau de liberdade igual a 9. Sua estatística de teste (2,5) encontra-se entre dois valores: 2, 262 (o 97,5° percentil) e 2, 821 (o 98° percentil).

Mas qual é o *p*-valor? Ele se encontra em algum ponto entre 100% - 97, 5% = 2,5% = 0, 025 e 100% - 98% = 2% = 0,02 (tenha em mente que para uma alternativa de superioridade, é preciso encontrar a diferença entre 100% e o percentil). Você não sabe qual é o *p*-valor exato, mas, devido ao fato de que os valores de 2% e 2,5% são menores do que o valor para o limite de corte mais comum, 5%, H_o deve ser rejeitado.

Observe que para a hipótese alternativa de inferioridade, sua estatística de teste seria um número negativo (à esquerda de 0 na distribuição t). Nesse caso, você deve encontrar a porcentagem abaixo, ou à esquerda, de sua estatística de teste para chegar ao *p*-valor. Ainda que as estatísticas de teste negativas não apareçam na Tabela 14-2, não se preocupe! A porcentagem à esquerda (ou abaixo) de um valor t negativo é a mesma porcentagem à direita (acima) dos valores t positivos, devido à simetria. Portanto, para encontrar um *p*-valor para sua estatística de teste negativa, consulte a versão positiva da estatística de teste na Tabela 14-2, encontre o percentil correspondente e subtraia 100%.

Por exemplo, se a estatística de teste for – 2,5 com 9 graus de liberdade, procure pelo número +2,5 na Tabela 14-2, e você descobrirá que ele se encontra entre os percentis 97,5° e 98°. Subtraindo 100% desses dois valores, seu *p*-valor estará em algum ponto entre 2% e 2,5% (observe que a abordagem dos números negativos aqui é diferente da maneira como essas situações foram abordadas na Tabela 8-1, no Capítulo 8. Mas aquela tabela foi estabelecida de maneira diferente).

Se sua hipótese alternativa (H_a) for uma alternativa de não igualdade, multiplique a porcentagem obtida por dois.

Para todos os tipos de hipóteses alternativas (de superioridade, inferioridade e não igualdade), transforme a porcentagem em probabilidade, dividindo-a por 100 ou movendo duas casas decimais para a esquerda.

A tabela *t* (Tabela 14-2) não inclui todas as estatísticas possíveis de teste, sendo assim, basta escolher a que for mais próxima da sua estatística de teste. Consulte a coluna onde ela se encontra e descubra o percentil correspondente. Depois, calcule o *p*-valor.

A última linha da Tabela 14-2 mostra os valores da distribuição normal padrão (distribuição Z) correspondentes aos percentis. Observe que, conforme os graus de liberdade da distribuição *t* aumentam (conforme você se movimenta para baixo na tabela), os números chegam cada vez mais próximos aos valores da última linha da tabela. Isso confirma o que você já sabia: conforme o tamanho amostral aumenta, as distribuições *t* e Z ficam mais cada vez mais parecidas.

Capítulo 15

Testes de Hipóteses mais Utilizados

Neste Capítulo

- Desmembrando alguns dos testes de hipótese mais comuns
- Calculando suas estatísticas de teste
- Utilizando os resultados para tomar decisões embasadas

Em propagandas publicitárias ou em notícias sobre as mais recentes descobertas da medicina, você sempre irá se deparar com alegações feitas sobre uma ou mais populações. Por exemplo, "Prometemos entregar suas encomendas em dois dias ou menos" ou "Recentemente, dois estudos mostraram que uma dieta rica em fibras diminui em 20% seu risco de desenvolver um câncer de cólon". Sempre que alguém fizer um argumento ou alegação (também chamada de hipótese nula) sobre uma população (por exemplo, uma alegação de que o tempo médio gasto pelas pessoas para ir e voltar do trabalho é de 6 horas por semana ou de que a porcentagem de pessoas nos Estados Unidos que gostam de reality shows é de 30%), você poderá testá-la utilizando o que os estatísticos chamam de teste de hipótese. Você também pode usar um teste de hipótese para comparar duas populações (por exemplo, o tempo médio gasto para ir e voltar do trabalho das pessoas que trabalham no primeiro turno versus o das pessoas que trabalham no segundo turno ou, ainda, a proporção de mulheres que têm celulares, comparada a dos homens). Veja o Capítulo 14, para informações sobre as ideias principais por detrás dos testes de hipótese.

Um teste de hipótese envolve o estabelecimento de hipóteses (uma alegação e sua alternativa), a seleção de uma amostra (ou amostras), a coleta de dados, o cálculo das estatísticas relevantes e o uso dessas estatísticas para decidir se a hipótese nula é ou não verdadeira. O que, de fato, você irá fazer é comparar sua estatística amostral ao parâmetro populacional alegado e verificar a distância entre eles. Por exemplo, se o tempo médio gasto para ir e voltar do trabalho das pessoas selecionadas em sua amostra for igual a 5,2 horas, então 5,2 é sua estatística amostral. Se a hipótese nula de que o tempo médio gasto para ir e voltar do trabalho da população de todos os trabalhadores for igual a 6 horas por semana, então 6 é o parâmetro populacional alegado e, nesse caso, também é a média populacional. Quanto mais perto a estatística de teste estiver do valor alegado para o parâmetro populacional, mais você poderá acreditar que a hipótese nula é verdadeira, ainda que a grande questão seja: "Que distância seria perto o suficiente?".

Neste capítulo, eu resumo as fórmulas usadas para alguns dos testes de hipótese mais comuns, também explico os cálculos necessários e mostro alguns exemplos.

Testando Uma Média Populacional

Este teste é usado quando a variável for numérica (por exemplo, idade, renda, tempo e outros) e somente uma população ou grupo está sendo estudado (por exemplo, todos os lares americanos ou todos os estudantes universitários). Por exemplo, a Dra. Ruth diz que o tempo médio que as mães que trabalham fora passam conversando com seus filhos é de 11 minutos por dia, em média (para os papais, o argumento é de 8 minutos). A variável, tempo, é numérica, e a população é representada por todas as mães que trabalham fora.

A hipótese nula é de que a média populacional, μ, é igual a determinado valor alegado, μ_o. A equação para a hipótese nula é (H_o: $\mu_o = \mu_o$). Portanto, a hipótese nula no exemplo da Dra. Ruth é H_o: $\mu = 11$ minutos e, μ_o, nesse exemplo, é 11. Observe que μ representa o número médio de minutos por dia que todas as mães que trabalham fora passam conversando com seus filhos. A hipótese alternativa, H_a, pode ser $\mu > \mu_o$, $\mu < \mu_o$ ou $\mu \neq \mu_o$. Nesse exemplo, as três possibilidades para H_a seriam: $\mu > 11$, $\mu < 11$ ou $\mu \neq 11$ (veja o Capítulo 14, para mais informações a respeito das hipóteses alternativas). Se você suspeitar que o tempo médio que as mães que trabalham fora passam conversando com seus filhos seja mais do que 11 minutos, sua hipótese alternativa deverá ser H_a: $\mu > 11$.

A fórmula para a estatística de teste para apenas uma média populacional é $\dfrac{\bar{x} - \mu_o}{s/\sqrt{n}}$. Para calculá-la, faça os seguintes procedimentos:

1. **Calcule a média amostral, $\bar{x}$, e o desvio padrão da amostra, s. Lembre-se de que *n* representa o tamanho amostral.**

   Veja o Capítulo 4, para o cálculo da média e do desvio padrão.

2. **Encontre a diferença entre $\bar{x}$ e μ_o.**

3. **Calcule o erro padrão: $s/\sqrt{n}$. Salve sua resposta.**

4. **Divida o resultado obtido no segundo passo pelo erro padrão encontrado no terceiro passo.**

   Para o exemplo da Dra. Ruth, suponha que uma amostra aleatória, composta por 100 mães que trabalham fora de casa, passe o tempo médio de 11,5 minutos conversando com seus filhos, com um desvio padrão de 2,3 minutos. Logo, $\bar{x}$ é 11,5, $n = 100$ e $s = 2,3$.

   Faça $11,5 - 11 = +0,5$.

   Divida 2,3 pela raiz quadrada de 100 (que é 10) e obtenha o erro padrão igual a 0,23.

Divida + 0,5 por 0,23 e obtenha o resultado 2,17 (que você pode arredondar para 2,2). Esse valor é a sua estatística de teste.

Isso significa que sua média amostral é igual a 2,2 erros padrões acima da média populacional alegada. Mas esses resultados amostrais seriam incomuns, caso a alegação (H_o: $\mu = 11$ minutos) fosse verdadeira? Para decidir se sua estatística de teste sustenta a hipótese nula, calcule o *p*-valor. Para calculá-lo, procure por sua estatística de teste (nesse caso, 2,2) na distribuição normal padrão (distribuição Z) – veja a Tabela 8-1, no Capítulo 8 – e subtraia 100% do percentil encontrado, pois sua H_a é uma hipótese de superioridade. Nesse caso, a porcentagem seria 100% - 98,61% = 1,39%. Então, o *p*-valor (dividindo por 100) seria igual a 0, 0139 (veja, no Capítulo 14, outras informações sobre o cálculo do *p*-valor). O *p*-valor de 0, 0139 (1,39%) é bem menor do que 0,05 (5%), o que significa que seus resultados amostrais seriam incomuns se a alegação fosse verdadeira. Logo, rejeite a hipótese nula ($\mu = 11$ minutos) e aceita a hipótese alternativa H_a ($\mu > 11$ minutos).

Sua conclusão: de acordo com essa amostra (hipotética), a alegação da Dra. Ruth de 11 minutos é um pouco baixa; a média real é maior do que 11 minutos por dia. Veja o Capítulo 14, para mais sobre os cálculos dos testes de hipóteses e suas conclusões.

Se o tamanho amostral, *n,* aqui fosse menor do que 30, você deveria procurar o valor de sua estatística de teste na distribuição t, ao invés de usar a distribuição normal padrão (distribuição Z). Veja a Tabela 14-2, no Capítulo 14, para mais detalhes sobre esse assunto. Para informações sobre como calcular os *p*-valores para hipóteses alternativas de inferioridade e de não igualdade, também veja o Capítulo 14.

Testando Uma Proporção Populacional

Este teste é usado quando a variável for categórica (por exemplo, sexo, partido político, situação/oposição e outros) e quando apenas uma única população ou grupo estiver sendo estudado (por exemplo, todos os cidadãos norte-americanos ou todos os eleitores registrados). O teste irá verificar a proporção (p) de indivíduos na população que possuem uma determinada característica, por exemplo, a proporção de pessoas que andam com seus celulares. A hipótese nula é H_o:$p = p_o$, em que p_o é um valor alegado (ou argumento feito). Por exemplo, se a alegação é de que 20% das pessoas andam com seus celulares, p_o é 0,20. A hipótese alternativa poderá ser uma das seguintes: $p > p_o$, $p < p_o$ ou $p \neq p_o$ (veja o capítulo 14, para mais detalhes sobre hipóteses alternativas).

A fórmula para a estatística de teste para uma única proporção é $\dfrac{\hat{p} - p_o}{\sqrt{\dfrac{p_o(1 - p_o)}{n}}}$. Para calculá-la, realize os seguintes procedimentos:

1. **Calcule a proporção amostral, $\hat{p}$, dividindo o número de pessoas na amostra que tenha a característica de interesse (por exemplo, o número de pessoas que andam com seus celulares) por *n*, o tamanho amostral.**

2. **Pegue $\hat{p}$ menos p_o.**

3. **Calcule o erro padrão: $\sqrt{\frac{p_o(1-p_o)}{n}}$. Salve sua resposta.**

4. **Divida o resultado obtido no segundo passo pelo resultado obtido no terceiro passo.**

Para interpretar a estatística de teste, procure-a na distribuição normal padrão (consulte a Tabela 8-1, no Capítulo 8) e calcule o *p*-valor (veja o Capítulo 14, para mais detalhes sobre como calcular o *p*-valor).

Por exemplo, suponha que o fabricante da pasta de dente "Cáries Free" alegue que quatro em cinco dentistas recomendam Cáries Free a seus pacientes. Nesse caso, a população é representada por todos os dentistas e*p* é a proporção de todos os dentistas que recomendaram Cáries Free para seus pacientes. A hipótese nula é de que*p* é igual a "quatro em cinco", o que significa que *p* é $4 \div 5 = 0{,}80$. Porém, você suspeita que essa proporção seja, na verdade, menor do que 0,80. Então, suas hipóteses serão H_o: $p = 0{,}80$ versus H_a: $p < 0{,}80$. Imagine que 150 em 200 pacientes da amostra receberam a recomendação de usar Cáries Free.

> Para encontrar a estatística de teste, comece com $\hat{p}$ igual a $150 \div 200 = 0{,}75$. Também $p_o = 0{,}80$ e $n = 200$.
>
> Subtraia $0{,}75 - 0{,}80 = -0{,}05$.
>
> A seguir, o erro padrão, que é a raiz quadrada de $[(0{,}80 \times [1 - 0{,}80]) \div 200]$ = a raiz quadrada de $(0{,}16 \div 200)$ = a raiz quadrada de $0{,}0008 = 0{,}028$.
>
> A estatística de teste é $-0{,}05$ dividido por 0,028, que é $-0{,}05 \div 0{,}028 = -1{,}79$, que pode ser arredondado para $-1{,}8$.

Isso significa que seus resultados amostrais estão 1,8 erros padrões abaixo do valor alegado para a população.

Mas com que frequência você esperaria conseguir resultados como esses se H_o fosse verdadeiro? A porcentagem de chances de ficar em -1,8, ou além desse valor (nesse caso à esquerda) é de 3,59% (consulte -1,8 na Tabela 8-1, no Capítulo 8, e use o percentil correspondente, pois H_a é uma hipótese de inferioridade. Veja mais a esse respeito, no Capítulo 14). Agora divida o percentil encontrado por 100 e obtenha seu *p*-valor, que é 0, 0359. Já que o *p*-valor é menor do que 0,05, você possui evidências suficientes para rejeitar H_o.

Segundo sua amostra, a alegação de que quatro em cinco (80% dos) dentistas recomendam Cáries Free para seus pacientes é falsa; a porcentagem real de recomendação é menor do que isso.

A maioria dos testes de hipótese envolvendo proporções é realizada com o uso de amostras bastante grandes, dado que eles geralmente se baseiam em pesquisas de opinião. Portanto, raramente você encontrará uma situação em que uma amostra muito pequena tenha sido utilizada. Para informações sobre como calcular o *p*-valor para as alternativas de superioridade e não igualdade, veja o Capítulo 14.

Comparando Duas Médias Populacionais Separadas

Este teste é usado quando a variável for numérica (por exemplo, renda, nível de colesterol ou quilômetros rodados por litro) e quando duas populações ou grupos estiverem sendo comparados (por exemplo, homens versus mulheres, atletas versus não-atletas ou carros de passeio versus utilitários). É preciso que se coletem duas amostras aleatórias separadas, uma para cada população, a fim de reunir os dados necessários para este teste. A hipótese nula é a de que as duas médias populacionais são iguais. Em outras palavras, que a sua diferença é igual a 0. A equação para a hipótese nula é H_o: $\mu_x - \mu_y = 0$, em que μ_x representa a média da primeira população e μ_y representa a média da segunda população.

A fórmula para a estatística de teste que compara as duas médias é

$$\frac{\bar{x}-\bar{y}}{\sqrt{\frac{s_x^2}{n_1}+\frac{s_y^2}{n_2}}}$$. Para calculá-la, execute o seguinte procedimento:

1. **Calcule as médias amostrais ($\bar{x}$ e $\bar{y}$) e os desvios padrões (s_x e s_y) para cada amostra separadamente. Use n_1 e n_2 para representar os dois tamanhos amostrais (eles não precisam ser iguais).**

   Veja o Capítulo 4 para informações sobre esses cálculos.

2. **Encontre a diferença entre as duas médias amostrais, $\bar{x} - \bar{y}$.**

3. **Calcule o erro padrão, $\sqrt{\frac{s_x^2}{n_1}+\frac{s_y^2}{n_2}}$. Salve sua resposta.**

4. **Divida o resultado obtido no segundo passo pelo resultado encontrado no terceiro passo.**

Para interpretar a estatística de teste, consulte-a na distribuição normal padrão (veja a Tabela 8-1, no Capítulo 8) e calcule o *p*-valor (veja no Capítulo 14 mais sobre os cálculos do *p*-valor).

Por exemplo, suponha que você queira comparar a absorvência de duas marcas de papel toalha (vamos chamá-las de Absorve Rápido e Enxuga Tudo). Você pode fazer essa comparação, ao observar o número médio de onças (mililitros) que cada marca consegue absorver até ficar saturada. H_o diz que a diferença entre a capacidade média de absorvência das duas marca é 0 (ou seja, não existe) e H_a diz que a diferença não é 0. Em outras palavras, H_o: $\mu_x - \mu_y = 0$ versus H_a: $\mu_x - \mu_y \neq 0$. Aqui não temos a indicação de qual papel toalha pode ser o mais absorvente, assim, estabeleceremos uma alternativa de não igualdade (veja o Capítulo 14).

Imagine que você tenha selecionado uma amostra aleatória de 50 papéis toalha de cada uma das marcas e medido a absorvência de cada uma delas. Suponha que a absorvência média da marca Absorve Rápido (x) seja igual a 3 onças (88,72 ml), com um desvio padrão de 0,9 onças (26,61 ml) e a média para a marca Enxuga Tudo (y) seja de 3,5 onças (103,50 ml), com um desvio padrão de 1,2 onças (35,48 ml).

> Dadas essas informações, você tem $\bar{x} = 3$, $s_x = 0{,}9$, $\bar{y} = 3{,}5$, $s_y = 1{,}2$, $n_1 = 50$ e $n_2 = 50$.
>
> A diferença entre as médias amostrais para as duas marcas é (3 – 3,5) = - 0,5 onças (14,78 ml). (Uma diferença negativa simplesmente significa que a segunda média amostral era maior do que a primeira).
>
> O erro padrão é $\sqrt{\dfrac{0.9^2}{50} + \dfrac{1.2^2}{50}} = \sqrt{\dfrac{0.81}{50} + \dfrac{1.44}{50}} = \sqrt{0.045} = 0.2121.$
>
> Divida a diferença, -0,5, pelo erro padrão, 0, 2121, que lhe dará o valor - 2,36, que arredondado fica – 2,4. Essa é sua estatística de teste.

Para encontrar o *p*-valor, procure o valor -2,4 na distribuição normal padrão (distribuição Z) – veja a Tabela 8-1, no Capítulo 8. As chances de ficar além, nesse caso, à esquerda de -2,4 é igual ao percentil, que aqui é 0,82%. Por H_a ser uma alternativa de não igualdade, você deve multiplicar a porcentagem por dois: 2× 0,82 = 1,64%. Por último, transforme esse valor em uma probabilidade para obter o *p*-valor de 0, 0164. Para isso, basta que você o divida por 100. Esse *p*-valor é menor do que 0,05, portanto você tem evidências suficientes para descartar $H_{o.}$

Sua conclusão é de que existe uma diferença estatisticamente significativa entre os níveis de absorvência dessas duas marcas de papel toalha, segundo suas amostras. E parece que os papéis toalha da Enxuga Tudo estão na frente, pois possuem a maior média.

Sendo o estatístico sagaz que você é, não caia na conversa de comerciais que mostram apenas uma folha de um único rolo de papel toalha (ou seja, uma amostra composta por 1 indivíduo) sendo mais absorvente do que a outra marca. Também não dê credibilidade para aqueles programas de TV em que os fabricantes vão às ruas para entrevistarem duas ou três pessoas a fim de coletar informações e fazer análises comparativas. As anedotas são interessantes, mas não podem ser generalizadas. Um teste de hipótese, desde que feito corretamente, irá lhe fornecer resultados que são interessantes e que também podem ser generalizados (veja o Capítulo 14, para mais informações sobre como evitar as anedotas).

A maior parte dos testes de hipótese que compara duas médias populacionais separadamente é realizada com o uso de amostras bastante grandes, uma vez que a maioria se baseia em pesquisas de opinião. Entretanto, se as duas amostras são menores do que 30, você deve utilizar a distribuição t (com graus de liberdade igual a $n_1 - 1$ ou $n_2 - 1$, independente de qual for o menor) para encontrar o *p*-valor (veja a Tabela 14-2, no Capítulo 14, para mais informações sobre a distribuição t).

Testando uma Diferença Média (Dados Pareados)

Este teste é usado quando a variável for numérica (por exemplo, renda, nível de colesterol ou quilômetros rodados por litro), quando os indivíduos na amostras estão de alguma forma igualados (gêmeos idênticos são, geralmente, usados para esse fim) ou quando as mesmas pessoas são usadas duas vezes (por exemplo, no uso de um pré-teste e de um pós-teste). Os testes pareados são, normalmente usados quando o objetivo do teste é saber se um novo tratamento, técnica, ou método funciona melhor do que um método já existente, sem ter que se preocupar com outros fatores em relação aos indivíduos que possam influenciar os resultados. Veja o Capítulo 17, para mais detalhes.

Por exemplo, suponha que uma pesquisadora queira saber se o ensino de leitura por meio do uso de um jogo de computador trará mais resultados do que o ensino por meio do método fonético tradicional. Ela seleciona aleatoriamente 20 alunos e os divide em 10 pares de acordo com seu nível de leitura, idade, QI e outras características. Aleatoriamente, ela também seleciona um aluno de cada par para aprender através do computador, e os outros aprenderão pelo método fonético tradicional. No final do estudo, todos os alunos fazem a mesma prova de leitura. Os dados são mostrados na Tabela 15-1.

Tabela 15-1 As Notas da Prova de Leitura dos Alunos Ensinados via Jogo de Computador versus Alunos Ensinados com o Método de Fonética

Par de Alunos #	*Nota da prova para o aluno ensinado via Método Computacional*	*Nota da prova para alunos ensinados via Método de Fonética*	*Diferenças Pareadas (Notas do Computador – Notas da Fonética)*
1	85	80	+5
2	80	80	+0
3	95	88	+7
4	87	90	-3
5	78	72	+6
6	82	79	+3
7	57	50	+7
8	69	73	-4
9	73	78	-5
10	99	95	+4

Os dados estão em pares, mas você realmente está interessado em saber a diferença das notas da prova de leitura (notas do ensino via computador – notas do ensino via método fonético) de cada par, e não na nota da prova de leitura propriamente dita. Sendo assim, você deve encontrar a diferença entre as notas para cada par e, então, essas diferenças pareadas constituirão seu novo conjunto de dados. Caso os dois métodos de leitura sejam iguais, a média entre as diferenças pareadas deve ser 0. Se o método computacional for melhor, a média das diferenças pareadas deve ser positiva (pois as notas do método computacional deverão ser maiores do que as do método fonético). Portanto, você realmente tem aqui um teste de hipótese para uma média populacional, em que a hipótese nula é de que a média (das diferenças pareadas) é igual a 0 e a hipótese alternativa é de que a média (das diferenças pareadas) é >0.

A equação para a hipótese nula é H_o: $\mu_d = 0$, onde μ_d é a média das diferenças pareadas (a letra d subscrita serve apenas para lembrar que você está trabalhando com diferenças pareadas).

A fórmula para a estatística de teste para as diferenças pareadas é $\frac{\bar{d}-0}{s/\sqrt{n}}$. Para calculá-la, realize os seguintes procedimentos:

1. **Para cada par de dados, faça a diferença entre os valores do primeiro e do segundo par para encontrar a diferença pareada.**

   Entenda as diferenças como sendo seu novo conjunto de dados.

2. **Calcule a média, $\bar{d}$, e o desvio padrão, *s*, de todas as diferenças.**

   Nesse caso, designe a letra *n* para o número de diferenças pareadas que você possui.

3. **Calcule o erro padrão: $s/\sqrt{n}$. Salve sua resposta.**

4. **Divida $\bar{d}$ pelo erro padrão obtido no terceiro passo.**

   LEMBRE-SE

   Lembre-se que $\mu_d = 0$ se H_o for verdadeiro, logo, μ_d não pode ser incluída na fórmula nesse ponto.

Para o exemplo das notas de leitura, é possível utilizar o procedimento acima mencionado para verificar se o método computacional é uma melhor alternativa para o ensino de leitura.

Calcule as diferenças para cada par; você pode ver essas diferenças na coluna 4 da Tabela 15-1. Note que o sinal em cada uma das diferenças é importante; ele indica qual dos métodos se saiu melhor para aquele par em particular.

A média e o desvio padrão das diferenças devem ser calculados (coluna 4 da Tabela 15-1). (Veja, no Capítulo 4, como calcular médias e desvios padrões). A média das diferenças foi +2 e o desvio padrão é 4,64. Observe que $n = 10$ nesse exemplo.

O erro padrão é igual a 4,64 dividido pela raiz quadrada de 10 (3,16). Portanto, temos 4,64 ÷ 3,16 = 1,47 (lembre-se que aqui, *n* é o número de pares, ou seja, 10).

Para o último passo, divida a média das diferenças, +2, pelo erro padrão, 1,47 e obtenha o número +1,36, a sua estatística de teste. Isso significa que a diferença média para essa amostra é igual a 1,36 erros padrões acima de 0. Mas seria essa uma diferença significativa a ponto de ser aplicada a toda uma população?

Já que *n* é menor do que 30, procure por 1,36 na distribuição t com 10 -1 = 9 graus de liberdade (veja a Tabela 14-2, no Capítulo 14) para calcular o *p*-valor. Nesse caso, o *p*-valor é maior do que 0,05, pois 1,36 fica próximo ao valor de 1,38 na tabela e, assim, seu *p*-valor seria maior do que 0,10 (o *p*-valor correspondente a 1,38). Devido ao fato de que 1,38 encontra-se na coluna referente ao 90° percentil e, também, porque H_a é uma alternativa de superioridade, você deve subtrair 100% - 90% = 10% = 0,10. Concluímos então que não há evidências o suficiente para rejeitar H_o e, assim, o jogo de computador não pode ser considerado como um método melhor para o ensino de leitura (isso pode ter ocorrido devido à falta de evidencias extras necessárias para uma hipótese com uma amostra pequena).

Em muitos experimentos pareados, os conjuntos de dados serão pequenos em virtude dos custos e do tempo associados à execução desses tipos de estudos. Isso quer dizer que a distribuição t será frequentemente usada no lugar da distribuição normal padrão (veja a Tabela 8-1, no Capítulo 8) para a determinação do *p*-valor.

Comparando Duas Proporções Populacionais

Este teste é usado quando a variável for categórica (por exemplo, fumantes/não fumantes, partidos políticos, a favor/contra uma opinião e outras) e você estiver interessado na proporção de indivíduos que tenham determinada característica – por exemplo, a proporção de fumantes. Neste caso, duas populações ou grupos estarão sendo comparados (tais como a proporção de mulheres fumantes versus a proporção de homens fumantes). Para a realização deste teste, duas amostras aleatórias precisam ser selecionadas, uma de cada população. A hipótese nula é a de que as duas proporções populacionais são iguais; ou seja, $H_o: p_1 - p_2 = 0$, em que p_1 é a proporção da primeira população e p_2 é a proporção da segunda população.

A fórmula para a estatística de teste que compara a duas proporções é:

$$\frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

. Para calculá-la, realize os seguintes procedimentos:

1. **Calcule as proporções amostrais $\hat{p}_1$ e $\hat{p}_2$ para cada exemplo. Deixe que n_1 e n_2 representem os dois tamanhos amostrais (eles não precisam ser iguais).**

2. **Encontre a diferença entre as duas proporções amostrais, $(\hat{p}_1 - \hat{p}_2)$.**

3. **Calcule a proporção amostral geral, $\hat{p}$, que é o número total de indivíduos das duas amostras que têm a característica de interesse (por exemplo, o número total de fumantes, homens e mulheres, da amostra) e divida-o pelo número total de indivíduos das duas amostras $(n_1 + n_2)$.**

4. **Calcule o erro padrão: $\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$. Salve sua resposta**

5. **Divida o resultado obtido no segundo passo pelo resultado do quarto passo.**

Para interpretar a estatística de teste, procure-a na distribuição normal padrão (Tabela 8-1, no Capítulo 8) e calcule o *p*-valor (veja o Capítulo 14, para mais a respeito dos *p*-valores).

Por exemplo, vamos considerar aqueles anúncios que as empresas farmacêuticas colocam nas revistas. A página da frente dos anúncios mostra uma paisagem serene iluminada pelos raios de sol, com flores e pessoas sorrindo – suas vidas mudaram por causa do remédio. A empresa alega que seu medicamento consegue reduzir os sintomas de alergia, ajudam as pessoas a dormirem melhor, diminui a pressão arterial ou conserta qualquer outra coisa para a qual for indicado. A alegação pode soar muito boa para ser verdade, mas quando você olha no canto inferior da página, consegue ver aquelas letrinha miúdas que a empresa usa para justificar o como foi capaz de chegar a essas alegações (é justamente aí que as estatísticas ficam escondidas!). Em algum lugar, naquela informação escrita em letras minúsculas, você provavelmente encontrará uma tabela mostrando os efeitos colaterais do medicamento em comparação ao grupo de controle (indivíduos que recebem uma droga falsa para que se possa fazer uma comparação sem interferências com aqueles que realmente receberam o medicamento. Veja o Capítulo 17, para mais a esse respeito). Por exemplo, o Adderall, medicamento para o transtorno do déficit de atenção/hiperatividade, relatou que 26 em 374 indivíduos (7%) que tomaram o medicamento apresentaram vômito como efeito colateral, comparados a 8 em 210 (4%) indivíduos que receberam o placebo (medicamento falso). Observe que os pacientes não sabiam qual tratamento estavam recebendo. Na amostra, o número de pessoas que apresentaram vômito foi maior entre os indivíduos que receberam o medicamento verdadeiro, mas seria essa porcentagem o suficiente para dizer que uma população inteira apresentaria vômito? Podemos testar para ver.

Nesse exemplo, temos H_o: $p_1 - p_2 = 0$ versus H_o: $p_1 - p_2 > 0$, em que p_1 representa a proporção de indivíduos que vomitaram ao usar Adderall e p_2 representa a proporção de indivíduos que vomitaram usando o placebo.

Por que H_a contém um sinal de “>” e não de “<”? H_a representa a situação em que o número de pacientes que vomitaram tomando Adderall foi maior do que o número dos que usaram o placebo – isso é o tipo de informação que o FDA gostaria de saber. Mas a ordem dos grupos também é importante. Você deve estabelecê-la de modo que o grupo do Adderall venha primeiro para que assim, quando você for encontrar a diferença entre a proporção do grupo do Adderall e a do grupo de controle, seu resultado seja um número positivo se H_a for verdadeiro. Se você alterar a ordem dos grupos, o sinal poderá ser negativo.

O próximo passo é calcular a estatística de teste:

Primeiro, $\hat{p}_1 = 26 \div 374 = 0.07$ e $\hat{p}_2 = 8 \div 210 = 0{,}04$. Os tamanhos amostrais são $n_1 = 374$ e $n_2 = 210$, respectivamente.

A seguir, encontre a diferença entre as proporções amostrais, 0,07 – 0,04 = 0,03.

A proporção amostral total, $\hat{p}$, é $(26 + 8) \div (374 + 210) = 34 \div 584 = 0{,}058$.

O erro padrão é:

$$\sqrt{0.058 \times (1 - 0.058) \times \left(\frac{1}{374} + \frac{1}{210}\right)} = \sqrt{0.058 \times 0.942 \times 0.0074} = \sqrt{0.0004} = 0.02.$$

Por fim, divida o resultado obtido no segundo passo, 0,03, por 0,02, para chegar a sua estatística de teste, 0,03 ÷ 0,02 = 1,5.

O *p*-valor é a porcentagem de chances de estar no valor 1,5 ou além dele (nesse caso, à esquerda dele), que é de 100% - 93,32% = 6,68%, que, expresso forma de probabilidade, é 0, 0668 (veja a Tabela 8-1, no Capítulo 8). O *p*-valor é apenas uma sombra de 0,05, portanto, tecnicamente, você não tem evidências o suficiente para rejeitar H_o. O que significa que o efeito colateral de vomitar não foi apresentado por muito mais pessoas que utilizaram o medicamento do que pelas pessoas que utilizaram o placebo (embora esse seja o tipo de resultado que os estatísticos chamariam de marginal).

FIQUE ATENTO

Um *p*-valor muito próximo daquele valor de corte mágico igual a 0,05 é o que os estatísticos costumam chamar de resultado marginal. No exemplo anterior, o *p*-valor de 0, 0668 é, normalmente, visto como um resultado marginal. Isso significa que o resultado está bem na linha limite entre aceitar e rejeitar H_o. Essa é, no entanto, a beleza de se relatar o *p*-valor; você pode analisá-lo e, a partir dele, tomar sua própria decisão. Quanto menor for o *p*-valor, mais evidências você terá contra H_o. Mas quantas evidências podem ser consideradas o suficiente? Cada pessoa é diferente da outra. Caso você se depare com um estudo em que alguém alega ter encontrado resultados estatisticamente significativos, e esses resultados forem importantes para você, pergunte pelo *p*-valor para que assim você possa chegar a sua própria conclusão. Veja mais no Capítulo 14.

FIQUE ATENTO

A maioria dos testes de hipótese que comparam duas proporções populacionais é realizada com o uso de amostras bastante grandes, uma vez que a maioria delas provém de pesquisas de opinião, e, portanto, você provavelmente não irá se deparar com casos em que amostras pequenas tenham sido utilizadas.

Parte VII

Estudos Estatísticos: Fique por Dentro

A 5ª Onda por Rich Tennant

"Ted e eu passamos 120 horas trabalhando juntos para analisar os dados da pesquisa e veja o que descobrimos: Ted pega emprestado minhas canetas e nunca devolve, ele range a cadeira de propósito só para me irritar e, aparentemente, eu falo enquanto durmo."

Nesta Parte...

Muitas estatísticas que você vê e ouve atualmente baseiam-se em resultados de pesquisas, experimentos e estudos observacionais. Infelizmente, você não pode acreditar em tudo o que lê ou ouve.

Nesta parte, você ficará por dentro do que realmente acontece por detrás, nos bastidores, desses estudos: como eles são planejados e conduzidos; como os dados são coletados (ou pelo menos, como deveriam ser coletados) e como identificar os resultados enganosos.

Capítulo 16

Pesquisas de Opinião, Pesquisas de Opinião e Mais Pesquisas de Opinião

Neste Capítulo

- Percebendo o impacto das enquetes e pesquisas de opinião
- Percorrendo os bastidores das enquetes e pesquisas de opinião
- Interpretando os resultados das pesquisas
- Identificando resultados parciais e imprecisos

As pesquisas de opinião pública parecem estar na moda dentro dessa explosão de informações que observamos hoje. Todos querem saber o que o público pensa sobre determinados assuntos, que podem variar desde os preços de medicamentos e métodos para criar seus filhos até a taxa de aprovação do presidente e dos programas de reality show. As enquetes e as pesquisas de opinião realmente fazem parte da vida dos americanos; elas servem como veículo para a obtenção rápida de informações sobre suas opiniões e o modo como você vive, além de servirem também como meio de difundir rapidamente informações sobre importantes questões. Essas pesquisas são utilizadas para destacar tópicos polêmicos, despertar a consciência, gerar opinião política, enfatizar a importância de um assunto, educar e persuadir o público.

Os resultados das pesquisas de opinião podem ser poderosos, pois, quando muitas pessoas começam a ouvir que "tal e tal porcentagem de pessoas faz isso ou aquilo", elas aceitam esses resultados como sendo verdades e, então, passam a tomar decisões e formar suas opiniões a partir dessas informações. Na realidade, muitas pesquisas não nos fornecem informações corretas, completas, ou, até mesmo, honestas e equilibradas. Neste capítulo, discutirei o impacto dessas pesquisas e como elas são utilizadas. Também mostrarei os bastidores: como as pesquisas são projetadas e conduzidas, para que, assim, você possa saber com o que deve ter cuidado na hora que examinar os resultados desses trabalhos. Além disso, também falarei sobre como interpretar os resultados e como identificar as informações tendenciosas e imprecisas, para que você possa determinar por si só em quais resultados deve ou não acreditar.

Reconhecendo o Impacto das Pesquisas

Uma pesquisa de opinião é um instrumento que coleta dados por meio de perguntas e respostas e é utilizada para reunir informações sobre as opiniões, comportamentos, dados demográficos, estilo de vida e outras características da população de interesse. Mas, qual é a diferença entre pesquisas de opinião e enquetes? Os estatísticos não fazem uma distinção clara entre as duas, mas notei que o que as pessoas normalmente chamam de enquete é uma pesquisa de opinião curta, que contém poucas perguntas (talvez seja assim que os pesquisadores conseguem mais pessoas para entrevistar – eles a chamam de enquete, ao invés de pesquisa de opinião!). Mas para todos os fins aqui perseguidos, pesquisas de opinião e enquetes são a mesma coisa.

Diariamente, você se depara com pesquisas de opinião e seus resultados. Essas pesquisas têm até um programa de televisão próprio: o programa Family Feud é totalmente baseado em enquetes e testa a habilidade do participante de listar as respostas mais dadas pelas pessoas entrevistadas. Os participantes desse programa devem identificar corretamente as respostas dadas pelos entrevistados a questões como: "Dê o nome de um animal que pode ser visto no zoológico" ou "Fale o nome de uma pessoa famosa chamada John."

Comparadas a outros tipos de estudos, como os experimentos médicos, as enquetes são relativamente fáceis de serem conduzidas e não são tão caras. Elas fornecem resultados rápidos que, com frequência, viram manchetes interessantes nos jornais ou matérias chamativas nas revistas. As pessoas relacionam-se com as enquetes, pois sentem que os resultados desses estudos representam a opinião de pessoas como elas (embora talvez nunca tenham sido convidadas a participar de uma enquete). E muitas pessoas também gostam de saber como os outros se sentem, o que eles fazem e com o que se importam. De alguma forma, observar os resultados das enquetes faz com que as pessoas sintam-se como parte de um grupo maior. É nisso que os peritos de sondagem de opinião pública (as pessoas que realizam as enquetes) acreditam e é por isso que passam tanto tempo fazendo pesquisas e enquetes e relatando os resultados desses estudos.

Chegando à fonte

Quem conduz as enquetes atualmente? Quase qualquer pessoa que queira fazer uma pergunta. Alguns dos grupos que realizam enquetes e relatam seus resultados são os seguintes:

- Agências de notícias (por exemplo, ABC News, CNN, Reuters)
- Partidos políticos (os da situação e os da oposição)
- Organizações de pesquisa profissional (tais como o Instituto Gallup, The Harris Poll, Zogby International e outras)
- Representantes de revistas, programas de TV e rádio

- Organizações profissionais (tais como a Associação Médica Norte-Americana, que sempre conduz pesquisas com seus associados)
- Grupos de interesses especiais (tais como a National Rifle Association, entidade que luta pelos direitos dos cidadãos americanos de possuir armas de fogo)
- Pesquisadores acadêmicos (que realizam estudos sobre vários tópicos)
- Os governos (nos Estados Unidos, por exemplo, o governo conduz a American Community Survey, the Crime Victimization Survey e outras inúmeras enquetes por meio do Census Bureau)
- O público (que pode facilmente conduzir suas próprias enquetes pela Internet).

Nem todos que realizam enquetes são honestos e dignos de confiança. Sendo assim, certifique-se de checar a fonte de qualquer enquete para a qual tenha sido convidado a participar e das quais recebeu os resultados. Grupos que possuem interesses especiais nos resultados devem contratar uma organização independente para realizar a pesquisa ou devem disponibilizar ao público as cópias das perguntas feitas na pesquisa. Esses grupos também devem explicar os detalhes de como a pesquisa foi projetada e realizada, para que, assim, você possa tomar uma decisão embasada a respeito da credibilidade dos resultados.

Pesquisado o que está em alta

Os tópicos de muitas pesquisas de opinião são direcionados por eventos, assuntos e áreas de interesse da atualidade; afinal de contas, a conveniência e a relevância ao público são as duas principais qualidades de atração de qualquer pesquisa. Veja aqui alguns exemplos de assuntos que estão sendo trazidos à tona pelas pesquisas de hoje, juntamente com alguns de seus resultados:

- O ativismo das celebridades influencia a opinião política do público americano? (Mais de 90% do público americano respondeu que não, de acordo com a CBS News)
- Qual a porcentagem de americanos que já namoram alguém pela Internet? (Apenas 6% dos usuários de Internet solteiros, segundo a CBS News)
- A dor é algo que atinge muitos americanos (de acordo com a CBS News, três quarto das pessoas com menos de 50 anos frequentemente sofrem com dores ou pelo menos de vez em quando)
- Quantas pessoas usam a Internet para encontrar informações relacionadas à saúde? (Cerca de 98 milhões, segundo The Harris Poll).

- Qual é o nível atual de otimismo dos investidores? (De acordo com o Instituto Gallup, ele deveria se chamar pessimismo do investidor).
- Qual foi o pior carro do milênio? (O Yugo, sendo os ouvintes do programa Car Talk da rádio NPR).

Quando você leu os resultados das pesquisas mencionadas acima, você se pegou pensando sobre o que esses resultados significam para você, ao invés de se perguntar primeiro se os resultados são mesmo válidos? Alguns dos resultados mencionados são mais válidos e precisos do que outros. Por isso, você deveria pensar se deve acreditar nos resultados antes de aceitá-los sem questionar.

Classificação dos piores carros do milênio

Talvez você já tenha ouvido falar de um programa de rádio chamado Car Talk, que vai ao ar todos os sábados pela manhã pela National Public Radio e é apresentado pelos irmãos Click e Calck, que dão conselhos malucos para os ouvintes que ligam para reclamar de problemas estranhos que eles têm com seus carros. O site do programa sempre lança algumas enquetes sobre os mais variados assuntos relacionados a carros, tais como: “Quem tem adesivos nos para-choques e o que eles dizem?”. Uma de suas enquetes mais recentes fazia a seguinte pergunta: “Em sua opinião, qual foi o pior carro do milênio?”. Milhares de ouvintes responderam à pergunta, mas é claro que essas pessoas não representam todos os proprietários de carros. Elas apenas representam os proprietários de carro que ouvem ao programa, acessam o site na Internet e respondem às enquetes.

Mas não é por causa disso que você vai ficar sem saber os resultados da pesquisa (e eu sei que você está morrendo de curiosidade para saber). Veja os resultados na tabela a seguir. Antes de olhar as respostas dos outros, talvez você deva votar também! (No entanto, lembre-se que os resultados apenas representam as opiniões dos fãs do Car Talk que tiraram um tempo para acessar o site do programa e participar da enquete). Observe que as porcentagens não somam 100%, pois os resultados apresentados na tabela mostram apenas os dez mais votados.

Classificação	*Tipo de Carro*	*Porcentagem de Votos*
1	Yugo	33,7%
2	Chevy Vega	15,8%
3	Ford Pinto	12,6%
4	AMC Gremlin	8,5%
5	Chevy Chevette	7,0%
6	Renault LeCar	4,3%
7	Dogde Aspen/Plymouth Volare	4,1%
8	Cadilac Cimarron	4,0%
9	Renault Dauphine	3,6%
10	Volkwagem (VW) Bus	2,7%

Impactando vidas

Enquanto algumas enquetes são divertidas de se ver, outras podem ter impactos diretos em sua vida pessoal e profissional. Essas pesquisas de peso decisório precisam ser minuciosamente analisadas antes que uma decisão importante seja tomada. Enquetes desse nível podem fazer com que políticos criem novas leis, podem motivar cientistas a buscarem soluções para problemas recentes, encorajar fabricantes a inventarem novos produtos ou mudar políticas e práticas econômicas, além de influenciar o comportamento e a maneira de pensar das pessoas.
Veja a seguir alguns resultados de pesquisas recentes que podem lhe influenciar:

- **Adolescentes dirigem bêbados:** uma pesquisa recente da agência Reuters, com 1.119 adolescentes, estudantes do ensino médio em Ontario, Canadá, mostrou que 15% deles já tinham dirigido depois de consumir pelo menos dois copos de bebida alcoólica.

- **Saúde infantil está em risco:** uma pesquisa recente feita com 400 pediatras pelo Children's National Medical Center, em Washington DC, relatou que os pediatras gastam apenas, em média, de 8 a 12 minutos com cada paciente.

- **Crimes em segredo:** segundo a Crime Victimization Survey de 2001 do U.S. Bureau of Justice, apenas 49,9% dos crimes violentos são relatados à polícia. As razões que as vítimas alegam para não dar queixa dos crimes à polícia estão listadas na Tabela 16-1.

Tabela 16-1 Razões Pelas Quais Vítimas Não Relatam Crimes à Polícia

Razões para não relatar	*Porcentagem*
Consideram ser essa uma questão pessoal	19,2
O criminoso não foi bem sucedido/não conseguiu completar o crime	15,0
Relatou o ocorrido a outro oficial	14,7
Não considerou o crime tão importante	5,5
Acha que a polícia não gostaria de ser incomodada	5,3
Falta de provas	5,0
Medo de represália	4,6
Muito inconveniente/perda de tempo	3,9
Pensam que a polícia pode ser imparcial/ineficiente	2,7
A propriedade roubada não tinha número de identificação	0,5
Não sabia que um crime havia ocorrido até então	0,4
Outras razões	22,3

A razão mais frequentemente alegada pelas vítimas para não dar queixa de um crime à polícia, foi a de que a vítima considerou o crime uma questão pessoal (19,2%). Observe que quase 12% das razões estão relacionadas à percepção do ato da queixa propriamente dita (por exemplo, de que levaria muito tempo, ou de que a polícia não gostaria de ser incomodada, ou de que a polícia é imparcial e ineficiente).

- **Tratamento contra o câncer de mama é desafiado:** qual tratamento uma mulher com câncer de mama deveria escolher, a lumpectomia (em que o tumor é removido, mantendo a maior parte da mama) ou a mastectomia (em que toda a mama é retirada)? A crença mais popular é a de que a lumpectomia seria a melhor opção para a maioria das mulheres. Porém, uma pesquisa recente feita com médicas cirurgiãs revelou que se elas mesmas tivessem a doença em seu estágio inicial, 50% delas optariam pela mastectomia.

- **Celular e trânsito, uma combinação perigosa:** uma pesquisa feita recentemente com compradores e vendedores de automóveis abordou as questões relacionadas a carros e direção que mais os preocupam. Os resultados mostraram que a questão que preocupa a maior porcentagem de entrevistados (53%) foi a dos motoristas que se distraem com o uso de celulares. Esse assunto foi mais debatido do que o alto preço dos combustíveis (preocupação de 50% dos entrevistados) e a direção ofensiva (que foi a preocupação de 38% dos entrevistados).

- **Os crimes virtuais causam prejuízos a empresas:** o Computer Security Institute (CSI) realizou uma pesquisa com algumas empresas americanas para avaliar a extensão dos problemas causados pelos crimes virtuais. Noventa por cento dos entrevistados declararam possuírem brechas em seus sistemas de segurança e 80% deles reconheceram ter tido prejuízos financeiros em virtude dessas falhas. Setenta e oito por cento dos entrevistados relataram abusos dos direitos de acesso à Internet, cometidos por seus funcionários que, por exemplo, acessaram conteúdos pornográficos, fizeram download de softwares pirateados e abusaram do acesso ao e-mail pessoal.

- **Assédio sexual no trabalho:** uma pesquisa recente publicada pela Equal Employment Opportunity Commision revelou que entre 40% e 70% das mulheres e que entre 10% a 20% dos homens relataram já terem sido vítimas de assédio sexual no trabalho. Nos últimos anos, o número de queixas feitas por homens triplicou.

FIQUE ATENTO

Os exemplos mencionados apontam algumas questões muito importantes, mas é necessário que primeiro você decida se pode ou não confiar nesses resultados. Você precisa ser capaz de separar o joio do trigo. Regra número um: nunca acredite logo de cara em tudo o que você lê!

Nos bastidores: Os Segredos das Pesquisas

As pesquisas e seus resultados fazem parte de sua vida diária e fazem com que você acabe usando essas informações para tomar decisões que podem mudar sua vida (algumas decisões podem, até mesmo, transformar radicalmente sua vida). Por isso, é importante que você examine criticamente essas pesquisas. Antes de tomar uma decisão ou uma atitude baseada em uma pesquisa, você deve determinar se esses resultados são realmente confiáveis, plausíveis e críveis. Uma boa maneira de começar a desenvolver suas habilidades investigativas é conhecer os bastidores e ver como as pesquisas e enquetes são planejadas, desenvolvidas, implantadas e analisadas.

O processo de pesquisa pode ser dividido em dez passos:

1. **Estabelecimento do objetivo da pesquisa.**
2. **Definição da população alvo.**
3. **Escolha do tipo de pesquisa.**
4. **Formulação das perguntas.**
5. **Estabelecimento do momento certo para a pesquisa.**
6. **Seleção da amostra.**
7. **Coleta de dados.**
8. **Insistência, insistência, e mais insistência.**
9. **Organização e análise dos dados.**
10. **Formulação das conclusões.**

Cada passo apresenta seu próprio conjunto de assuntos a ser abordados. Mas cada procedimento é essencial para a produção de resultados honestos e precisos. Esses passos lhe ajudarão a projetar, planejar e implantar uma pesquisa, mas também podem ser usados para uma avaliação crítica da pesquisa feita por outras pessoas, caso os resultados sejam relevantes para você.

Planejando e projetando uma pesquisa

O objetivo de uma pesquisa é responder às perguntas a respeito de uma população alvo. A população alvo é todo grupo de indivíduos sobre os quais você se interessa em chegar a alguma conclusão. Na maioria das situações, pesquisar toda a população alvo é impossível, pois os pesquisadores necessitariam gastar muito tempo e dinheiro para desempenhar essa tarefa. (a pesquisa de toda uma população alvo é denominada censo). Em geral, o melhor que você pode fazer é selecionar uma amostra de indivíduos retirados da população alvo, entrevistá-los e tirar conclusões acerca da população, baseando-se nos dados da amostra.

Parece fácil, certo? Errado. Muitos problemas em potencial surgem depois que você se dá conta de que não pode pesquisar todos os indivíduos de uma população. Infelizmente, muitas pesquisas são executadas sem o tempo necessário para se pensar nessas questões, o que acaba gerando erros, resultados enganosos e conclusões equivocadas.

Estabelecendo o objetivo da pesquisa

Isso parece óbvio, mas, na realidade, muitas pesquisas foram projetadas e executadas sem nunca cumprirem seu propósito principal ou nenhum de seus objetivos; mas nem todas. Ficar perdido com as perguntas e esquecer o que você realmente está tentando descobrir são coisas que realmente podem acontecer. Ao estabelecer o propósito de sua pesquisa, seja o mais específico possível. Pense nos tipos de conclusões a que gostaria de chegar se tivesse que escrever um relatório e deixe que isso o ajude a determinar as metas de sua pesquisa.

DICA

Quanto mais específico você puder ser acerca dos propósitos de sua pesquisa, mais fácil será a formulação das perguntas que atendam seus objetivos e mais prático será o relato de seus resultados.

Definindo a população alvo

Suponha, por exemplo, que você queira conduzir uma pesquisa para determinar quantos funcionários checam seus e-mails pessoais no horário de expediente. Em um primeiro momento, poderíamos pensar que a população alvo, nessa situação, seria todos os indivíduos que checam seus e-mails no local de trabalho. Entretanto, o que você deseja é determinar o número de funcionários que checam seus e-mails pessoais durante o horário de expediente e, sendo assim, você não pode levar em conta apenas os que checam seus e-mails, pois, desta forma, seus resultados seriam desfavoráveis aos que não checam seus e-mails durante o expediente. Mas será que você deveria também incluir aqueles que nem têm acesso a um computador durante o horário de expediente? (Viu o como as pesquisas podem ser cheias de pegadinhas?).

A população alvo que nesse caso faria mais sentido seria a de todas as pessoas que usam computadores com acesso à Internet durante o expediente. Todos nesse grupo têm acesso às contas de e-mails, embora apenas alguns com esse acesso realmente façam uso dele, e dos que o usam, apenas alguns o usam para checar seus e-mails pessoais (e é isso o que você deseja descobrir – quantos funcionários realmente acessam seus e-mails pessoais durante o expediente).

LEMBRE-SE

Você deve ser muito claro ao definir sua população alvo. Essa definição é o que irá lhe ajudar a selecionar amostras adequadas, além de guiar-lhe até as conclusões finais, para que você não super generalize os resultados. Se o pesquisador não definir claramente a população alvo, ele pode gerar outros problemas para a pesquisa.

Escolhendo o tipo de pesquisa

O próximo passo na projeção de sua pesquisa é a escolha do tipo de pesquisa mais apropriado para a situação. Há várias formas de se conduzir uma pesquisa, como, por exemplo, por telefone, por correspondência, por visitas às casas dos entrevistados ou pela Internet. Porém, nem todo tipo de pesquisa é adequado para todos os tipos de situações. Suponha que você queira determinar alguns dos fatores relacionados ao analfabetismo nos Estados Unidos. Para tanto, você não poderia enviar o questionário da pesquisa por correspondência, pois as pessoas que não sabem ler não conseguiriam participar do estudo. Nesse caso, a melhor opção seria uma entrevista por telefone.

LEMBRE-SE

Escolha o tipo de pesquisa mais apropriado para a população, visando a obtenção de dados o mais confiáveis e informativos possíveis. Ao examinar os resultados de uma pesquisa, observe se o tipo de pesquisa usado foi o mais apropriado para a situação em questão.

Formulando as questões

Depois que o propósito de sua pesquisa for devidamente esboçado e você já tiver escolhido o tipo de pesquisa a ser usado, o próximo passo é formular as questões. A maneira como as perguntas são feitas pode fazer uma grande diferença na qualidade dos dados que serão coletados. Uma das fontes mais comuns de tendenciosidade nas pesquisas de opinião é a maneira como as perguntas são escritas. As questões indutivas (perguntas formuladas a favor de determinada reposta em detrimento de outra) podem influenciar muito a resposta das pessoas, fazendo com que estas não reflitam a verdadeira opinião acerca do assunto questionado. Veja a seguir duas formas que você poderia usar para formular uma pergunta para uma pesquisa sobre investimentos na educação (ambas levando a questões indutivas):

> *Você não concorda que um pequeno aumento dos impostos sobre as operações comerciais seria um investimento que valeria a pena para a melhoria da qualidade educacional de nossas crianças?*
>
> *Você não acha que nós deveríamos parar de aumentar a carga tributária dos contribuintes e parar de pedir mais aumentos dos impostos sobre as operações comerciais para investirem em nosso dispendioso sistema educacional?*

A partir da escrita de cada uma dessas questões indutivas, é possível perceber o que o entrevistador quer que você responda. Estudos demonstram que a maneira como uma pergunta é feita realmente influencia os resultados das pesquisas de opinião. A melhor maneira de formular uma pergunta é usar uma forma neutra. Por exemplo, a questão deveria ter sido escrita desta forma:

> *A secretaria de educação está propondo um aumento de 0,01% das taxas sobre operações comerciais para obter fundos para a construção de uma nova escola de ensino médio em nossa cidade. Qual é a sua opinião sobre a proposta de aumento do imposto? (Possíveis respostas: totalmente a favor, a favor, neutro, contra, totalmente contra).*

Em uma boa pesquisa de opinião, as perguntas são sempre escritas de forma neutra, de tal maneira que se evite ser parcial. A melhor maneira de avaliar a neutralidade de uma pergunta é questionar se é possível perceber a maneira como o entrevistador quer que você responda. Se sua resposta for positiva, isso significa que a pesquisa possui questões indutivas, o que pode gerar resultados enganosos.

DICA

Se os resultados da pesquisa forem importantes para você, peça ao pesquisador um cópia das perguntas usadas, para que, assim, você possa avaliar melhor a qualidade das questões.

Determinando o momento certo da pesquisa

Na pesquisa, assim como na vida, o momento certo é fundamental. Os acontecimentos atuais formam a opinião pública a todo o momento e, enquanto alguns pesquisadores tentam determinar a opinião das pessoas com relação a esses eventos, outros tiram proveito, especialmente dos eventos negativos, e os usam como plataforma política ou como base para manchetes e assuntos polêmicos. Por exemplo, suponha que a população alvo de uma pesquisa seja as pessoas que trabalham período integral. Se você conduz uma pesquisa telefônica para entrevistar pessoas que trabalham em escritórios a respeito do uso do e-mail pessoal durante o expediente e telefonar para suas casas entre 9 da manhã e 5 da tarde, você terá resultados tendenciosos, pois esse período compreende o horário de expediente da maioria das pessoas que trabalham em escritórios.

FIQUE ATENTO

Verifique a data e o horário em que a pesquisa foi realizada e veja se é possível determinar alguns eventos relevantes que tenham ocorrido naquele período e que possam ter influenciado os resultados. Também confira se a pesquisa foi conduzida no período do dia mais conveniente para a população alvo responder.

Selecionando a amostra

Depois que a pesquisa foi projetada, o próximo passo é selecionar as pessoas que participarão dela. Uma vez que, normalmente, não é possível ter o dinheiro e o tempo necessários para a realização de um censo (uma pesquisa feita com toda a população alvo), é necessário selecionar um subconjunto da população, conhecido como amostra. A maneira como a amostra é selecionada pode fazer toda a diferença em se tratando de precisão e qualidade dos resultados.

Três critérios são fundamentais para a seleção de uma boa amostra:

- **Uma boa amostra deve representar a população alvo.** Para representar a população alvo, a amostra deve ser selecionada a partir da população alvo, de toda a população alvo e de nenhum outro lugar que não a população alvo. Suponha que você queira descobrir quantas horas por dia, em média, os americanos passam assistindo televisão. Pedir que alguns estudantes de um alojamento da universidade local registrem seu hábito de assistir TV não funcionaria. Esses estudantes apenas representam uma porção da população alvo. Pedir para que as pessoas telefonem e deem sua opinião em um programa de rádio também não lhe ajudará a coletar uma amostra que represente a população alvo; os resultados apenas representarão as pessoas que estavam ouvindo o programa, que podiam telefonar naquele momento e que se sentiram interessados o suficiente para telefonar. Da mesma forma, uma pesquisa feita pela internet irá representar somente as pessoas que têm acesso à internet e que estavam conectadas ao site onde a pesquisa foi postada.

Infelizmente, muitas pessoas que realizam pesquisas de opinião não gastam o tempo e o dinheiro suficientes para selecionar uma amostra representativa de pessoas para participar do estudo. Tal comportamento acaba gerando resultados tendenciosos. Quando você se deparar com resultados de pesquisas e enquetes, descubra como a amostra foi selecionada antes de examiná-las.

- **Uma boa amostra é sempre selecionada aleatoriamente.** Uma amostra aleatória é um tipo de amostra em que todos os membros da população alvo têm chances iguais de serem selecionados. O jeito mais fácil de visualizar essa situação é imaginando um chapéu (ou uma caixa) contendo pedaços de papel, cada um com o nome de uma pessoa. Se os papéis forem bem misturados antes de serem sorteados, você obterá uma amostra aleatória da população alvo (nesse caso, a população de pessoas cujos nomes estão na caixa).

  As organizações de renome, tais como o Instituto Gallup, usam um procedimento de discagem digital para telefonar aos membros de suas amostras. É claro que esse procedimento exclui as pessoas que não possuem telefone, mas como a maioria das famílias americanas possui pelo menos uma linha telefônica, esse método mostra-se adequado e os prejuízos causados por essa exclusão são relativamente pequenos.

  Tenha cuidado com as pesquisas com amostras grandes mas não selecionadas aleatoriamente. As pesquisas feitas pelas internet são os maiores exemplos. Alguns podem dizer que 50.000 pessoas acessaram o site para responder à enquete e que isso significa que o webmaster desse site reuniu muita informação. Mas essa informação é parcial, pois representa apenas as opiniões daqueles que sabiam da pesquisa, optaram por participar e tiveram que acessar a internet. Em casos como esse, menos teria sido mais: o projetista da pesquisa deveria ter selecionado uma amostra menor, mas de maneira aleatória.

- **Uma boa amostra deve ser grande o suficiente para que os resultados sejam precisos.** Se você possui uma amostra que seja grande, represente a população alvo e que tenha sido selecionada aleatoriamente, pode contar que a informação obtida a partir dela será bastante precisa. A precisão dependente do tamanho da amostra, e quanto maior for a amostra, mais precisas serão as informações obtidas (desde que a informação seja uma boa informação). A precisão da maioria das questões de pesquisas de opinião é medida em porcentagens. Essa porcentagem é chamada de margem de erro e representa o quanto o pesquisador espera que seus resultados variem, caso fosse necessário repetir a pesquisa várias vezes usando diferentes amostras com o mesmo tamanho. Leia mais sobre o assunto no Capítulo 10.

DICA

Uma forma rápida e prática de estimar a precisão de uma pesquisa é dividir 1 pela raiz quadrada do tamanho amostral.

Por exemplo, uma pesquisa com 1.000 pessoas selecionadas aleatoriamente tem uma precisão de $\frac{1}{\sqrt{1,000}} = 0.032$, ou 3,2 pontos percentuais (observe que nos casos em que nem todos selecionados participaram da pesquisa, você deve substituir o tamanho amostral pelo número de pessoas que realmente responderam às perguntas. Veja a seção "Insistência, insistência e mais insistência", ainda neste capítulo).

Executando a pesquisa

A pesquisa já foi projetada e os participantes selecionados. Agora, você deve conhecer o processo de execução da pesquisa, outro passo muito importante, no qual muitos erros e parcialidades podem acontecer.

Coletando os dados

Durante a pesquisa propriamente dita, os participantes podem ter problemas para entender as perguntas, podem dar respostas que não estão entre as opções (no caso de questionários de múltipla escolha) ou podem, ainda, decidir dar respostas imprecisas ou falsas (como exemplo desse terceiro tipo de erro, em que os respondentes dão informações falsas, pense na dificuldade envolvida em fazer as pessoas dizerem a verdade com relação à falsificação de seus formulários de declaração de imposto de renda). Esse terceiro tipo de erro é conhecido como viés de resposta – o respondente dá uma resposta enviesada.

Alguns dos problemas em potencial com o processo de coleta de dados podem ser minimizados ou evitados com o treinamento do pessoal responsável pelas entrevistas. Com treinamento adequado, quaisquer questões que possam surgir durante uma pesquisa serão resolvidas de maneira clara e consistente, sem que erros sejam feitos durante o registro dos dados. Problemas com questões confusas e com opções incompletas para as perguntas de múltipla escolha podem ser resolvidas com a realização de um estudo piloto com poucos participantes antes da realização da pesquisa e, depois, de acordo com esse feedback, os problemas com as perguntas deverão ser corrigidos. Os entrevistadores também podem ser treinados para criar um ambiente em que os respondentes sintam-se seguros o bastante para falar a verdade; a garantia de que a privacidade será protegida também ajuda a encorajar mais pessoas a responderem.

Insistência, insistência, e mais insistência

Qualquer um que já tenha jogado fora uma pesquisa recebida por correspondência ou que tenha se recusado a responder algumas perguntas por telefone sabe que reunir pessoas para participar de uma pesquisa não é uma tarefa fácil. Se o pesquisador quiser evitar ser tendencioso, a melhor maneira é conseguir o máximo de respondentes possíveis insistindo por uma, duas ou, até mesmo, três vezes. Ofereça dinheiro, cupons, cartas respostas, chances de concorrer a prêmios e outros incentivos. Qualquer esforço é válido.

O que já motivou você a responder o questionário de uma pesquisa de opinião? Se o incentivo oferecido pelo pesquisador não persuadiu você (ou se o ato de pegar o brinde que vem junto com a pesquisa e jogá-la no lixo não faz você se sentir culpado), talvez o assunto possa atrair seu interesse. É aí que o viés entra em cena. Se apenas aqueles que se interessam pelo assunto responderem à pesquisa, isso significa que apenas as suas opiniões serão contadas, pois as pessoas que realmente não se interessaram pelo assunto não responderam à pesquisa e o seu voto de "eu não ligo" não foi contabilizado. Ou, talvez, eles até se importassem, mas não tiveram tempo o suficiente para contar a ninguém. De qualquer maneira, o voto deles não conta.

Por exemplo, suponha que 1.000 pessoas foram selecionadas para participar de uma pesquisa sobre se as regras dos parques deveriam ser alteradas, a fim de permitir a entrada de cães na coleira. Quem responderia? Muito provavelmente, os respondentes seriam as pessoas que fossem totalmente a favor ou totalmente contrárias às regras propostas. Suponha que 100 pessoas de cada lado foram as únicas a responder. Isso significaria que 800 opiniões não foram contabilizadas. Suponha que nenhuma dessas 800 pessoas realmente se importasse com o assunto. Caso fosse possível contar suas opiniões, os resultados seriam 800 ÷ 1000 = 80% "não opinaram", 100 ÷ 1000 = 10% foram a favor das novas regras e 100 ÷ 1000 = 10% foram contra as novas regras. Mas sem os votos dos 800 não-respondentes, os pesquisadores relatariam, "das pessoas que responderam à pesquisa, 50% foram a favor e 50% foram contra as novas regras". Isso nos dá a impressão de um resultado bastante diferente (e enviesado) do resultado que teríamos obtido se todas as 1000 pessoas tivessem respondido.

Mentira: O que eles sabem?

Um estudo publicado no Journal of Applied Social Psychology concluiu que quando uma mentira é contada para o próprio bem da pessoa que está ouvindo a mentira, ela se torna mais sociamente aceitável, mas quando uma mentira visa os interesses da pessoa que a contou, ela se torna menos aceitável perante a sociedade. Isso parece interessante e também parece fazer sentido, mas será que serve para todo mundo? O modo como os resultados foram relatados faz parecer que sim. Entretanto, se observarmos as pessoas que realmente participaram do estudo, começaremos a sentir que essas conclusões parecem ser, no mínimo, um pouco ambiciosas.

Os autores começaram com 1.105 mulheres selecionadas para participar do estudo. Dessas, 659 se recusaram a cooperar dizendo que não tinham tempo. 233 foram consideradas ou "jovens demais" ou "velhas demais" pelos pesquisadores e 22 foram julgadas inadequadas pelo motivo de barreira linguística. No final, apenas 180 foram entrevistadas e a média de idade desse grupo de participantes foi de 34,8 anos.

Nossa!, por onde eu começo aqui? O tamanho amostral original de 1.105 parece grande o suficiente, mas será que foi coletado aleatoriamente? Observe que a amostra inteira é constituída por mulheres, fato interessante, já que as conclusões não dizem que as mentiras são mais ou menos aceitas de acordo, apenas, com mulheres nessas situações. Depois, 659 das mulheres selecionadas se recusaram a participar (causando viés). Essa é uma porcentagem bastante grande (60%), mas dado o assunto do estudo, não deveria nos surpreender. Os pesquisadores deveriam ter minimizado o problema com a garantia de que as respostas seriam anônimas, por exemplo. Nenhuma informação é fornecida com relação à realização de acompanhamentos (o que, provavelmente, não deve ter sido feito).

Descartar 233 pessoas porque elas foram consideradas muito jovens ou muito velhas é, simplesmente, errado, a menos que sua população alvo esteja limitada a uma determinada faixa etária. Mas, se esse tivesse sido o caso, as conclusões deveriam ter sido feitas apenas considerando a faixa etária estudada. Por fim, a gota d'água: considerar 33 pessoas inadequadas (nos termos dos pesquisadores) devido a barreiras linguísticas. Eu diria, "Chamem um tradutor", pois as conclusões não se limitaram apenas àquelas que falavam inglês. Você não pode ter os participantes de uma pesquisa representando um microcosmo tão minúsculo (mulheres jovens que falam inglês) e sair por aí tirando conclusões a respeito de toda uma sociedade, baseando-se somente nos dados obtidos a partir desse microcosmo minúsculo. Começar com uma amostra composta por 1.105 mulheres e terminar com apenas 180 só poderia ter resultado em uma péssima estatística.

FIQUE ATENTO

A taxa de resposta de uma pesquisa é a razão encontrada através da divisão do número de respondentes pelo número de pessoas que foram originalmente selecionadas para participar. Os estatísticos acreditam que uma boa taxa de resposta deve ficar um pouco acima de 70%. Entretanto, muitas taxas de resposta ficam bem abaixo disso, com exceção das pesquisas feitas por organizações de renome como o Instituo Gallup. Procure a taxa de resposta da próxima vez em que for examinar um resultado. Caso ela seja muito baixa (muito menor do que 70%), os resultados podem estar enviesados e devem ser ignorados. Não se deixe enganar por pesquisas que alegam ter um grande número de respondentes, mas que, na verdade, possuem uma taxa de resposta bastante baixa, pois, nestes casos, muitas pessoas possam ter participado, mas muitas mais foram selecionadas e não responderam.

Observe que muitas fórmulas estatísticas (incluindo as fórmulas presentes neste livro) assumem que seu tamanho amostral é igual ao número de respondentes, pois os estatísticos querem que você saiba o quanto é importante contatar as pessoas para não acabar com dados enviesados por causa da falta de respondentes. Entretanto, na realidade, os estatísticos sabem que nem sempre é possível fazer com que todos respondam, independente do quanto você tenha se esforçado. Assim, qual número devemos considerar como n nas fórmulas: o tamanho amostral pretendido (o número de pessoas contatadas) ou tamanho real da amostra (o número de pessoas que realmente responderam à pesquisa)? Use sempre o número de pessoas que responderam. Observe, entretanto, que para qualquer pesquisa com uma taxa de resposta baixa, os resultados não deveriam ser relatados, pois, provavelmente, estariam enviesados. É por isso que insistir é realmente tão importante (mas será que as pessoas costumam seguir esse conselho quando relatam seus resultados? Geralmente não).

LEMBRE-SE

Em se tratando da qualidade dos dados, a seleção de uma amostra inicialmente menor com contatos mais insistentes é muito melhor do que a seleção de um grande grupo de respondentes em potencial que acaba apresentando uma baixa taxa de resposta.

Interpretando resultados; detectando problemas

O propósito de uma pesquisa é o de reunir informações a respeito de uma população alvo; tais informações podem incluir opiniões, informação demográfica, tipos de comportamentos e estilos de vida. Se a pesquisa tiver sido projetada e conduzida de maneira correta e precisa, com seus objetivos sempre em mente, os dados podem fornecer boas informações com relação ao que está se passando com a população alvo (dentro da margem de erro estabelecida). Os próximos passos são organizar os dados, a fim de se ter uma ideia melhor do que está acontecendo; analisá-los de modo a encontrar ligações, diferenças ou outros tipos de relações de interesse e, depois, esboçar conclusões baseados nos resultados.

Organizando e analisando

Depois que a pesquisa estiver pronta, o próximo passo é organizar e analisar os dados (em outras palavras, calcular alguns números e fazer alguns gráficos). Há muitas maneiras diferentes de exibir os dados e sintetizar as estatísticas, dependendo do tipo de informação coletada (dados numéricos, tais como renda, possuem características diferentes e, geralmente, são representadas de forma diferente dos dados categorizados, como o sexo). Para mais informações sobre a organização e a sinterização dos dados, veja os Capítulo 4 e 5, respectivamente. Dependendo das questões da pesquisa, diferentes tipos de análises podem ser empregadas, incluindo a obtenção de estimativas para a população, o teste de hipóteses ou a observação de relações, apenas para citar algumas. Veja os Capítulos 13, 15 e 18, para mais detalhes sobre cada uma delas, respectivamente.

Preste atenção nos gráficos e nas estatísticas enganosas. Nem todas as pesquisas são conduzidas de maneira honesta e correta. Veja o Capítulo 2 para mais informações acerca de como as estatísticas podem dar errado.

Anonimato versus confidencialidade

Se você tivesse que realizar uma pesquisa para determinar o número de pessoas que utilizam seus e-mails pessoais durante o expediente, a taxa de resposta provavelmente seria uma questão a ser considerada, pois muitas pessoas se recusariam a falar sobre o assunto, ou, talvez não falassem a verdade. Você poderia encorajá-los a responder, garantindo-lhes sua privacidade durante e depois da pesquisa.

Ao relatar os resultados de uma pesquisa, você geralmente não relaciona a informação coletada ao nome dos respondentes, pois, se o fizer, irá violar a privacidade dos respondentes. Você provavelmente já ouviu os termos "anônimo" e "confidencial" antes, mas o que talvez você ainda não saiba é que essas duas palavras são completamente diferentes, em se tratando de assuntos relacionados à privacidade. Os resultados confidenciais significam que eu posso relacionar sua informação a seu nome, mas há a promessa de que eu não farei isso. Os resultados anônimos significam que eu não tenho como relacionar sua informação a seu nome, mesmo se eu quisesse.

Quando você for convidado a participar de uma pesquisa, assegure-se sobre o que os pesquisadores realmente pretendem fazer com suas respostas e se seu nome será ou não relacionado à pesquisa (as boas pesquisas sempre deixam isso muito claro). Depois escolha se deseja ou não participar.

Esboçando conclusões

As conclusões são a melhor parte de qualquer pesquisa. Afinal, elas são o que motiva os pesquisadores a fazer todo o trabalho anterior. Se a pesquisa for projetada e conduzida adequadamente, a amostra terá sido selecionada com cuidado – e os dados estarão corretamente organizados e sintetizados – e os resultados representarão a realidade da população alvo de maneira justa e precisa. Mas é claro que nem todas as pesquisas são feitas da maneira correta e, ainda que elas fossem, os pesquisadores poderiam interpretar mal ou superestimar os resultados, de maneira que acabariam dizendo mais do que realmente deveriam dizer. Você conhece o ditado: "Ver é acreditar"? Alguns pesquisadores são os responsáveis por este ditado: "Acreditar é ver". Ou seja, eles alegam ver o que eles querem acreditar sobre os resultados. Mais uma razão para você saber onde fica o limite entre as conclusões racionais e os resultados enganosos, e identificar quando esse limite é ultrapassado.

Eis alguns erros mais comuns cometidos no momento de esboçar as conclusões:

- Fazer projeções para uma população maior do que a representada pelo estudo.
- Alegar a existência de uma diferença entre dois grupos que, na verdade, não existe.
- Dizer que "esses resultados não são científicos, mas..." e continuar a apresentá-los como realmente sendo científicos.

Excite ou Yahoo!

Em 1998, um comunicado de impressa emitido pela ferramenta de busca Excite declarava que eles tinham sido eleitos o melhor Website por uma pesquisa publicada pelo USA Today e conduzida pelo Intelliquest. A pesquisa baseou-se em 300 usuários da web selecionados a partir de um grupo de 30.000 funcionários que trabalhavam para a Intelliquest (observe que essa não é uma amostra aleatória dos usuários da Web!). As conclusões afirmavam que Excite ganhou a categoria de preferência dos consumidores, fazendo com este seja o melhor site da internet, derrotando o Yahoo! E outros concorrentes.

A Excite alegou que era melhor do que o Yahoo!, baseando-se nessa pesquisa. Os resultados reais, entretanto, contam outra história. A nota média da qualidade geral, em uma escala de 0% a 100%, foi de 89% para o Excite e 87% para o Yahoo!. A nota da Excite realmente foi muito boa e ligeiramente maior do que a obtida pelo Yahoo!; entretanto, a diferença entre os resultados para as duas empresas está, na verdade, bem na margem de erro da pesquisa, que é de 3,5% para mais ou para menos. Ou seja, Excite e Yahoo! estão estatisticamente empatadas e, portanto, dizer qual empresa está em primeiro lugar não é possível nesse caso (veja os Capítulos 9 e 20, para mais detalhes acerca da variação amostral e da margem de erro).

Para evitar alguns erros comuns cometidos ao esboçar as conclusões, faça o seguinte:

1. **Verifique se a amostra foi selecionada de maneira apropriada e que as conclusões não vão além da população apresentada pela amostra.**

2. **Se for possível, procure pelo termo de responsabilidade da pesquisa antes de ler os resultados.**

   Dessa forma, você será menos influenciado pelos resultados que estiver lendo, caso estes realmente não se baseiem em uma pesquisa científica. Agora que você já sabe o que uma pesquisa científica (termo utilizado pela mídia para as pesquisas precisas e não enviesadas) realmente envolve, pode usar esses critérios para julgar se os resultados são ou não confiáveis.

3. **Esteja atento a conclusões estatisticamente incorretas.**

   Se alguém relatar uma diferença entre dois grupos usando os resultados de uma pesquisa, certifique-se de que a diferença seja maior do que a margem de erro registrada. Caso a diferença esteja dentro da margem de erro, você deve esperar que os resultados variem de acordo com aquele valor apenas devido à casualidade, e a chamada "diferença" pode, na verdade, não existir para toda a população (veja o Capítulo 14, para mais sobre o assunto).

4. **Afaste-se de qualquer um que disser, "Esses resultados não são científicos, mas..."**

Aqui está a verdade sobre as pesquisas: conheça as limitações de qualquer pesquisa e fique atento a quaisquer informações que venham de pesquisas nas quais essas limitações não são respeitadas. Uma pesquisa ruim é barata e fácil de ser executada, mas tudo tem seu preço. Antes de analisar os resultados de qualquer pesquisa, investigue o modo como ela foi projetada e conduzida, para que, assim, você possa julgar a qualidade dos resultados.

Capítulo 17

Experimentos: Avanços Médicos ou Resultados Ilusórios

Neste Capítulo

- Considerando as limitações dos estudos observacionais
- Descobrindo como os experimentos funcionam
- Observando os resultados enganosos

Os avanços médicos aparecem e somem muito rapidamente na atual era da informação. Um dia você ouve sobre um novo tratamento promissor para uma doença e depois descobre que o medicamento não atingiu as expectativas durante o último estágio de testes. As indústrias farmacêuticas bombardeiam os telespectadores com comerciais de remédio, fazendo com que milhões de pessoas recorram a seus médicos para pedir pela inovadora cura para seus males, muitas vezes sem mesmo saber para que o medicamento é indicado. Qualquer um pode usar a Internet para pesquisar detalhes sobre tipos de mal-estar, doenças ou sintomas e conseguir toneladas de informações e conselhos. Mas serão essas informações realmente dignas de confiança? E como decidir quais opções são as melhores, caso você fique doente e precise de uma cirurgia ou tenha alguma emergência?

Neste capítulo, você conhecerá os bastidores dos experimentos, a força condutora dos estudos médicos e outras pesquisas investigativas nas quais comparações são feitas – comparações que testam, por exemplo, os melhores materiais para a construção civil, o refrigerante preferido dos adolescentes, o utilitário mais seguro durante uma colisão e outros. Você descobrirá a diferença entre os experimentos e os estudos observacionais e saberá o que este último pode fazer por você, a maneira como eles deveriam ser executados, como eles podem dar errado e como você pode identificar os resultados enganosos. Com tantas manchetes nos jornais, flashes de notícias na televisão e conselhos vindo até você por todas as direções, é necessário que você utilize todas as suas habilidades críticas para avaliar as informações, algumas vezes conflitantes, com as quais você pode frequentemente se deparar.

Determinando o que Diferencia os Experimentos

Embora existam muitos tipos diferentes de estudos, você pode reduzi-los a, basicamente, dois tipos: experimentos e estudos observacionais. Esta seção examina o que, exatamente, diferencia os experimentos de outros estudos.

Um estudo observacional é o que seu próprio nome diz: um estudo em que o pesquisador simplesmente observa os indivíduos e registra informações. Nenhuma invenção ocorre, nenhuma mudança é introduzida e nenhuma restrição ou controle é imposto. Um experimento é um estudo que, ao invés de observar os indivíduos em seu estado natural, aplica-lhes tratamentos e situações controladas e registra os resultados.

Examinando os experimentos

O objetivo básico de um experimento é descobrir se um determinado tratamento provoca uma mudança da resposta (a palavra de ordem aqui é "causa"). A maneira utilizada pelo experimento para atingir seu objetivo é a criação de um ambiente muito controlado – tão controlado que o pesquisador pode detectar se certo fator ou se a combinação de fatores causa a mudança na variável de reposta e, se esse for o caso, também se pode detectar a dimensão em que tal fator influencia a resposta.

Por exemplo, a fim de ganhar a aprovação do governo para um medicamento proposto, os pesquisadores das indústrias farmacêuticas estabelecem experimentos para determinar se aquele medicamento auxilia a controlar a pressão arterial, qual o nível de dosagem mais apropriado para cada população diferente de pacientes, quais os efeitos colaterais (se houver) e qual é a dimensão em que tais efeitos podem ocorrem em cada população.

Observando os estudos observacionais

Em certas situações, os estudos observacionais são um ótimo caminho a seguir. Os estudos observacionais mais comuns são as pesquisas de opinião e as enquetes (veja o Capítulo 16). Quando o objetivo for apenas descobrir o que as pessoas pensam e coletar algumas informações demográficas (tais como sexo, idade, renda e outras), as pesquisas de opinião e as enquetes são imbatíveis, desde que projetadas e conduzidas corretamente.

Em outras situações, especialmente naquelas em que se procura por uma relação de causa e efeito (discutida em detalhes no Capítulo 18), os estudos observacionais não são os mais indicados. Por exemplo, suponha que você tenha tomado alguns comprimidos de vitamina C na semana passada, será que foi isso que lhe ajudou a evitar o resfriado que estava rondando o escritório? Talvez o fato de ter dormido um pouco mais recentemente ou de ter lavado melhor as mãos também tenham lhe ajudado a evitar o resfriado. Ou ainda, você apenas teve sorte. Com tantas variáveis possíveis, como dizer qual realmente foi a responsável por você não ter ficado resfriado?

Ao examinar os resultados de quaisquer estudos, determine primeiro o propósito desse estudo e, depois, se o tipo de estudo usado é adequado a esse propósito. Por exemplo, se um estudo observacional foi usado no lugar de um experimento, a fim de estabelecer uma relação de causa-efeito (veja o Capítulo 18), quaisquer conclusões a que se cheguem deveriam ser minuciosamente examinadas.

Respeitando questões éticas

O problema dos experimentos é que alguns projetos experimentais nem sempre são éticos. É por isso que tantas evidências são necessárias para mostrar que o fumo causa câncer de pulmão e o porquê das indústrias de cigarro apenas recentemente começaram a pagar indenizações às vítimas. Você não pode forçar os participantes de uma pesquisa a fumarem para que se possa observar o que lhes vai acontecer. A única coisa possível é observar os indivíduos com câncer de pulmão e fazer o caminho inverso para ver que fatores (variáveis a serem estudadas) possam ter contribuído para a doença. Mas como não é possível controlar os vários fatores pelos quais você se interessa – ou quaisquer outras variáveis para a questão estudada –, escolher qualquer causa se torna difícil por meio dos estudos observacionais.

Embora as causas do câncer e de outras doenças não possam ser determinadas eticamente por meio de experimentos com humanos, os tratamentos para o câncer podem ser (e são) testados por meio de experimentos. Os estudos médicos que envolvem experimentos são denominados ensaios clínicos. Visite o site: `www.clinicaltrials.gov`; para mais informações.

As pesquisas de opinião, as enquetes e outros estudos observacionais são indicados para os casos em que você quer saber as opiniões das pessoas, examinar seus estilos de vida sem nenhuma intervenção ou examinar algumas variáveis demográficas. Caso você queira determinar a causa de certos resultados ou comportamentos (isso é, a razão pela qual algo aconteceu), um experimento é muito mais indicado. Nos casos em que um experimento não é possível (por causa de questões éticas, de orçamento ou de viabilidade), um grande corpo de estudos observacionais que examinem os mais diferentes fatores e cheguem a conclusões semelhantes seria a sua segunda melhor opção (veja o Capítulo 18, para mais detalhes sobre as relações de causa-efeito).

Projetando um Bom Experimento

A maneira como um experimento é projetado pode significar a diferença entre os bons e os péssimos resultados. Devido ao fato de que a maioria dos pesquisadores irá escrever os comunicados de impressa mais atraentes possíveis a respeito de seus experimentos, você precisa ser capaz de ver além da propaganda, de modo a determinar se pode ou não confiar no que está sendo falado.

Para decidir se um experimento é confiável ou não, verifique se ele segue estes passos:

1. **Inclui uma amostra grande o suficiente para garantir a precisão dos resultados.**
2. **Escolhe os indivíduos que melhor representam a população alvo.**
3. **Seleciona aleatoriamente os indivíduos que farão parte do grupo de tratamento e do grupo de controle.**
4. **Evita as possíveis variáveis de confusão.**
5. **Utiliza o método duplo-cego para evitar a parcialidade.**
6. **Coleta bons dados.**
7. **Contém a análise adequada dos dados.**
8. **Não esboça conclusões que vão além do escopo e das limitações do estudo.**

Nas seções a seguir, cada um desses critérios é explicado de maneira mais detalhada e ilustrado com o uso de vários exemplos.

Selecionando o tamanho amostral

O tamanho de uma amostra influencia muito a precisão dos resultados. Quanto maior for o tamanho amostral, mais preciso serão os resultados e mais poderosos serão os testes estatísticos (em se tratando da capacidade de detectar os resultados reais quando eles existem). Veja os Capítulos 10 e 14, para mais detalhes.

Entendendo que amostras pequenas não chegam a grandes conclusões

Você poderá se surpreender ao saber o número de pesquisas que foram conduzidas com base em amostras muito pequenas. Essa é uma questão de grande preocupação entre os estatísticos, que sabem que para se detectar a maioria das diferenças entre grupos é necessário o uso de amostras grandes (com pelo menos 30 indivíduos; veja o Capítulo 10). Quando grandes conclusões são feitas a partir de amostras muito pequenas, ou os pesquisadores não utilizaram o teste de hipótese correto durante a análise de seus dados (eles deveriam ter usado a distribuição t ao invés da distribuição Z; veja o Capítulo 14) ou as diferenças existentes eram tão grande que mesmo uma amostra pequena foi suficiente para detectá-las. No entanto, esta última opção não é tão frequente.

Tome cuidado com conclusões que chegam a resultados significativos a partir do uso de amostras pequenas (especialmente as amostras com valores muito menores do que 30). Caso os resultados sejam importantes para você, peça uma cópia do relatório da pesquisa e veja que tipo de análise de dados foi empregado. Também veja se os indivíduos que compunham a amostra realmente representavam a população sobre a qual as conclusões foram feitas.

Verificando sua definição de "tamanho amostral"

Ao questionar sobre o tamanho amostral, seja específico sobre o que você quer dizer por tamanho amostral. Por exemplo, você pode perguntar quantos indivíduos foram selecionados para participar e também pode perguntar pelo número de pessoas que realmente completaram o experimento; esse dois números podem ser muito diferentes. Certifique-se de que os pesquisadores possam explicar as situações em que os indivíduos estudados decidiram abandonar o experimento ou foram considerados inaptos (por alguma razão) a terminá-lo.

Por exemplo, um artigo publicado no New York Times e intitulado "Maconha tem Efeitos Analgésicos no Tratamento do Câncer" dizia, em seu primeiro parágrafo, que a maconha era muito mais eficiente do que qualquer outro medicamento para a diminuição dos efeitos colaterais causados pela quimioterapia. Quando entramos nos detalhes, descobrimos que os resultados se baseavam em apenas 29 pacientes (15 recebendo tratamento e 14 recebendo placebo). Para aumentar ainda mais a confusão, descobrimos que apenas 12 dos 15 pacientes do grupo de tratamento realmente completaram o estudo; portanto, o que aconteceu com os outros três indivíduos?

Às vezes, alguns pesquisadores esboçam conclusões baseadas apenas nos indivíduos que completaram o estudo. Isso pode levar a erros, uma vez que os dados não incluem a informação sobre aqueles que abandonaram o experimento (e o porquê), além de gerar dados enviesados. Para uma conversa sobre o tamanho amostral necessário para a obtenção de certo nível de precisão, veja o Capítulo 12.

A precisão não é a única questão a ser considerada quando se trata de conseguir "bons" dados. Você ainda precisa se preocupar em eliminar o viés por meio da seleção aleatória (veja o Capítulo 3, para mais sobre como as amostras aleatórias são coletadas).

Escolhendo os indivíduos

O primeiro passo na projeção de um experimento é a seleção da amostra de participantes, chamados de indivíduos de pesquisa. Embora os pesquisadores desejem que seus indivíduos sejam selecionados de maneira aleatória, a partir de suas respectivas populações, na maioria dos casos, isso é, simplesmente, inviável. Por exemplo, suponha que um grupo de pesquisadores oftalmologistas queira testar uma nova cirurgia a laser em pessoas com hipermetropia e precisam de uma amostra aleatória de indivíduos. Para isso, eles selecionam aleatoriamente alguns médicos oftalmologistas ao redor do país e selecionam aleatoriamente alguns pacientes desses médicos. Então, ligam para cada uma das pessoas selecionadas e dizem: "Estamos experimentando uma nova cirurgia a laser para o tratamento de hipermetropia, e você foi selecionado para participar de nosso estudo. Quando podemos marcar a cirurgia?"

Algo me diz que essa abordagem não vai funcionar com muitas pessoas (embora algumas realmente aceitem a oportunidade, especialmente se não tiverem que pagar nada pela cirurgia). A questão aqui é que conseguir uma amostra realmente aleatória de pessoas para participar em um experimento é, geralmente, mais difícil do que conseguir uma amostra aleatória de pessoas para participar de uma enquete.

O voluntariado pode causar efeitos colaterais

Para encontrar indivíduos para seus experimentos, os cientistas geralmente publicam anúncios procurando voluntários e lhes oferecem incentivos, tais como dinheiro, tratamentos gratuitos e acompanhamento médico pela sua participação. A pesquisa médica com seres humanos é complicada e difícil, mas necessária para que se saiba realmente se um tratamento funciona, como ele funciona, qual dosagem deve ser empregada e quais são seus efeitos colaterais. A fim de prescrever e receber os tratamentos corretos em doses corretas na vida real, médicos e pacientes dependem de que esse estudos representem a população em geral. A fim de recrutar esses indivíduos representativos, cientistas necessitam fazer grandes campanhas publicitárias e selecionar número suficiente de pacientes, com características diferentes, para representar um cruzamento das populações que poderão futuramente receber o tratamento.

O U.S National Institute of Health possui um site na Internet (`www.clinicaltrials.gov`) que fornece informações sobre estudos clínicos.

Dividindo os indivíduos de maneira aleatória

Depois que a amostra foi selecionada, os pesquisadores dividem os indivíduos de pesquisa em diferentes grupos. Eles são mandados para um ou mais grupos de tratamento, onde recebem vários níveis de dosagem do medicamento ou tratamento estudado, e para um grupo de controle, onde não recebem nenhum tratamento ou recebem um tratamento falso.

A importância da divisão aleatória

Suponha que um pesquisador queira determinar os efeitos causados pelos exercícios físicos sobre os batimentos cardíacos. Os indivíduos em seu grupo de tratamento correrão 8 km e seus batimentos cardíacos serão medidos antes e depois do exercício. Os indivíduos em seu grupo de controle permaneceram sentados o tempo todo em um sofá assistindo as reprises de "Os Simpsons". Em que grupo você preferiria ficar? Os apaixonados por exercícios físicos não hesitariam em candidatar-se ao grupo de tratamento. Se a ideia de correr 8 km não lhe agradar muito, provavelmente você optaria pelo mais fácil, e seria um candidato ao grupo de controle (ou talvez, você odeie "Os Simpsons" e prefira a ideia de correr 8 km a assistir um único episodio da série). Mas, qual impacto esse voluntariado seletivo teria sobre os resultados do estudo? Se apenas os apaixonados por esportes (que, provavelmente, já têm batimentos cardíacos excelentes) candidatarem-se ao grupo de tratamento, o pesquisador irá apenas observar o efeito do tratamento (a corrida de 8 km) em pessoas saudáveis e ativas. Ele não verá os efeitos da corrida de 8 km sobre os batimentos cardíacos de pessoas sedentárias. Assim, podemos notar que a divisão não aleatória dos indivíduos de pesquisa poderia causar um grande impacto sobre as conclusões tiradas a partir desse estudo.

Para evitar a parcialidade nos resultados de um experimento, os indivíduos devem ser divididos aleatoriamente e não se deve permitir que eles escolham o grupo em que querem ficar. Tenha isso em mente ao avaliar os resultados de um experimento.

Controle do efeito placebo

O falso tratamento deve levar em consideração o que se convencionou chamar de efeito placebo. O efeito placebo é a reação que as pessoas têm (ou pensam que têm) apenas pelo fato de estarem recebendo um tratamento (mesmo quando este seja um tratamento falso, como com pílulas de farinha). Veja o Capítulo 3, para mais informações.

Quando você vê uma propaganda de medicamento em uma revista, procure pelas letras miúdas no final da página. Ali você verá uma tabela que lista os efeitos colaterais relatados pelo grupo que recebeu o tratamento, comparados aos relatados pelos indivíduos do grupo de controle. Se o grupo de controle estivesse recebendo um placebo, o esperado seria que eles não relatassem nenhum efeito colateral, mas é aí que você se engana. Os grupos que recebem placebo sempre relatam efeitos colaterais em porcentagens bastante altas; isso porque suas mentes os enganam e eles experimentam o chamado efeito placebo. Se você quiser ter certeza sobre os efeitos colaterais de um tratamento, você também deve levar em consideração os efeitos colaterais do grupo controle – efeitos colaterais apenas causados pelo efeito placebo.

Em algumas situações, como quando os indivíduos de pesquisa apresentam doenças graves, oferecer um tratamento falso como uma das opções não é nada ético. Em 1997, o governo americano foi extremamente criticado por financiar um estudo sobre HIV que examinava níveis de dosagem do AZT, uma droga conhecida, até aquele momento, por diminuir em dois terços o risco de transmissão do HIV das mães para seus bebês durante a gravidez. Esse estudo em particular, do qual participaram 12.000 mulheres grávidas que tinham o vírus HIV na África, Tailândia e na República Dominicana, foi projetado de maneira cruel. Os pesquisadores deram à metade das mulheres várias doses de AZT, mas, para a outra metade, eles apenas ofereceram comprimidos de açúcar. É claro que se o governo americano soubesse que os pesquisadores dariam um placebo à metade dos indivíduos da pesquisa, ele não teria financiado esse estudo.

Mas como lidar com situações como essa? Quando questões éticas entravam o uso de tratamentos falsos, o novo tratamento deve ser comparado ao tratamento padrão já existente e conhecido como sendo um tratamento eficiente. Depois que se tenham reunidos dados suficientes para observar que um dos tratamentos realmente funciona melhor do que o outro, então, por questões éticas, o experimento deve ser interrompido e o melhor tratamento deve ser oferecido a todos os indivíduos.

Ao examinar os resultados de um experimento, verifique se os pesquisadores compararam um grupo de tratamento a um grupo de controle, a fim de ter certeza de que as experiências com o grupo de tratamento foram além das experiências vividas pelo grupo de controle. O grupo de controle pode tanto receber um tratamento falso quanto um tratamento padrão, dependendo da situação.

Controlando as variáveis de confusão

Suponha que você esteja participando de um estudo cujo objetivo seja observar os fatores que influenciam as causas de um resfriado. Se o pesquisador apenas registrar o período em que você ficou resfriado e lhe fizer perguntas sobre seus hábitos (quantas vezes ao dia você lava suas mãos, quantas horas você dorme por noite, etc.), ele estará conduzindo um estudo observacional. O problema com esse tipo de estudo é que, sem o controle dos outros fatores que puderam influenciar as causas do resfriado e sem regulamentação da ação tomada para evitá-lo, o estudo não será capaz de especificar exatamente quais ações (se é que houve uma) realmente tiveram um impacto sobre o resultado.

A grande limitação dos estudos observacionais é que eles não podem realmente mostrar a verdadeira relação de causa e efeito, em virtude do que os estatísticos chamam de variável de confusão. A variável de confusão é uma variável ou fator que não foi controlada durante o estudo, mas que pode influenciar os resultados.

Por exemplo, uma manchete de jornal dizia: "Estudo relaciona mães mais velhas a uma maior expectativa de vida". O parágrafo de introdução dizia que as mulheres que têm seu primeiro filho aos 40 anos têm muito mais chances de viver até os 100 anos, quando comparadas às mulheres que têm seus primeiros filhos antes dessa idade. Quando entramos nos detalhes do estudo (feito em 1996), descobrimos, primeiramente, que ele se baseava em 78 mulheres que moravam em subúrbios de Boston e viveram até 100 anos, comparadas a 54 mulheres que nasceram no mesmo ano (1896), mas morreram em 1969 (o ano em que os pesquisadores conseguiram encontrar os primeiros registros de óbito). Este denominado grupo controle viveu exatamente 73 anos, nada a mais e nada a menos. Das mulheres que viveram até pelo menos 100, 19% deram a luz depois dos 40 anos, enquanto apenas 5,5% das mulheres que morreram aos 73 anos tinham dado á luz depois dessa idade.

Eu realmente tenho um problema com esse tipo de conclusão. O que dizer sobre o fato de que o "grupo de controle" baseava-se apenas nas mulheres que morreram em 1969 aos 73 anos? E todas as outras mães que morreram antes de completar 73 ou com idade entre 73 e 100 anos? Talvez o grupo de controle (apresentado em um propósito tão limitado) tenha incluído mulheres que tivessem algum tipo de relação; talvez essa relação tenha sido a causa para que muitas delas morressem no mesmo ano e, talvez, essa relação também estivesse ligada à razão para que a maioria delas tivesse tido seus filhos quando mais jovens. Quem pode saber? E quanto as outras variáveis que podem influenciar tanto a idade das mães no nascimento de seus filhos quanto as expectativas de vida mais longas – variáveis como as condições financeiras, a estabilidade do casamento ou outros fatores socioeconômicos? As mulheres, nesse estudo, estavam com 33 anos durante o período da Grande Depressão; isso pode ter influenciado sua expectativa de vida e se (ou quando) deveriam ter filhos.

Como os pesquisadores lidam com as variáveis confusas? A palavra de ordem é "controle". Eles controlam quantas variáveis confusas são possíveis prever. Em experimentos envolvendo humanos, os pesquisadores têm que lutar contra muitas variáveis confusas. Por exemplo, em um estudo que tenta determinar o efeito que diferentes tipos de música e volume causam sobre o tempo que os compradores passam dentro de um supermercado (sim, eles realmente pensam em coisas desse tipo), os pesquisadores têm que antecipar o máximo de variáveis confusas possíveis, para, então, tentar controlá-las. Que outros fatores além do volume e do tipo de música podem influenciar o tempo que você gasta dentro de um mercado? Eu consigo pensar em vários fatores: sexo; idade; período do dia; se estou com meus filhos; quanto dinheiro posso gastar; o dia da semana; a limpeza do estabelecimento; a educação dos funcionários e (muito importante) o motivo para eu estar ali – estou fazendo compras para a semana ou, simplesmente, passei para comprar uma barra de chocolate?

Como os pesquisadores podem começar a controlar tantas possíveis variáveis confusas? Algumas delas, tais como o período do dia, o dia da semana e a razão para a compra, podem ser controladas durante o projeto do estudo. Porém, outros fatores (tais como a percepção da atmosfera da loja) dependem totalmente do indivíduo participando do estudo. A única forma de controle dessas variáveis confusas subjetivas é usar pares de pessoas que se combinem segundo variáveis importantes, ou, também, usar a mesma pessoa duas vezes: uma vez com o tratamento e a outra sem o tratamento. Esse tipo de experimento é chamado de projeto de pares combinados. (Veja o Capitulo 14, para mais detalhes).

Antes de acreditar em qualquer manchete médica (ou qualquer manchete relacionada a esse assunto), veja como o estudo foi realizado. Os estudos observacionais não podem controlar as variáveis confusas, assim seus resultados não possuem significado estatístico (não importa o que os estatísticos digam) como os de um experimento bem projetado. Nos casos em que não é possível realizar um experimento (afinal de contas, ninguém pode lhe obrigar a ter um filho antes ou depois dos 40 anos), certifique-se de que o estudo observacional se baseia-se em uma amostra grande o bastante para representar um corte transversal da população.

Estudo duplo-cego

Experimentos bem projetados são realizados com a metodologia do estudo duplo-cego. Duplo-cego significa que nem os indivíduos nem os pesquisadores sabem quem está recebendo o tratamento e quem está no grupo de controle. Os indivíduos de pesquisa precisam estar inconscientes do tratamento que estão recebendo para que os pesquisadores possam medir o efeito placebo. Mas por que os pesquisadores também não podem saber quem está recebendo o tratamento? Porque assim eles não tratarão os indivíduos de maneira diferente pelo fato de esperar (ou não) certas reações de cada um dos grupos. Por exemplo, se uma cientista sabe que você está no grupo de tratamento para estudar os efeitos colaterais de um remédio, ela poderia esperar que você ficasse doente e, por isso, poderia lhe dar mais atenção do que ela daria aos indivíduos do grupo de controle. Esse tipo de comportamento pode levar a dados enviesados e resultados enganosos.

Quando o pesquisador sabe quem recebeu o tratamento, mas os indivíduos não, a metodologia usada é chamada de estudo cego (no lugar de estudo duplo-cego). Os estudos cegos são melhores do que nada, mas os duplo-cegos são melhores. Caso você esteja se perguntando: em um estudo duplo-cego, existe alguém que saiba qual tratamento está sendo dado aos indivíduos? Relaxe; geralmente essa parte é realizada por um assistente de laboratório terceirizado.

Ao analisar um experimento, procure saber se ele é um estudo duplo-cego. Caso não seja, os resultados podem estar enviesados.

Coletando bons dados

O que seriam "bons" dados? Os estatísticos usam três critérios para avaliar a qualidade dos dados; cada um dos critérios está estritamente relacionado com a qualidade do instrumento de medição usado no processo de coleta de dados.

Para você saber se está ou não diante de bons dados de um experimento, procure por estas características:

- **Confiável (você obtém os mesmos resultados para várias medições subsequentes):** muitas balanças de banheiro não são confiáveis. Você sobe na balança e ela lhe mostra um número. Você não acredita nesse número, desce, volta a subir e ela lhe dá outro número (se o segundo número for menor, provavelmente você irá parar por aí mesmo: mas, se ele não for, você talvez continue repetindo a operação até ver o número que deseja).

  A questão é: dados não confiáveis provêm de instrumento de medição não confiáveis. Tais instrumentos podem ir além de balanças para abranger os instrumentos de medição mais intangíveis, como as perguntas de uma enquete, que podem nos dá resultados irreais se forem escritas de maneira ambígua (veja o Capítulo 16, para mais detalhes sobre o assunto).

  Ao examinar os resultados de um experimento, descubra como os dados foram coletados. Se a medida não for confiável, os dados poderão ser imprecisos.

- **Não enviesados (os dados não contêm erros sistemáticos que possam tanto aumentar quanto diminuir os valores reais):** os dados enviesados são aqueles que, sistematicamente, aumentam ou diminuem o resultado real. O viés pode ocorrer em quase todas as etapas de projeção e implantação de um estudo. O viés pode ser causado por um instrumento de medição ruim (como as balanças de banheiro que "sempre" aumentam uns quilinhos), como as questões de enquetes que induzem o participante a responder de certa maneira ou como os pesquisadores que sabem qual tratamento cada indivíduo recebeu e que preconceberam as expectativas.

  O viés é, provavelmente, o problema número um com relação aos dados. E o pior de tudo é que ele não pode ser quantificado (por exemplo, a margem de erro não mede o viés. Veja o Capítulo 10, para mais detalhes sobre a margem de erro). Entretanto, algumas medidas podem ser tomadas para minimizar o viés, como o discutido no Capítulo 16 e na seção "Dividindo os indivíduos de maneira aleatória", no início deste capítulo.

Tome cuidado com as várias formas que o viés pode tomar durante a projeção e implantação de quaisquer estudos, e sempre avalie um estudo ficando de olho no viés. Se um estudo contiver muita tendenciosidade, seus resultados devem ser ignorados.

- **Validade (os dados realmente medem o que deveriam medir):** verificar a validade dos dados requer que você volte alguns passos e veja o todo. É necessário que você se pergunte: "Será que esses dados realmente medem o que eles deveriam medir?". Ou os pesquisadores deveriam ter coletado dados diferentes? O uso de um instrumento de medição adequado também é importante. Por exemplo, pedir que os alunos do ensino médio relatem suas notas de matemática pode não ser uma maneira válida de medir as reais notas. Uma medida mais válida seria olhar a nota de cada aluno. Medir a prevalência da criminalidade utilizando o número de crimes também não é válido; a taxa de criminalidade (número de crimes per capita) seria a melhor opção.

Antes de aceitar os resultados de um experimento, descubra o que os dados medem e como eles foram medidos. Certifique-se de que os pesquisadores coletaram dados apropriados ao propósito do estudo.

Analisando os dados de maneira adequada

Depois que os dados são coletados, eles são colocados naquela misteriosa caixa chamada análise estatística. A escolha da análise é tão importante quanto (em se tratando da qualidade dos resultados) qualquer outro aspecto do estudo. A análise adequada deveria ser planejada com antecedência durante a fase de projeção do experimento, pois, desta forma, depois que os dados estiverem coletados, você não enfrentaria problemas durante a análise.

Aqui vai uma dica para que você selecione a análise adequada. Antes de tudo, faça a pergunta: "Depois que os dados forem analisados, eu conseguirei responder a pergunta que propus?.Se sua resposta for negativa, a análise não é a mais adequada.

Os tipos básicos de análises estatísticas incluem os intervalos de confiança (usados quando se está tentando estimar um valor populacional ou a diferença entre duas populações); os testes de hipótese (usados quando se quer testar uma alegação a respeito de uma ou duas populações, tal como a alegação de que um medicamento é mais eficaz do que outro); e as análises de correlação e regressão (usadas quando o objetivo é se e/ou como uma variável pode prever ou causar alterações em outra variável). Veja os Capítulos 13, 15 e 18, respectivamente, para mais detalhes sobre cada um desses tipos de análise.

Ao escolher o modo como você irá realizar a análise de seus dados, você deve se certificar de que seus dados e sua análise sejam compatíveis. Por exemplo, caso você deseje comparar um grupo de tratamento a um grupo de controle com relação à perda de peso obtida graças a um novo programa alimentar (versus um programa já existente), é necessário que se coletem dados sobre o peso que cada pessoa perdeu (e não apenas o peso de cada indivíduo ao final do estudo).

Esboçando conclusões apropriadas

Em minha opinião, os maiores erros que os pesquisadores costumam cometer ao esboçar as conclusões sobre seus estudos são:

- Superestimar seus resultados.
- Fazer relações e dar explicações que não são sustentadas pelas estatísticas.
- Ir além do escopo do estudo no que se refere a quem os resultados se aplicam.

Cada um desses problemas será discutido nas seções seguintes.

Superestimando os resultados

Muitas vezes, as manchetes publicadas pela mídia superestimam os reais resultados das pesquisas. Ao ler uma manchete ou escutar algo sobre um estudo, procure por mais informações, de modo a descobrir os detalhes de como o estudo foi realizado e quais são suas reais conclusões.

Com frequência, os comunicados de impressa também superestimam resultados. Por exemplo, em um comunicado recentemente emitido pelo National Institute for Drug Abuse, os pesquisadores alegavam que o uso do Êxtase diminuiu em 2002 com relação a 2001. Porém, quando olhamos os resultados estatísticos reais, descobrimos que a porcentagem de adolescentes da amostra que disseram ter usados Êxtase foi menor em 2002 do que em 2001, mas a diferença não pode ser considerada estatisticamente significativa (embora muitos outros resultados fossem considerados estatisticamente significativos). Isso significa que, embora menos adolescentes na amostra tenham usado Êxtase em 2002, a diferença não foi o suficiente para descartar as chances de variabilidade de uma amostra para outra (veja o Capítulo 14, para mais detalhes sobre a significância estatística).

Os títulos e os parágrafos de introdução em comunicados de imprensa e artigos de jornais frequentemente superestimam os resultados reais de um estudo. Os grandes resultados, as descobertas surpreendentes e os principais avanços médicos são os responsáveis atualmente hoje em dia e a mídia constantemente seleciona o que vale ou não a pena ser publicado. Como separar a verdade do exagero? A melhor coisa a fazer é ler o que está escrito em letras bem miúdas.

Levando os resultados um passo além dos dados reais

Um estudo que relaciona a idade com que você tem seus filhos a sua expectativa de vida ilustra outra questão a respeito dos resultados de pesquisas. Os resultados desse estudo observacional são mesmo capazes de mostrar que quanto mais tarde você tem um filho, maior será sua expectativa de vida? "Não", dizem os pesquisadores. A explicação para os resultados é de que ter um filho com idade mais avançada pode ser consequência de um relógio biológico mais lento, o que presumidamente resulta em um processo de envelhecimento desacelerado.

Minha pergunta para esses pesquisadores é: "Então, por que vocês não estudam isso, ao invés de apenas observar as idades?". Eu não vejo nenhum dado nesse estudo que me levaria a concluir que as mulheres que tiveram filhos depois dos 40 anos tivessem envelhecido a uma taxa mais lenta do que as outras. Portanto, do meu ponto de vista, ou os pesquisadores não deveriam ter chegado a essa conclusão ainda, ou eles deveriam ter deixado mais claro que essa visão é apenas uma teoria e ainda necessita de mais estudos. Baseada nos dados desse estudo, a teoria dos pesquisadores me parece mais uma amostra de fé (embora sendo uma mamãe de 41 anos, também espero o melhor!).

Frequentemente, em comunicados de imprensa e notícias, o pesquisador irá dar uma explicação sobre porque ele acredita que seu estudo tenha demonstrado os presentes resultados e quais seriam suas implicações para a sociedade como um todo (essas explicações podem vir em resposta às perguntas feitas por um jornalista e, depois, serem editadas para apenas mostrarem as citações mais interessantes do pesquisador). A maioria dessas explicações não passa de teorias que ainda necessitam ser testadas. Nesses casos, fique atento com as conclusões, explicações ou relações feitas pelos pesquisadores e que não se sustentam por seus estudos.

Generalizando os resultados para populações além do escopo do estudo

Você só pode tirar conclusões acerca da população representada por sua amostra. Caso tenha selecionado apenas homens, não pode tirar conclusões sobre mulheres. Se você somente selecionou pessoas jovens e saudáveis, não pode tirar conclusões a respeito de todo mundo. Mas muitos pesquisadores fazem isso. Essa prática comum pode levar a resultados enganosos. Cuidado com ela!

Eis algumas dicas para que você possa determinar se as conclusões de um pesquisador estão à altura:

- Descubra qual é a população alvo (ou seja, o grupo sobre o qual o pesquisador quer tirar conclusões).
- Descubra como a amostra foi selecionada e se ela realmente representa a população alvo (e não qualquer outra população mais restrita).
- Verifique as conclusões feitas pelos pesquisadores; certifique-se de que eles não estão tentando aplicar seus resultados a uma população mais ampla do que a realmente estudada.

Tomando Decisões Embasadas sobre os Experimentos

Só porque alguém alega ter conduzido um "experimento cientifico" ou um "estudo científico", isso não significa que esse estudo ou experimento tenha sido conduzido corretamente ou que seus resultados sejam dignos de confiança. Infelizmente, eu me deparei com muitos experimentos ruins em meus dias como consultora estatística. A pior parte de um experimento mal feito é que não há nada a ser feito a não ser ignorar seus resultados – e é exatamente isso o que você deve fazer.

Aqui vão algumas dicas de como tomar uma decisão embasada sobre se você deve ou não confiar nos resultados de um experimento, especialmente se esses resultados são muito importantes para você.

- Ao ouvir ou ver um resultado pela primeira vez, pegue um lápis e anote o máximo de informações possíveis, tais como o lugar onde você ouviu ou leu, quem fez a pesquisa e quais foram os principais resultados (eu sempre tenho lápis e papel na minha sala de TV e na minha bolsa para situações como essa).

- Continue a pesquisar suas fontes até que encontre a pessoa que fez a pesquisa original e, depois, peça a ela uma cópia do relatório ou do artigo.

- Analise o relatório e avalie o experimento de acordo com os oito passos para um bom experimento, descritos na seção "Projetando um Bom Experimento", neste capítulo (você não necessita entender tudo o que está escrito no relatório para fazer isso).

- Analise minuciosamente as conclusões feitas pelo pesquisador com relação a seus achados. Muitos pesquisadores tendem a superestimar seus resultados, tirar conclusões além das evidências estatísticas ou tentar aplicar seus resultados a uma população maior do que a estudada.

- Nunca tenha receio de fazer perguntas à mídia, aos pesquisadores e até mesmo aos seus próprios especialistas. Por exemplo, se você tiver alguma dúvida sobre um estudo na área da medicina, pergunte a seu médico. Ele ficará contente em ter um paciente tão bem informado!

Capítulo 18

Procurando Vínculos: Correlações e Associações

Neste Capítulo

- Entendendo as estatísticas de relações
- Distinção entre associações, correlações e causalidade
- Fazendo previsões baseadas em relações conhecidas
- Detectando os resultados enganosos

Todos parecem querer lhe contar as últimas relações, correlações, associações ou ligações encontradas. Muitas dessas ligações vêm da pesquisa médica, como você já pode esperar. O trabalho dos pesquisadores médicos é falar o que você pode e não pode fazer para viver uma vida mais longa e saudável.

Veja a seguir alguns dos novos lançamentos fornecidos pelo National Institutes of Health (NIH):

- As atividades sedentárias (como assistir TV) estão associadas ao aumento da obesidade e dos riscos de diabetes em mulheres.
- A expressão de raiva pode estar inversamente relacionada ao risco de ataques cardíacos e derrames (aqueles que expressam sua raiva têm um risco menor para essas complicações).
- Beber moderadamente reduz os riscos de doenças cardíacas em homens.
- O tratamento imediato ajuda a retardar o progresso do glaucoma.

Os jornalistas adoram contar esses tipos de relações para as pessoas, pois essas histórias acabam virando grandes notícias. Algumas recomendações, no entanto, parecem mudar de tempos em tempos; por exemplo, em um minuto o zinco é recomendado para evitar o resfriado e, um minuto, depois, ele já não serve mais. Muitas das relações que você vê na mídia são consideradas como sendo relações de causa e efeito, mas será que podemos confiar nessas notícias? (Por exemplo, o simples fato de uma mulher ter seu primeiro filho depois dos 40 anos pode mesmo ser a causa para que ela viva por mais tempo?). Será que você está tão céptico que, simplesmente, não acredita mais em nada?

Se você se sente um consumidor de informação confuso ao ouvir sobre as mais recentes relações e correlações médicas, fique tranquilo; este capítulo poderá lhe ajudar. Você vai descobrir o que realmente significa a correlação, a associação ou a causalidade de dois fatores, além de saber quando e como fazer previsões, baseando-se nessas relações. Você também vai ganhar habilidades para dissecar e avaliar as alegações de uma pesquisa, a fim de tomar suas próprias decisões acerca das manchetes e flashes de notícias que alertam você sobre a mais nova correlação.

Retratando as Relações: Enredos e Tabelas

Um artigo na revista Garden Gate chamou minha atenção. O título dizia: "Conte o Cricrilar dos Grilos e Descubra a Temperatura". Segundo o artigo, tudo o que você precisava fazer era encontrar um grilo, contar quantas vezes ele cricrila em 15 segundos, somar 40 e voila! Você acabou de prever a temperatura em Fahrenheit.

O National Weather Service Forecast Office até mesmo publicou seu próprio Conversor do Cricrilar dos Grilos. Basta você digitar o número de vezes que o grilo cricrila em 15 segundos e o conversor lhe mostrará a temperatura estimada em quatro unidades, incluindo Fahrenheit e Celsius.

Várias boas pesquisas dão suporte à alegação de que há uma relação entre a frequência do cricrilar dos grilos e a temperatura. Para ilustrar, tirei um subconjunto de alguns dados e o representei na Tabela 18-1. Note que cada observação é composta por duas variáveis dependentes, nesse caso, o número de vezes que o grilo cricrila em 15 segundos e a temperatura no momento (em Fahrenheit). Os estatísticos costumam chamar esse tipo de dados de duas dimensões de dados bivariados. Cada observação contém um par de dados coletado simultaneamente.

Tabela 18-1 O Cricrilar dos Grilos e a Temperatura (trecho selecionado)

Número de vezes que o grilo cricrila em 15 segundos	*Temperatura (em Fahrenheit)*
18	57
20	60
21	64
23	65
27	68
30	71
34	74
39	77

Um comunicado de impressa recentemente publicado pelo Departamento Médico da Ohio State Unversity também chamou minha atenção. O título dizia que a aspirina podia prevenir pólipos em pacientes com câncer de cólon. Como tive um parente próximo que faleceu por causa desta doença, fiquei contente em saber que os pesquisadores estavam fazendo progressos nessa área e decidi ver mais de perto. Os dados do estudo da aspirina versus pólipos colorretais estão resumidos na tabela 18-2.

Tabela 18-2 Resumo dos Resultados Pólipos versus Aspirina

Grupo	*% do Desenvolvimento de Pólipos**
Aspirina	17
Sem aspirina (placebo)	27

**Tamanho amostral total = 635 (aproximadamente a metade foi dividida aleatoriamente em cada um dos grupos).*

Os dados crus para este estudo contêm 635 linhas. Cada linha representa uma pessoa e inclui seu número de identificação, o grupo do qual participou (aspirina ou sem aspirina) e desenvolveu-se ou não pólipos durante o período do estudo (sim ou não). Por exemplo, a linha 1 para esse conjunto de dados seria algo como:

NI#22292 GRUPO = ASPIRINA DESENVOLVIMENTO DE PÓLIPOS = NÃO

Se olharmos para esse conjunto de dados bivariados enorme, provavelmente teremos problemas para deduzir qualquer relação entre as variáveis: 635 linhas de dados seriam muita informação para qualquer um (exceto para um computador) analisar e chegar a alguma conclusão. Para os dados que relacionavam o cricrilar dos grilos com a temperatura, mesmo que você consiga ver um padrão geral nos dados crus (por exemplo, ao notar que conforme a frequência do cricrilar dos grilos aumenta, a temperatura também parece aumentar), é muito difícil detectar a relação exata.

Para chegar a alguma conclusão a partir de quaisquer dados, primeiro é necessário organizá-los em tabelas ou gráficos (veja o Capítulo 4). Quando os dados forem bivariados e você estiver procurando por relações entre duas variáveis, os gráficos e as tabelas também precisarão ter duas dimensões, assim como os dados. Essa é a única maneira de conseguir investigar as possíveis ligações entre as variáveis.

Exibindo os dados numéricos bivariados

Nos casos em que ambas as variáveis são quantitativas (isso é, numéricas, tais como peso e altura), os dados bivariados são, normalmente, organizados em um gráfico que os estatísticos chamam de Diagramas de Dispersão (ou de Scatterplot, o termo em inglês). O Diagrama de dispersão possui duas dimensões, uma dimensão horizontal (chamada de eixo-x) e uma dimensão vertical (chamada de eixo-y). Ambas são numéricas; cada uma contém uma linha de números.

Fazendo um diagrama de dispersão

Marcar as observações (ou pontos) em um diagrama de dispersão é semelhante a procurar uma cidade em um mapa que usa letras e números para marcar as seções do mapa. Cada observação possui duas coordenadas; a primeira corresponde à primeira parte dos dados de um par (ou seja, a coordenada *x*. o valor que você coloca à esquerda ou à direita). A segunda coordenada corresponde à segunda parte dos dados de um par (ou seja, a coordenada *y*, o valor que será colocado acima ou abaixo). No local onde as duas coordenadas se encontram é que você deverá marcar o ponto que representa a observação. A Figura 18-1 mostra o diagrama de dispersão para os dados da frequência do cricrilar dos grilos versus a temperatura listados na Tabela 18-1. Devido ao fato de ter colocado os dados em ordem de acordo com seus valores-x (número de vezes do cricrilar), quando criei a Tabela 18-1, os pontos no diagrama (da esquerda para a direita) correspondem à ordem das observações feitas na Tabela 18-1.

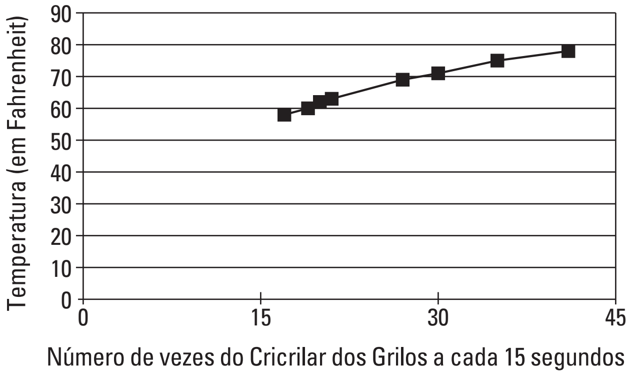


Figura 18-1: Diagrama de Dispersão para os dados da frequência do cricrilar dos grilos versus a temperatura externa.

Interpretando um diagrama de dispersão

Interpretamos um diagrama de dispersão ao examinarmos a tendência dos dados a partir da esquerda para a direita:

- Se os dados formarem uma linha ascendente, conforme você se movimenta da esquerda para a direita, isso indica a existência de uma relação linear (ou proporcional) positiva. Conforme *x* aumenta (movendo-se uma unidade para a direita), *y* aumenta (movendo-se para cima) um determinado valor.

- Se os dados formarem uma linha descendente conforme você se movimenta da esquerda para a direita, isso indica a existência de uma relação linear (ou proporcional) negativa. Significa que conforme *x* aumenta (movimentando-se uma unidade para a direita), *y* diminui (movimentando-se para baixo) uma determinada quantia.

- Se os dados não se parecerem com nenhum tipo de linha (nem mesmo vagamente), isso indica que não existe nenhum tipo de relação linear entre as variáveis.

Se observarmos a Figura 18-1, veremos que parece existir uma relação linear positiva entre o número de vezes que um grilo crica e a temperatura externa. Ou seja, quanto mais o grilo cricrilar, maior é a temperatura. Mas, será essa uma relação de causa-efeito (ou seja, o aumento da temperatura faz os grilos cricrilar mais?). Não podemos afirmar isso, pois esses dados vieram de um estudo observacional e não de um experimento (veja o Capítulo 17).

Os recursos visuais utilizados para exibir os dados mostram as possíveis associações ou relações entre duas variáveis. No entanto, apenas porque seu gráfico ou tabela está mostrando que algo está acontecendo, isso não significa que exista uma relação de causa e efeito. Por exemplo, se olharmos o diagrama dispersivo do consumo de sorvete e das taxas de homicídio, veremos que essas duas variáveis também possuem uma relação linear positiva. No entanto, ninguém irá alegar que o consumo de sorvete causa homicídios ou que as taxas de homicídios afetam o consumo de sorvete. Se alguém estiver tentando mostrar uma relação de causa e efeito por meio de um gráfico ou tabela, vá mais fundo e descubra como o estudo foi projetado e como os dados foram coletados e, só então, avalie adequadamente o estudo usando os critérios destacados no Capítulo 17.

Existem outras tendências além das relações lineares positivas e negativas. As variáveis podem se relacionar umas com as outras por meio de relações em curvas ou exponenciais; mas essas estão além do escopo deste livro. A boa notícia é, no entanto, que muitas relações podem ser caracterizadas como relações lineares positivas e negativas.

Exibindo os dados bivariados categóricos

Nos casos em que ambas as variáveis são categóricas (como o sexo dos respondentes e se o respondente apoia ou não o presidente), os dados são, normalmente, sintetizados no que os estatísticos chamam de tabela de dupla entrada (tabelas cujas linhas representam as categorias da primeira variável e cujas colunas representam as categorias da segunda variável).

Por exemplo, no estudo da aspirina versus os pólipos colorretais, as duas variáveis eram categóricas: se o paciente com câncer de cólon recebeu ou não aspirina (sim ou não) e se ele desenvolveu ou não mais pólipos (sim ou não). Observe que a Tabela 18-2, na seção "Retratando as Relações: Enredos e Tabelas", é uma tabela de dupla entrada.

Uma maneira mais agradável de organizar os dados com duas dimensões é usar um gráfico de barras ou uma série de gráficos pizza. A Figura 18-2 mostra um gráfico de barras que indica a porcentagem de paciente do grupo de tratamento (medicados com a aspirina) que desenvolveram pólipos, comparados com pacientes do grupo de controle (medicados com o placebo). A Figura 18-3 mostra dois gráficos pizza, um para o grupo da aspirina e outro para o grupo do placebo. Cada gráfico pizza mostra a porcentagem de pessoas que desenvolveu pólipos em cada um dos grupos.

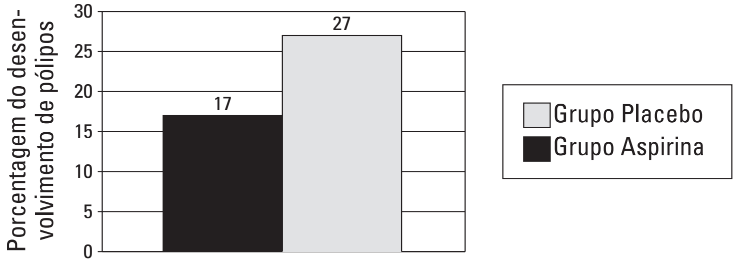


Figura 18-2: Gráfico de barra mostrando os resultados do estudo pólipos versus aspirina.

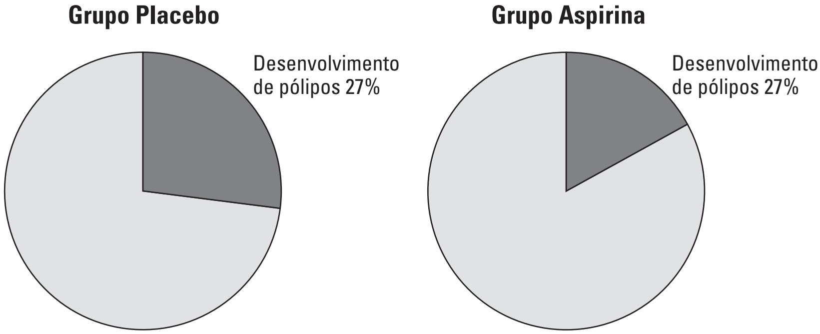


Figura 18-3: Gráficos pizza mostrando os resultados do estudo pólipos versus aspirina.

Devido aos diferentes tamanhos das barras do gráfico de barras e à diferença na aparência dos dois gráficos pizza, realmente parece existir uma relação entre o tratamento com a aspirina e o desenvolvimento de pólipos entre os indivíduos desse estudo (a palavra chave aqui é "parece". É necessário realizar testes de hipótese para as duas proporções, a fim de se ter certeza de que essas diferenças na amostra podem ser aplicadas às suas respectivas populações. Veja o Capítulo 15, para mais informações sobre o assunto).

FIQUE ATENTO

Desconfie de qualquer um que queira tirar conclusões a respeito de relações entre duas variáveis apenas usando um gráfico ou tabela. As aparências enganam (veja o Capítulo 4). As medições e os testes estatísticos devem ser realizados para demonstrar que tais relações são estatisticamente significativas (veja o Capítulo 14).

DICA

No caso da aspirina versus os pólipos, você pode achar que a segunda variável (os pólipos) é numérica, pois seu valor é representado por uma porcentagem. Mas, na verdade, o caso não é esse. As porcentagens são apenas uma maneira prática de resumir os dados de uma variável categórica. Nesse exemplo, a segunda variável é o desenvolvimento ou não dos pólipos (sim/não) pelos pacientes (essa é uma variável categórica). A porcentagem apenas resume todos os pacientes nas categorias sim e não.

Quantificando as Relações: As Correlações e Outras Medidas

Depois que os dados bivariados tiverem sido organizados, o próximo passo é obter algumas estatísticas que possam quantificar ou medir a dimensão e a natureza da relação.

Quantificando a relação entre duas variáveis numéricas

Se as duas variáveis são numéricas ou quantitativas, os estatísticos conseguem medir a direção e a força da relação linear entre as variáveis *x* e *y*. Os dados que se assemelham a uma linha ascendente têm uma relação linear positiva, mas não possuem necessariamente uma relação forte. A força da relação depende de quanto os dados assemelham-se a uma linha reta. É claro que existem níveis de variação para a "semelhança a essa linha". Além disso, você tem que distinguir as relações positivas das negativas. E existe uma estatística para medir tudo isso? Com certeza!

Os estatísticos usam o que eles chamam de coeficiente de correlação para medir a força e a direção das relações lineares entre *x* e y.

Calculando o coeficiente de correlação (r)

A fórmula para o coeficiente de correlação (denotado r) é:

$$r = \frac{1}{n-1}\sum\frac{(x-\overline{x})(y-\overline{y})}{s_x s_y}.$$

Para calcular o coeficiente de correlação:

1. **Encontre a média de todos os valores *x* (vamos chamá-la de $\overline{x}$) e a média de todos os valores *y* (vamos chamá-la de $\overline{y}$).**

   Veja no Capítulo 5 como fazer esses cálculos.

2. **Encontre o desvio padrão de todos os valores *x* (s_x) e o desvio padrão de todos os valores *y* (s_y).**

   Consulte o Capítulo 5.

3. **Para cada par (*x, y*) no conjunto de dados, encontre a diferença entre *x* e $\overline{x}$ e entre *y* e $\overline{y}$ e, depois, multiplique essas diferenças uma pela outra.**
4. **Some os produtos**
5. **Divida a soma por $s_x \times s_y$.**
6. **Divida o resultado por *n* – 1, em que *n* é o número de pares (*x, y*).**

Por exemplo, suponha que você tenha o conjunto de dados (3,2), (3,3) e (6,4). Seguindo os passos mencionados, podemos calcular o coeficiente de correlação. Observe que os valores de *x* são 3, 3 e 6 e os valores de *y* são 2, 3 e 4.

1. **$\bar{x}$ é 12 ÷ 3 = 4 e $\bar{y}$ é 9 ÷ 3 = 3.**

2. **Os desvios padrões são *s* = 1,73 e *s* = 1,00.**

3. **As diferenças encontradas no terceiro passo e multiplicadas umas pelas outras é: (3 – 4) (2 – 3) = (-1) (-1) = 1; (3 – 4) (3 – 3) = (-1) (0) = 0; (6 – 4) (4 – 3) = (+2) (+1) = +2.**

4. **Os resultados do terceiro passo somados são iguais a 1 + 0 + 2 = 3.**

5. **Dividindo o resultado do quarto passo por s_x s_y, você obtém 3 ÷ (1,73 1,00) = 3 ÷ 1,73 = 1,73.**

6. **Dividindo o resultado obtido no quinto passo por 3 -1 (2), teremos 0,87.**

Este é o coeficiente de correlação.

Interpretando a correlação

A correlação r sempre é um valor entre -1 e +1.

- A correlação exatamente igual a -1 indica a existência de uma relação linear negativa perfeita.
- A correlação próxima a -1 indica a existência de uma relação linear negativa forte.
- A correlação próxima a 0 significa a não existência de relações lineares.
- A correlação próxima de +1 indica a existência de uma relação linear positiva forte.
- A correlação exatamente igual a +1 indica a existência de uma relação linear positiva perfeita.

Muitas pessoas cometem o erro de pensar que uma correlação igual a -1 é algo ruim, indicando a não existência de relações. Na verdade, é o contrário! Uma correlação de -1 significa que os dados estão alinhados em uma reta perfeita, é a relação linear mais forte que se pode ter. A única coisa é que a reta está em sentido descendente, por isso o sinal negativo!

Qual a distância que você deve ficar de -1 a +1 para indicar uma relação linear forte? Muitos estatísticos gostam de ver correlações acima de +0,6 (ou abaixo de -0,6) antes de entusiasmarem-se. Entretanto, não espere que uma correlação sempre fique em +0,9 ou -0,9; esses são dados reais e dados reais não são perfeitos.

A Figura 18-4 mostra exemplos da aparência de várias correlações, em termo de força e direção das relações.

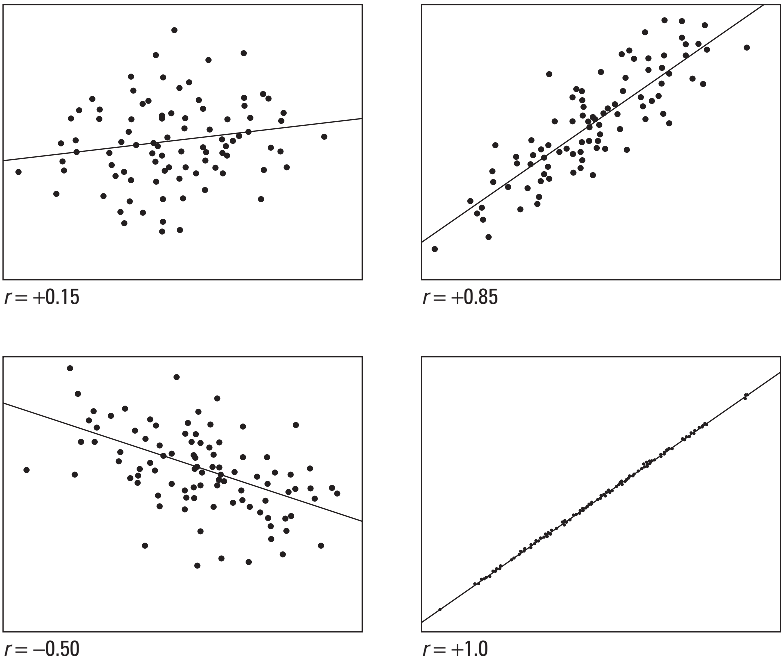


Figura 18-4: Diagramas de Dispersão exibindo várias correlações.

Para o meu subgrupo dos dados do estudo da relação entre a frequência do cricrilar dos grilos e temperatura externa, eu calculei um coeficiente de correlação igual a 0,98, um coeficiente quase nunca encontrado no mundo real (esses grilos são realmente feras!).

Entendendo as propriedades do coeficiente de correlação

Veja aqui algumas propriedades úteis das correlações:

- A correlação é uma medida sem unidade. Isso significa que se você alterar as unidades de x ou y, a correlação continuará a mesma. Por exemplo, mudar a unidade da temperatura de Fahrenheit para Celsius não influenciará a correlação entre o cricrilar dos grilos e a temperatura.
- A ordem dos valores de x e y podem se alterar dentro do conjunto de dados e a correlação ainda permanecerá a mesma.

Quantificando a relação entre duas variáveis categóricas

Se as duas variáveis são categóricas (tais como o tratamento ou não do paciente com aspirina ou se o paciente desenvolveu ou não pólipos), você não pode usar a palavra correlação para descrever a relação existente, pois a correlação mede a força da relação linear entre variáveis numéricas (esse erro acontece o tempo todo na mídia e deixa os estatísticos loucos de raiva!).

A palavra usada para descrever a relação entre dois valores categoriais é associação. Duas variáveis categoriais (tais como o grupo de tratamento e os resultados) associam-se se a porcentagem de indivíduos que apresentaram um determinado resultado em um grupo for significativamente diferente das porcentagens dos que apresentaram os mesmos resultados no outro grupo. No exemplo dos pólipos versus a aspirina, discutido na primeira seção deste capítulo, os pesquisadores descobriram que no grupo medicado com a aspirina, 17% dos pacientes com câncer de cólon desenvolveram pólipos, enquanto no grupo tratado com placebo, 27% dos pacientes tiveram o problema. Devido à razoável diferença desses resultados, as duas variáveis estão associadas.

Mas, qual precisa ser a diferença entre as porcentagens para que se possa determinar uma associação significativa entre as variáveis? A diferença encontrada na amostra deve ser estatisticamente significativa. Desta forma, as mesmas conclusões acerca de uma relação podem ser feitas para toda uma população e não apenas para um conjunto de dados em particular. Um teste de hipótese das duas proporções funcionaria muito bem para esse fim (veja o Capítulo 15, para detalhes sobre esse tipo de teste). Eu analisei os dados do estudo aspirina versus pólipos usando um teste e obtive um *p*-valor menor do que 0, 0001, o que significa que os resultados são altamente significativos (veja o Capítulo 14, para mais detalhes sobre *p*-valor). Você consegue entender agora por que os pesquisadores interromperam o estudo e decidiram medicar todos os pacientes com a aspirina!

Explicando a Relação: Associação e Correlação versus Causalidade

Se duas variáveis se mostram associadas ou correlacionadas, isso não significa necessariamente que exista uma relação de causa e efeito entre elas. A descoberta da existência de uma relação de causalidade entre as variáveis depende de como o estudo foi realizado. Apenas um experimento muito bem conduzido (veja o Capítulo 17) ou vários estudos observacionais podem mostrar que uma associação ou correlação tenha evoluído para uma relação de causa e efeito.

A aspirina realmente parece funcionar

Senti confiança nas conclusões a que os pesquisadores chegaram com o estudo da aspirina versus os pólipos, discutido na primeira seção deste capítulo. Esse estudo foi um experimento bem projetado, segundo os critérios estabelecidos no Capítulo 17. Ele incluía a divisão aleatória dos pacientes aos tratamentos, apresentava tamanhos amostrais grandes o suficiente para a obtenção de informação precisa e controlava as variáveis confusas. Isso significa que os pesquisadores estavam realmente fazendo jus ao título do comunicado de impressa, "Aspirina previne Pólipos em Paciente com Câncer de Cólon". Graças ao projeto do estudo, é possível dizer que existe uma relação de causa e efeito (associação) entre o uso diário da aspirina pelos pacientes com câncer de cólon e o desenvolvimento de pólipos.

Esquentando o cricrilar dos grilos

A temperatura externa pode, realmente, influenciar a frequência com que os grilos cricrilam? (É obvio que o inverso não é possível, mas será que existe alguma possibilidade nesse sentido?) Algumas pessoas afirmam que as mudanças na temperatura externa fazem com que os grilos cricrilem em diferentes frequências. Porém, não conheço nenhum dado baseado em experimento (em oposição a estudos observacionais) que confirmariam ou negariam uma relação de causa e efeito para esse caso. Talvez você devesse fazer seu próprio experimento e esquentar alguns grilos para ver o que acontece! (Mas antes de pular de cabeça nessa ideia – sim, o trocadilho foi pretendido – projete um bom experimento, seguindo os critérios descritos no Capítulo 17).

Fazendo Previsões: Regressão e Outros Métodos

Depois de ter encontrado uma relação entre duas variáveis e tiver alguma forma de quantificá-la, você pode criar um modelo que permita usar uma variável para prever a outra.

Fazendo previsões com dados correlatos

No caso de duas variáveis numéricas, se uma forte correlação tiver sido estabelecida, os pesquisadores provavelmente usarão a relação entre *x* e *y* para fazer previsões. Uma vez que *x* é correlato a *y*, isso significa que existe uma relação linear entre os dois e que você pode descrevê-la usando uma reta. Se você souber a inclinação e a intercepção-y dessa linha, então pode dar um valor para *x* e prever o valor médio para *y*. Ou seja, você pode prever *y* a partir de *x.*

Uma vez que a correlação entre o cricrilar dos grilos e a temperatura é tão alta (r = 0,98), é possível encontrar uma reta que se ajuste aos dados. Isso significa que você deve encontrar a reta que melhor se ajuste aos dados (em termos de distância média de todos os pontos em um conjunto de dados à linha que você criar). Os estatísticos chamam essa busca pela linha que melhor se ajusta de análise de regressão.

Nunca faça uma análise de regressão, a menos que você já tenha encontrado uma forte correlação (positiva ou negativa) entre duas variáveis. Já vi casos em que os pesquisadores vão em frente, fazendo previsões, mesmo tendo encontrado um correlação menor do que 0,02! Isso não faz nenhum sentido! Se o diagrama dispersivo dos dados não se parece com uma reta, como você pode tentar usar uma reta para se ajustar aos dados e fazer previsões sobre a população.

LEMBRE-SE

Antes de examinar qualquer modelo que preveja uma variável a partir da outra, encontre primeiro a correlação; se a correlação for muito fraca, pare imediatamente.

Você deve estar pensando que terá que tentar várias e várias linhas diferentes até encontrar a que se ajuste melhor. Felizmente, não é esse o caso (embora medir uma linha a olho no diagrama dispersivo realmente ajuda a imaginar qual seria a resposta esperada). A linha que melhor se ajustar tem uma inclinação e uma intercepção-y diferentes que podem ser calculadas através de fórmulas (e, além disso, essas fórmulas não são tão difíceis de serem calculadas).

A fórmula para a linha que melhor combina

A fórmula para a linha que melhor se ajusta aos dados (a linha de regressão) é $y = mx + b$, em que m é a inclinação da linha e b é a intercepção-y. A inclinação de uma linha é a alteração ocorrida em *y* depois de uma alteração em *x.* Por exemplo, uma inclinação de 103 significa que, conforme você se move de um ponto a outro na linha, a cada 3 unidades que *x* se mover para a direita, o valor de *y* se moverá 10 unidades para cima. A intercepção-y é o local no eixo *y* onde as linhas se cruzam. Por exemplo, na equação $y = \frac{10}{3}x - 6$, a linha cruza o eixo *y* no ponto -6. As coordenadas para esse ponto são (0, -6) – e, uma vez que você está cruzando o eixo *y,* o valor *x* da intercepção-y é sempre igual a 0. De modo a obter a linha de regressão, é preciso encontrar os valores de m e b para que você possa ter a real equação de uma linha (por exemplo, $y = 2x + 3$; ou $y = -10x - 45$).

DICA

Para não perder tempo calculando a linha de regressão, tenha em mente que as cinco estatísticas de síntese mais conhecidas são tudo o que você precisa saber para fazer todos os cálculos necessários. Os estatísticos as chamam de as super cinco estatísticas de síntese:

- A média dos valores *x* (representada por $\bar{x}$).
- A média dos valores *y* (representada por $\bar{y}$).
- O desvio padrão dos valores *x* (representado por s_x).
- O desvio padrão dos valores *y* (representado por s_y).
- A correlação entre *x* e *y* (representada por r).

(Este capítulo e o Capítulo 5 contêm fórmulas e instruções passo a passo para o cálculo dessas estatísticas).

Encontrando a inclinação da linha de regressão

A fórmula para a inclinação, *m*, da linha de regressão é $m = r\left(\frac{s_y}{s_x}\right)$, onde *r* é a correlação entre *x* e *y* e s_y e s_x são os desvios padrões dos valores *y* e *x*, respectivamente (veja o Capítulo 5, para mais detalhes sobre o desvio padrão).

Para calcular a inclinação, *m*, da linha de regressão:

1. **Divida s_y por s_x.**

2. **Multiplique o resultado obtido por r.**

A inclinação da linha de regressão pode ser um número negativo, pois a correlação pode ser negativa. Uma inclinação negativa indica que a linha está em descendência.

A fórmula para a inclinação apenas dá uma unidade à correlação. Pense em $s_y \div s_x$ como sendo a alteração ocorrida a y em função da alteração ocorrida em x, com os desvios padrões cada um em sua unidade original.

Encontrando a intercepção-y da linha de regressão

A fórmula para a intercepção-y, b, da linha de regressão é $b = \bar{y}$ - m$\bar{x}$, em que $\bar{y}$ e $\bar{x}$ são as médias dos valores y e x, respectivamente, e m é a inclinação (cuja fórmula foi dada na seção anterior).

Para calcular a intercepção-y, b, da linha de regressão:

1. **Encontre a inclinação, m, da linha de regressão usando os passos listados na seção anterior.**

2. **Multiplique m por $\bar{x}$.**

3. **Subtraia $\bar{y}$ do resultado obtido no segundo passo.**

Sempre calcule a inclinação antes de calcular a intercepção-y. A fórmula para a intercepção-y necessita do valor da inclinação. Logo, você precisará de m para calcular b.

Descobrindo a linha de regressão para o cricrilar dos grilos e a temperatura

Embora a fórmula para a linha de regressão da relação entre o cricrilar dos grilos e a temperatura seja objeto de discussões (veja o Apêndice), o senso comum parece ser de que um bom modelo de trabalho para essa relação é $y = x + 40$, ou a temperatura = 1 × (número de cricrilos a cada 15 segundos) + 40, em que a temperatura está em Fahrenheit. Note que a inclinação dessa linha é igual a 1, x = os números de cricrilos a cada 15 segundos e y = a temperatura em Fahrenheit.

Observe que as fórmulas para a inclinação e para a intercepção-y estão em forma de x e y, portanto você precisa decidir qual das suas variáveis será a x e qual será a y. Ao fazer correlações, a escolha de qual será a variável x ou y não importa, desde que seus dados sejam consistentes; mas, ao determinar a linha de regressão e fazer previsões, a escolha de x e y faz toda a diferença. Dê uma olhada nas fórmulas apresentadas até agora – alterar a ordem de x e y faz com que todas elas se alterem.

Então, como determinar qual variável é qual? De modo geral, x é a variável que vem antes. Os estatísticos chamam x de variável explanatória, pois se você alterar x. explicará porque e como y irá se alterar. Nesse exemplo, x é o número de vezes que os grilos cricrilam a cada 15 segundos, a variável y é chamada de variável resposta; pois reage às alterações ocorridas com x. Em outras palavras, y está sendo previsto por x. Nesse caso, y é a temperatura.

Comparando o modelo de trabalho ao subconjunto de dados

As "super cinco" estatísticas de síntese para o subconjunto dos dados dos grilos são mostradas na Tabela 18-3.

Tabela 18-3 As super cinco estatísticas de síntese para os dados dos grilos

Variável	*Média*	*Desvio Padrão*	*Correlação*
# Cricrilar (x)	$\bar{x}$ = 26,5	s_x = 7,4	r = +0,98
Temperatura (y)	$\bar{y}$ = 67	s_y = 6,8	

A inclinação m da linha de regressão para o subconjunto de dados da frequência do cricrilar dos grilos versus a temperatura é $r \times (s_y \div s_x) = 0{,}98 \times (6{,}8 \div 7{,}4) = 0{,}98 \times 0{,}919 = 0{,}90$. Agora, para encontrar a intercepção-y, b, você deve fazer $\bar{y} - m \times \bar{x}$, ou 67 – (0,90) (26,5) = 67 – 23,85 = 43,15. Assim, a linha de regressão para prever a temperatura a partir do cricrilar dos grilos, baseando-se nesses dados, é $y = 0{,}9x + 43{,}2$, ou

> Temperatura (em Fahrenheit) = 0,9 × (o número de cricrilos a cada 15 segundos) + 43,2.

Observe que a equação anterior é semelhante, mas não é a mesma do modelo de trabalho: $y = x + 40$. Por que a equação mencionada acima não é exatamente igual ao modelo de trabalho? Duas razões me vêm à mente. Primeiro, "modelo de trabalho" é um eufemismo para: "não necessariamente preciso, mas muito prático". E estou imaginando que, por muitos anos, a inclinação foi arredondada para o número inteiro mais próximo de 1 e a intercepção-y para a dezena mais próxima (por isso o número 40) apenas para ficar mais fácil de ser lembrado e mais divertido de ser contado (essa não é uma boa prática estatística). Em segundo lugar, os dados que usei são apenas um subconjunto aleatório retirado do conjunto de dados originais (com a finalidade de ilustração), e estarão um pouco por acaso (veja o Capítulo 9, para mais informações sobre a variação entre amostras). Entretanto, já que os dados estão tão altamente correlacionados, a diferença entre uma amostra de dados de outra não seria muita.

Prevendo a temperatura usando o cricrilar dos grilos

A linha de regressão para a previsão da temperatura com o uso do cricrilar dos grilos que obtive com meu subconjunto de dados foi $y = 0{,}9x + 43{,}2$. Qualquer equação ou função usada para estimar ou prever a relação entre duas variáveis é chamada de modelo estatístico. Usando esse modelo, você consegue prever a temperatura a partir do cricrilar dos grilos. Mas, como fazer isso? Escolha um valor relevante para x, coloque-o no modelo e encontre o valor esperado para y.

Por exemplo, se você quiser prever a temperatura e souber que o grilo que está no seu jardim cricrilou 35 vezes em 15 segundos, atribua 35 a x

e encontre o valor de *y*. Veja como: $y = 0{,}9(35) + 43{,}2 = 31{,}5 + 43{,}2 = 74{,}7$. Logo, sabendo que o grilo cricrilou 35 vezes em 15 segundos, você pode prever que a temperatura esperada seja de, aproximadamente, 75 graus Fahrenheit (24°C).

O fato de você ter um modelo não necessariamente significa que você possa atribuir qualquer valor a *x* e obter uma boa previsão para *y*. Por exemplo, você não pode atribuir valores maiores do que 39 ou menores do que 18 para esse caso. E por que não? Porque, para esse caso, você não tem nenhum dado para *x* dentro dessa variação (veja a Tabela 18-1). Quem vai dizer que a linha de regressão ainda funcionará fora da área em que os dados foram coletados? Você realmente acredita que conforme a temperatura aumenta, os grilos irão cricrilar cada vez mais e mais rápido, sem nunca parar? Em certo ponto, os coitadinhos iriam morrer por causa da exposição excessiva ao calor. E, da mesma forma, eles não sobreviveriam a temperaturas extremamente baixas e, por isso, você não pode atribuir valores extremamente baixos para *x* e ainda esperar que o modelo funcione.

Fazer previsões usando valores *x* que fiquem fora da variação dos dados é errado. Os estatísticos chamam essa prática de extrapolação; fique atento com pesquisadores que tentam fazer alegações além da variação de seus resultados.

Uma vez que a linha de regressão é um modelo que descreve a relação geral entre *x* e *y*, você não está realmente prevendo *y*, mas está prevendo o valor esperado (ou a média) de *y* para um dado valor *x*.

Fazendo previsões com duas variáveis categóricas associadas

Depois que duas variáveis tenham sido consideradas associadas, é possível fazer previsões (estimativas) acerca da porcentagem em cada grupo com relação a uma das variáveis, ou, também, se pode estimar a diferença entre as porcentagens dos dois grupos. Nos dois casos, você deve usar os intervalos de confiança para chegar a essas estimativas.

Para o exemplo da aspirina versus o desenvolvimento de pólipos (dado na primeira seção deste capítulo), entre os indivíduos do grupo que recebeu aspirina, a porcentagem de pacientes com câncer de cólon que desenvolveu pólipos foi de 17%, comparada a 27% dos indivíduos do grupo que não recebeu a aspirina (consulte a Tabela 18-2). A previsão que você pode fazer aqui é a seguinte: se você for um paciente com câncer de cólon, suas chances de desenvolver pólipos serão menores se você ingerir 325 mg de aspirina ao dia.

Você ainda pode ser mais específico a respeito dessa previsão, podendo estimar as chances de um paciente com câncer de cólon desenvolver os pólipos caso ele tome aspirina diariamente, usando o intervalo de confiança. Uma vez que 17% da amostra dos pacientes que receberam aspirina desenvolveram os pólipos, isso significa que a chance de qualquer indivíduo dessa população (ou seja, pacientes com câncer de cólon) possa desenvolver os pólipos, se ele tomar aspirina diariamente, é 0,17 mais ou menos a margem de erro, que, nesse, caso é igual a 0,04 para mais ou para menos. Ou seja, para um paciente com câncer de cólon que tomar 325 mg de aspirina todos os dias, a chance de que ele desenvolva pólipos em seguida fica em torno de 17% - 4% a 17% + 4%, ou de 13% a 21% (veja o Capítulo 13, para a fórmula do intervalo de confiança para uma única proporção populacional, *p*, e o Capítulo 10, para como encontrar a margem de erro.)

Outra maneira de fazer previsões para essa situação é estimando a diminuição do risco para o desenvolvimento de pólipos, caso os pacientes tomem aspirina diariamente. Podemos conseguir esse resultado por meio do cálculo de um intervalo de confiança para as diferenças entre as duas proporções, em que p_1 é a proporção de pacientes do grupo de controle que desenvolveram o problema, e p_2 é a proporção de pacientes do grupo de tratamento que também apresentaram os pólipos. A diferença pela qual você se interessa aqui é $p_1 - p_2$. Esse intervalo de confiança é 0,27 – 0,17 = 0,10, mais ou menos a margem de erro ou, nesse caso, 0,03 para mais ou para menos. Logo, se você for um paciente com câncer de cólon que toma aspirina todos os dias, seu risco de desenvolver os pólipos diminui de 10% - 3% a 10% + 3%, ou de 7% a 13% (veja o Capítulo 13 para as fórmulas do intervalo de confiança ou para a diferença entre duas proporções populacionais).

Capítulo 19

Estatística e Pasta de dente: Controle de Qualidade

Neste Capítulo

- Observando como a estatística ajuda a melhorar produtos
- Dentro das especificações: fundamentos do controle de qualidade
- Monitorando o processo

As indústrias mais bem sucedidas são as que se preocupam com o controle de qualidade. Tudo o que elas querem é que você, como cliente, esteja satisfeito com seus produtos e compre-os mais e mais. Elas querem que você fique tão contente que fale para seus amigos, vizinhos, colegas de trabalho, e, até mesmo, para as pessoas nas ruas, o quanto os produtos daquela empresa são maravilhosos. Mas, como essas empresas podem garantir que você ficará satisfeito com seus produtos? Um critério usado para obter a satisfação do consumidor é a qualidade do produto e, acredite ou não, a estatística desempenha uma função vital no processo de avaliação e de melhoria da qualidade dos produtos. Este capítulo mostra-lhe como a estatística faz isso.

Satisfazendo as Expectativas

Os clientes esperam que os produtos satisfaçam suas expectativas, e uma das expectativas é de que a embalagem contenha a quantidade de produto prometida. Outra expectativa é a de que exista um certo nível de consistência cada vez que o produto é comprado. Qual o peso esperado para um pacote de batatas chips? Não parece estranho que um pacote de 230g de batata chips pareça ser tão grande e conter tão poucas batatas? (Os fabricantes dizem que colocam ar dentro da embalagem antes de fechá-la para proteger o produto contra danos). Caso o rótulo da embalagem esteja indicando que ela contém 230g e ela realmente contiver 230g, você não pode reclamar. Mas como você se sentirá se a embalagem contiver menos do que o indicado?

Suponha que a embalagem diga que contém 230g, mas, na verdade, contém apenas 220 g; você ficará chateado? Provavelmente, você nem notará essa diferença. Mas, e se a embalagem contiver apenas 170g de batatas? Ou, então, 110g? Em algum momento, você irá notar. E qual será sua reação? Você pode:

- Deixar para lá (desde que o problema não se repita).
- Devolver o produto na loja e exigir reembolso.
- Escrever uma carta de reclamação para a empresa.
- Decidir não comprar mais o produto.
- Reclamar no Procon ou em qualquer outro órgão de defesa do consumidor
- Organizar um boicote ao produto
- Tentar conseguir um emprego na empresa para "fazer parte da solução e não do problema".

Agora, algumas dessas opções parecem ser um pouco exageradas, especialmente se você foi lesado com apenas algumas batatas. Mas, imagine que você tenha comprado um carro novo, que acabou virando um grande abacaxi, ou seu filho quase tenha se engasgado com uma peça que veio solta no berço ou, ainda, que você tenha passado mal depois de comer um lanche comprado no dia anterior. A qualidade pode ser uma questão muito séria e decisiva. Embora os padrões para muitos produtos sejam desenvolvidos e reforçados por órgãos governamentais (como o Inmetro), problemas realmente surgem de tempos em tempos durante o processo de fabricação. Veja a seguir apenas alguns dos fatores que podem influenciar a qualidade de um produto durante a sua fabricação:

- O desempenho inconsistente dos funcionários (em virtude de níveis diferentes de capacidades, treinamento, condições de trabalho ou em virtude de alterações do turno, erro humano, indisposição, etc.).
- Os gerentes ou supervisores são inconsistentes e/ou confusos em relação a expectativas e/ou respostas aos problemas surgidos.
- Inconsistência do desempenho das máquinas e equipamentos (em virtude da falta de manutenção, do desgaste das peças, da quebra ou do mau funcionamento das máquinas, ou, simplesmente, por causa das diferenças individuais entre as diferentes máquinas e linha de produção).
- Falta de precisão ou sensibilidade suficiente das máquinas e equipamentos.
- Inconsistência das matérias-primas usadas no processo de fabricação.
- Falta de controle consistente do meio externo (temperatura, umidade, pureza do ar, etc).
- Insuficiência ou ineficiência da monitoração do processo.

Pressionados pela necessidade de agradar os clientes e seguir as regulamentações governamentais, os fabricantes bem sucedidos estão sempre em busca de maneiras para melhorar a qualidade de seus produtos. Um conceito popularmente usado pela indústria é o gerenciamento de qualidade total ou TQM (do termo em inglês Total Quality Management). O gerenciamento de qualidade total está focado no desenvolvimento de formas para monitorar, avaliar e melhorar continuamente o processo de fabricação, do começo ao fim. O TQM foi popularizado nos Estados Unidos pelo famoso e muito querido estatístico Dr. W. Edwards Deming, que desenvolveu uma lista conhecida em todo o país como "Os 14 pontos do Gerenciamento". A filosofia de Deming era de que se você, em primeiro lugar, construísse a qualidade de seu produto, teria custos mais baixos, aumento da produtividade e seria mais competitivo.

Qual é o papel da estatística no processo de qualidade de um produto? A estatística é utilizada para determinar e estabelecer especificações, além de monitorar todos os aspectos do processo de manufatura para garantir que as especificações sejam cumpridas. A estatística também é usada para ajudar a decidir quando um processo deve ser interrompido, assim como identificar os problemas antes que eles ocorram. De modo geral, os dados estatísticos fornecem um feedback (de forma contínua) ao fabricante sobre a qualidade de seus produtos, como parte da filosofia de gerenciamento de qualidade total. O papel da estatística em monitorar e melhorar a qualidade dos produtos ao longo de todo o processo de fabricação chama-se controle de processo estatístico (CEP). O controle de processo estatístico pode, sozinho, fornecer assunto para um livro inteiro; Entretanto, você terá uma boa noção geral das funções da estatística dentro do controle de qualidade, ao entender e aplicar algumas ideias básicas discutidas nas seções a seguir.

Espremendo Qualidade de um Tubo de Pasta de Dente

Embora os consumidores pareçam ter chegado a um nível de aceitação de que as embalagens de batatas chips sempre estarão (ou, pelo menos, parecerão estar) mal cheias, eles ainda colocam os tubos de pasta de dente em um padrão mais elevado, esperando que estejam preenchidos consistentemente até o máximo (os fabricantes de pasta de dente devem saber o quanto você luta para espremer até o fim a pasta de dente do tubo).

Felizmente, a indústria de creme dental (que inclui seu próprio Conselho especializado) levantou a questão e leva todo esse conceito muito a sério. Você até consegue encontrar "As Perguntas Mais Frequentes sobre o Creme Dental", no site de uma empresa especializada nessa área. Uma dessas perguntas refere-se à quantidade garantida no tubo da pasta.

Os objetivos dos equipamentos de preenchimento do creme dental, segundo a indústria, são a precisão e a consistência. O mecanismo de dosagem é a chave para o cumprimento desses objetivos (mecanismo de dosagem é o jargão usado pela indústria para se referir à máquina que realmente enche os tubos). A seguir, confira algumas características importantes a respeito dos equipamentos de dosagem de tubos de pasta de dente:

- Mecanismo que garante a correta interrupção do fluxo, eliminando vazamentos.
- Sistema que elimina a entrada de ar durante o processo.
- Mecanismo para evitar que a máquina tente encher os tubos que, por algum motivo, estejam faltando.
- Sistema projetado para facilitar a limpeza e a troca do produto.

Se esse nível de complexidade e atenção aos detalhes é necessário para garantir a qualidade do preenchimento de tubos de pasta de dente, imagine então o que está envolvido na construção de qualidade de um avião!

Ficou claro que a qualidade do preenchimento dos tubos de pasta de dente é afetada por vários fatores, incluindo os listados acima. Os problemas que os fabricantes de pasta de dentes precisam evitar incluem o preenchimento inferior (principalmente em virtude de bolhas de ar) e o excesso de preenchimento (que resulta em tubos "traidores", para usar o jargão da indústria; o excesso de preenchimento também resulta na liberação de produto grátis, o que irá refletir nos lucros). As dimensões da parte interna do tubo também podem desempenhar um importante papel na qualidade do preenchimento. Por exemplo, os tubos menores do que o normal (mesmo que sejam preenchidos com a quantidade correta de pasta de dente) irão inchar quando forem selados, e os tubos maiores do que o normal darão a impressão de que contém menos do que deveria.

Entendo a fórmula: qualidade = precisão + consistência

A estatística certamente está envolvida no fornecimento de dados necessários à avaliação dos equipamentos da pasta de dente, segundo cada critério listado na seção anterior. Entretanto, o papel da estatística no processo de fabricação fica mais claro pelo critério do fabricante para a qualidade do processo de enchimento dos tubos: consistência e precisão. As palavras consistência e precisão representam a estatística melhor do que qualquer outra palavra no contexto industrial, elas praticamente dizem tudo sozinhas.

A precisão e a consistência são estatisticamente monitoradas pelo uso de gráficos de controle. O gráfico de controle é um gráfico de linhas específico, que exibe os valore dos dados na ordem em que eles foram coletados ao longo do tempo (veja o capítulo 4, para mais informações a respeito do gráfico de linhas). Os gráficos de controles usam uma linha para denotar onde se encontra o valor especifico do fabricante – ou o valor alvo (isso se relaciona às questões de precisão) – e limites para indicar a distância acima ou abaixo do alvo esperada para os valores (isso se refere à questão da consistência). Os valores a serem exibidos no gráfico representam o peso, o volume ou a contagem de produtos individuais ou, mais comumente, eles representam o peso médio, o volume médio ou a contagem média das amostras dos produtos. Os limites superiores e inferiores de um gráfico de controle são chamados de Limite de Controle Superior (LCS) e Limite de Controle Inferior (LCI).

Por exemplo, suponha que um fabricante de doces esteja ensacando pacotes de balas e seu valor alvo seja de 50 balas por pacote, com LCI = 45 balas e LCS = 55 balas. Imagine que 8 pacotes são coletados para amostra, as balas são contadas e os resultados encontrados foram: 51 balas, 53 balas, 49 balas, 51 balas, 54 balas, 47 balas, 52 balas e 45 balas. A Figura 19-1 mostra uma imagem do gráfico de controle resultante para esse processo. O processo (pelo menos em seu início) parece estar sob controle.

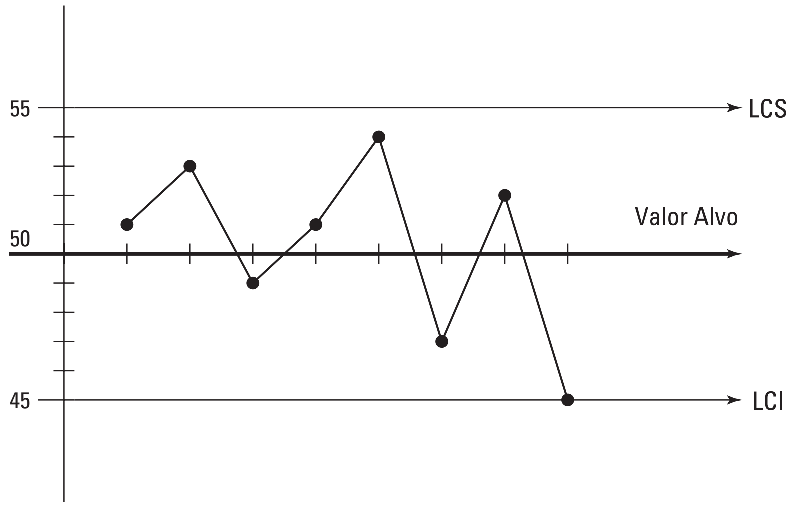


Figura 19-1: Gráfico de Controle para processo de ensaque de balas.

Usando gráficos de controle para monitorar a qualidade

Para usar a estatística na monitoração da qualidade, primeiro você tem que definir e medir a precisão e a consistência. Depois, é preciso estabelecer o valor alvo, determinar os limites superiores e inferiores e coletar os dados do processo. Os dados, então, precisam ser registrados no gráfico de controle e o processo necessita ser monitorado, a fim de determinar se ele está ou não sob controle. Essa última parte pode ser uma decisão complicada. Por um lado, você não quer interromper o processo por causa de um alarme falso (o que você poderia estar fazendo se interrompesse o processo na primeira vez em que um valor ficasse fora dos limites de controle). Por outro lado, você não quer deixar que o processo saia do controle se ele começar a produzir produtos inferiores.

Definindo a precisão

O que significa precisão para uma máquina que enche tubos de pasta de dente? Significa que o tubo deve pesar, no final, o que se espera que ele pese, em média. Observe, eu disse "em média". Você concordaria, acredito, que mesmo o processo de fabricação mais sofisticado não é perfeito e, sendo, assim, um pouco de variação, em qualquer processo, é normal devido a flutuação aleatória. Isso significa que não podemos esperar que todos os tubos de pasta de dente pesem exatamente 6,4 onças (181,44g). Entretanto, se o peso do tubo começar a ficar cada vez menor com relação a seu peso real ou se, de repente, começar a encher de ar e não conter a quantidade que deveria, os consumidores irão notar e a qualidade do produto será comprometida. Da mesma forma, se o peso médio começa a ficar maior, o fabricante irá perder dinheiro, pois estará dando produto a mais.

Falando do ponto de vista estatístico, o peso dos produtos é preciso se não contiver nenhum viés ou erro sistemático (os erros sistemáticos levam a valores consistentemente altos ou baixos, comparados aos valores esperados). Isso de encaixa perfeitamente dentro do conceito de precisão da indústria de pasta de dente. Nesse caso, o valor esperado (valor alvo) é o conjunto de especificação estabelecido pelo fabricante, tal como 6,4 onças (181,44g).

Definindo a consistência

O que significa consistência para uma máquina que enche tubos de pasta de dente? Significa que o peso dos tubos deve ficar dentro dos limites de controle, na maioria das vezes. Observe que eu disse "na maioria das vezes". Mais uma vez, um pouco de variação é normal em qualquer processo em virtude da flutuação aleatória. Entretanto, se o peso do tubo começar a pular de um lado para o outro, a qualidade do produto fica comprometida.

Falando do ponto de vista estatístico, o peso é consistente se seu desvio padrão for pequeno (veja o Capítulo 4). Qual deveria ser o tamanho do desvio padrão? Isso depende das especificações do fabricante e de algumas limitações do processo. Os operadores de máquina que preenchem os tubos de pasta de dente dizem que seus equipamentos têm precisão para ficar dentro de 0,5%. Isso significa que eles esperam que a maioria dos tubos com rótulo de 6,4 onças (181,44 g) fique 0, 032 onças (0,90 g) dentro do valor alvo (pois 0,5% de 6,4 é $0{,}005 \times 6{,}4 = 0{,}032$ onças). Suponha que você considere que eles queiram que 95% dos tubos fiquem dentro dessa variação. Quantos desvios padrões são necessários para abranger 95% dos valores ao redor do alvo?

Segundo a regra empírica (que pode ser usada, pois consideraremos que os pesos dos tubos de pasta de dente assumirão uma distribuição normal padrão – veja o Capítulo 8), 95% dos pesos estarão dentro de 2 desvios padrões com relação ao valor alvo. Portanto, 0,032 = 2 vezes o desvio padrão dos pesos. Isso significa que cada desvio padrão deveria ser igual, no máximo, a $0{,}032 \div 2 = 0{,}016$ onças (0,45 g), segundo as especificações do fabricante da máquina. Essa é uma estimativa conservadora; as especificações do fabricante podem ser mais amplas do que os reais limites de controle que eles estabelecem durante seu teste de controle de qualidade.

Esperando uma distribuição normal

Embora se estabeleça que os tubos devam pesar 6,4 onças (181,44g), nem todos os tubos pesarão exatamente esse valor, é claro. Alguns ficarão acima, outros abaixo, mas, se o processo estiver sob controle, você deve esperar que a maioria pese o mais próximo possível do valor determinado, apresentando a mesma porcentagem acima e abaixo do valor alvo (dentro dos limites de controle). Além de ter a forma de uma montanha no centro, os pesos dos tubos da pasta de dente devem realmente apresentar uma distribuição normal padrão (veja o Capítulo 8, para mais detalhes a respeito da distribuição normal). Se o processo estiver preciso, a média da distribuição será o valor alvo, indicado por μ, e você saberá que o fabricante não quer que o desvio padrão seja maior do que 0,016 onças (0,45 g) para os propósitos de consistência. Isso significa que $\mu = 6{,}4$ onças e $\sigma = 0{,}016$ onças. Veja a Figura 19-2.

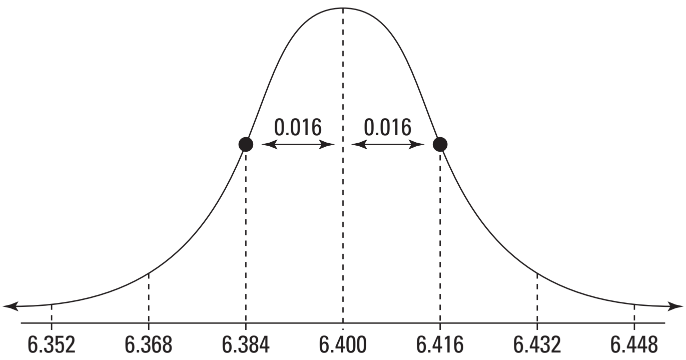


Figura 19-2: Distribuição dos pesos individuais dos tubos de pasta de dente (quando o processo está sob controle).

PAPO TÉCNICO

O desvio padrão de uma população é denotado por σ e o desvio padrão de uma amostra é denotado por *s* (veja o Capítulo 5, para mais detalhes a respeito de desvios padrões). Na maioria das situações estatísticas, o desvio padrão de uma população é desconhecido e, por isso, você deve usar *s* para estimá-lo a partir da amostra. Porém, no caso do controle de qualidade, o desvio padrão para uma população de produtos é estabelecido pelo fabricante. Espera-se que o desvio padrão para a população seja o valor estabelecido, caso o processo esteja sob controle.

Encontrando os limites de controle

Depois que o valor tiver sido estabelecido e o desvio padrão determinado, o próximo passo no controle de processo estatístico é estabelecer os limites de controle para o processo.

Se o fabricante irá pesar os tubos de pasta de dente individualmente, o limite de controle irá ser estabelecido a partir do valor alvo (média) mais ou menos 2 desvios padrões (para 95% de confiança) ou a partir do valor alvo (média) mais ou menos 3 desvios padrões (para 99% de confiança). As fórmulas para os limites de controle para os pesos individuais são: $\mu \pm 2\sigma$ ou $\mu \pm 3\sigma$, respectivamente.

Os fabricantes de pasta de dente estabeleceram a média em 6,4 e o desvio padrão em 0,016. Suponha que eles desejem 95% de confiança. Isso significa que os limites de controle seriam: $6{,}4 \pm 2\ (0{,}016) = 6{,}4 \pm 0{,}032$. O LCI é $6{,}4 - 0{,}032 = 6{,}368$ onças (180,53) e o LCS é $6{,}4 + 0{,}032 = 6{,}432$ onças (182,34 g). Se os fabricantes quiserem 99% de confiança, os limites de controle para os pesos individuais dos tubos seriam: $6{,}4 \pm 3\ (0{,}016) = 6{,}4 \pm 0{,}048$. O LCI é $6{,}4 - 0{,}048 = 6{,}352$ onças (180,07) e o LCS é $6{,}4 + 0{,}048 = 6{,}448$ onças (182,80 g).

Entretanto, muitos processos são monitorados através da retirada de amostras e do peso médio encontrado para cada amostra, ao invés de observar os produtos individualmente. Isso significa que os limites de controle incluirão o valor alvo (média), mais ou menos 2 erros padrões (para 95% de confiança), ou o valor alvo (média), mais ou menos 3 erros padrões (para 99% de confiança). O erro padrão é o desvio padrão das médias amostrais e é calculado a partir do desvio padrão dos pesos dividido pela raiz quadrada de*n,* onde *n* é o tamanho amostral (veja o Capítulo 9 e 10 para mais informações sobre o erro padrão). A equação para o erro padrão é $\frac{\sigma}{\sqrt{n}}$. As fórmulas para os limites de controle das médias amostrais são dadas por: $\mu \pm 2\frac{\sigma}{\sqrt{n}}$, or $\mu \pm 3\frac{\sigma}{\sqrt{n}}$, respectivamente.

PAPO TÉCNICO

O erro padrão sempre será menor do que o desvio padrão. Isso porque as médias são mais consistentes do que os valores individuais, uma vez que se baseiam em mais dados e, portanto, não irão variar tanto de uma amostra para outra. Quanto maior for o tamanho amostral, menor será o erro padrão (pois com *n* sendo o denominador, conforme *n* aumenta, o valor do desvio padrão dividido pela raiz quadrada de *n* diminui). Veja mais detalhes no Capítulo 10.

DICA

Se o processo for monitorado através da pesagem de cada tubo que sai da linha de produção, isso é o mesmo que monitorar tamanhos amostrais de $n=1$. Portanto, para $n=1$, as fórmulas para o desvio padrão e para o erro padrão são iguais.

Suponha que o tamanho de cada amostra de tubos de pasta de dente seja igual a 10. Dado que o desvio padrão está estabelecido em 0,016, o erro padrão para as médias amostrais (cada uma com $n=10$) é $\frac{0.016}{\sqrt{10}} = 0.005$ onças (0,14 g). Se o processo estiver sob controle, a distribuição dos pesos médios será normal, com média 6,4 e erro padrão igual a 0,005. Veja a Figura 19-3.

Se considerarmos que os fabricantes de pasta de dente estão usando 3 desvios padrões para seu nível de consistência aceitável (para ser conservadores), a fórmula para os limites de controle serão $\mu \pm 3\frac{\sigma}{\sqrt{n}} = 6.4 \pm 3 \times \frac{0.016}{\sqrt{10}} = 6.4 \pm 3(0.005)$. O limite de controle inferior

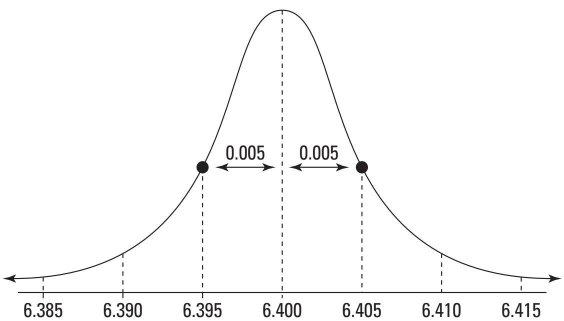


Figura 19-3: Distribuição das médias amostrais de 10 tubos de pasta de dente (em que o processo está sob controle)

(LCI) é 6,4 – 3 (0,005) = 6,4 – 0,015 = 6,385 onças (181,01g) e o limite de controle superior é 6,4 + 3(0,005) = 6,4 + 0,015 = 6,415 onças (181,86 g). O valor alvo e os limites de controle para esse processo, em particular, são ilustrados no gráfico de controle da Figura 19-4.

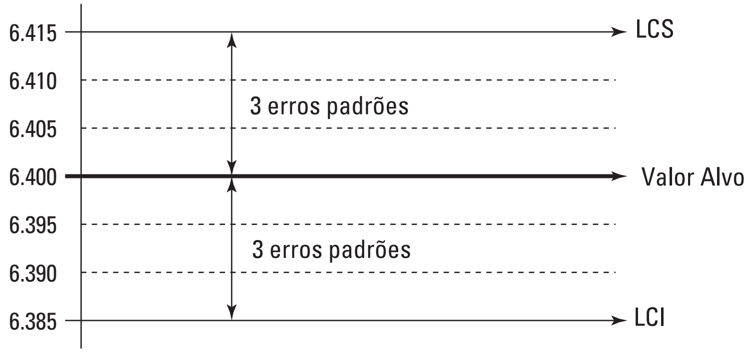


Figura 19-4: Limites de controle para as médias amostrais de dez tubos de pasta de dente (com 99% de confiança).

Monitorando o processo

Depois que os limites de controle tiverem sido estabelecidos, o próximo passo é monitorar o processo. Na maioria dos casos, isso envolve coletar amostras dos produtos em vários momentos, descobrir seu peso médio e marcar essas médias em um gráfico de controle. Quando parecer que o processo estiver saindo dos trilhos, tanto em termos de precisão quanto de consistência, ele deve ser interrompido, o problema identificado e reparos ou ajustes devem ser feitos.

Se um processo estiver sob controle, você poderá ver cerca de 68% das médias amostrais se enquadrando dentro de 1 erro padrão, 95% das médias dentro de 2 erros padrões e 99% em 3 erros padrões em relação ao valor alvo, de acordo com a regra empírica (veja o Capítulo 8). A média amostral total deveria ser o valor alvo, e você deveria observar o máximo de valores abaixo e acima do alvo, sem padrões específicos.

No atual processo de produção, quando uma máquina de preenchimento de tubos tem que parar, toda a linha de produção é interrompida. Tempo valioso e capacidade de produção são perdidos enquanto o pessoal da manutenção localiza e repara o problema. Isso pressiona ainda mais o processo estatístico para que sempre providencie informações precisas e confiáveis. Antes de decidir interromper um processo, os responsáveis pela qualidade do produto devem estar absolutamente seguros de que realmente existe um problema; por outro lado, se um problema ocorrer na linha de produção, eles não devem deixá-lo continuar por muito tempo antes de fazer os reparos necessários. Essa situação exige que um equilíbrio delicado seja encontrado.

Então, a próxima pergunta é: como podemos determinar quando o processo está “fora de controle”? E como fazer isso sem causar muitos alarmes falsos, custos e desperdício de tempo para a empresa? Assim como a maioria das questões que envolvem a estatística, não existe uma resposta única e definitiva para essa questão (como acontece com a matemática, por exemplo). Algumas pessoas dizem que é isso que elas adoram na estatística e outras dizem que é isso que elas odeiam.

Se os limites de controle forem estabelecidos dentro de 2 erros padrões para mais ou para menos, então, para um processo sob controle, 95% das médias deverão se encontrar dentro desses limites. Porém, isso significa que você deve esperar que 5% dos resultados fiquem de fora desses limites, simplesmente por obra do acaso e isso não representa problema algum! Aqui está o segredo. Você não deseja interromper o processo na primeira vez que uma média sair dos limites de controle; isso poderá acontecer em 5% das vezes, simplesmente por casualidade (veja o Capítulo 10, para mais detalhes sobre esse assunto). Portanto, para evitar tanto falsos alarmes, é necessário que mais do que uma média fique fora dos limites de controle para que você tome a decisão de interromper o processo.

O que você deseja é interromper o processo apenas quando estiver muito seguro de que algo de errado esteja acontecendo, e uma única amostra fora dos limites não é nada extraordinário. Agora, e se você visse 2, 3, 4, 5 ou mais resultados de uma só vez ficarem fora de seus limites pré-estabelecidos? Onde está o limite de corte? Bem-vindo ao maravilhoso mundo da inexatidão da estatística! Veja a seguir quatro exemplos de regras geralmente usadas para determinar se um processo está ou não fora de controle e se deve ou não ser interrompido:

- Cinco médias amostrais seguidas ficaram todas acima ou abaixo do valor alvo (como na Figura 19-5a). Suspeita: preenchimento sistemático em excesso ou preenchimento insuficiente em virtude de problemas no processo.

- Seis médias amostrais seguidas aumentaram ou diminuíram constantemente (como na Figura 19-5b). Suspeita: os produtos que estão saindo da linha estão lentamente ficando cada vez mais longe do valor médio pretendido, provavelmente em virtude de problemas com uma ou mais máquinas.

- Quatorze amostras seguidas se alternaram entre acima e abaixo do valor alvo (como na Figura 19-5c). Suspeita: dois operadores, máquinas ou abastecedores diferentes estão utilizando o mesmo sistema, mas não estão em acordo.

- Quinze médias amostrais seguidas estão dentro de 1 erro padrão em relação ao alvo (como na Figura 19-5d). Suspeita: o processo está mais consistente do que o exigido pelas especificações. (se essa super consistência do processo custar muito tempo e dinheiro, ela deve ser amenizada. Se a super consistência do processo não estiver trazendo prejuízos financeiros nem desperdiçando tempo, descobrir o porquê de o processo ter se alterado – e, futuramente, repita essa alteração – pode valer mais a pena).

Essas regras baseiam-se em probabilidade; você deve interromper o processo quando as chances de o processo ainda estar sob controle forem muito pequenas, segundos os dados obtidos. Observe que as chances de que qualquer uma das médias amostrais enquadre-se acima ou abaixo do alvo é de 50% ou 0,5. Sendo assim, para a primeira regra listada acima, a probabilidade de conseguir cinco médias amostrais seguidas que estejam todas do mesmo lado do alvo é $(0{,}5) \times (0{,}5) \times (0{,}5) \times (0{,}5) \times (0{,}5) = 0{,}03 = 3\%$. Esse valor é menor do que o limite de corte comumente usado de 5% (veja o Capítulo 14). Sendo assim, você conclui que o processo não está sob controle. As chances de que o processo ainda esteja sob controle, segundo os dados, são muito baixas.

Da próxima vez em que você abrir um tubo de pasta de dente, pense em toda a estatística usada para assegurar que ele tenha sido enchido com qualidade.

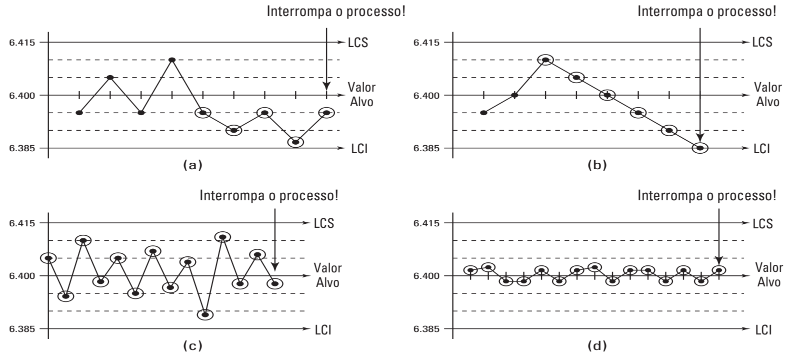


Figura 19-5: Processo de enchimento de tubos de pasta de dente que estão fora de controle.

Parte VIII
A Parte dos Dez

A 5ª Onda por Rich Tennant

Nesta Parte...

O que seria de um livro sobre estatística sem as suas próprias estatísticas? Esta parte contém dez critérios para uma boa pesquisa de opinião e os dez erros estatísticos mais comuns.

Nesta parte você também encontrará uma referência prática e concisa que pode ser usada para ajudar-lhe a avaliar ou projetar uma pesquisa e detectar os abusos estatísticos mais praticados.

Capítulo 20

Os Dez Critérios para uma Boa Pesquisa de Opinião

Neste Capítulo

- Avaliando criticamente uma pesquisa de opinião
- Planejando uma boa pesquisa

As pesquisas de opinião estão ao seu redor: garanto que em algum momento de sua vida, você será convidado a participar de uma. Isso significa que você também está sendo bombardeado com os resultados dessas pesquisas e, antes de consumir essa informação, é necessário avaliar se a pesquisa ou enquete foi projetada e executada corretamente – ou seja, não suponha que uma pesquisa é boa até que você realmente consiga provar (veja o Capítulo 16, para mais informações sobre o assunto). Os dois objetivos mais importantes de uma pesquisa são a precisão (ou seja, basear-se em dados suficientes para que os resultados não se alterem muito caso outra amostra seja coletada) e a obtenção do mínimo possível de viés (a super ou a subestimação sistemática do resultado verdadeiro, como a balança do banheiro que insiste em colocar uns quilinhos a mais!). Neste capítulo, você conhecerá dez critérios para avaliar ou planejar uma pesquisa.

A População Alvo Deve ser Bem Definida

A população alvo é todo o grupo de indivíduos pelos quais você se interessa. Por exemplo, suponha que você queira saber o que as pessoas na Grã-Bretanha acham dos reality shows. A população alvo neste caso seria todos os habitantes da Grã-Bretanha.

LEMBRE-SE

Observe que, às vezes, a população alvo precisa de um pouco de refinamento para uma maior clareza. Por exemplo, que faixas etárias você gostaria de incluir em sua população alvo? Para o caso dos reality shows, provavelmente você não quer incluir crianças abaixo de uma certa idade, digamos, 12 anos. Sendo assim, sua população alvo é, na verdade, toda a população da Grã-Bretanha acima de 12 anos.

Muitos pesquisadores não fazem um bom trabalho ao definir sua população alvo. Por exemplo, se o Egg Board (espécie de associação dos produtores de ovos nos Estados Unidos) quiser dizer "Ovos fazem bem à sua saúde!", eles devem especificar a saúde de quem eles estão se referindo. Por exemplo, será que eles estariam preparados a dizer que os ovos fazem bem à saúde de pessoas com o colesterol alto? (um dos estudos realizados por esse grupo baseava-se apenas em pessoas jovens e saudáveis que seguiam uma dieta de pouca gordura – é da saúde dessas pessoas que eles se referiam?).

Se a população alvo não for bem definida, os resultados da pesquisa provavelmente serão enviesados. Isso ocorre quando a amostra estudada contém pessoas que não fazem parte da população pretendida ou quando a pesquisa exclui alguns indivíduos que deveriam ser incluídos.

A Amostra Deve Retratar a População Alvo

Quando você conduz uma pesquisa, normalmente não conseguirá entrevistar todos os membros da população alvo, a fim de obter as informações pela quais está procurando. Em geral, você não tem tempo nem dinheiro necessários para desempenhar essa tarefa. O melhor que pode ser feito, então, é coletar uma amostra (um subconjunto dos indivíduos de uma população) e obter as informações a partir dela. Uma vez que essa amostra será seu único vínculo com toda a população alvo, ela deve ser realmente boa.

Uma boa amostra representa a população alvo. A amostra não deve sistematicamente favorecer certo grupo da população nem excluir outro. Parece fácil, certo? Tudo o que você precisa fazer é pegar uma lista de todos os indivíduos da população alvo (chamada de painel amostral) e selecionar uma amostra a partir dela. O que poderia dar errado?

Muita coisa. Suponha que sua população alvo seja todos os eleitores registrados nos Estados Unidos que, provavelmente, irão votar na próxima eleição presidencial. Conseguir a lista desses eleitores não é fácil. Você pode consultar as listas de registro de eleitores, mas, ainda assim, não saberá quais pessoas provavelmente votarão na próxima eleição. Você pode consultar os que votaram na eleição passada, mas muitas dessas pessoas mudaram-se ou morreram e, além disso, dessa forma, você não incluirá as pessoas que completaram 18 anos desde a última eleição. De repente, a situação fica um pouco complicada. Bem-vindo ao mundo das pesquisas de opinião!

Para resolver esse problema, você poderia obter as listas atualizadas de registro de eleitores, retirar uma amostra de indivíduos dessas listas e perguntar-lhes se eles pretendem ou não votar na próxima eleição. Quando alguém responder que não, interrompa o questionário e não conte aquela pessoa em sua pesquisa. Para os que realmente pretendem votar, pergunte para quem eles pretendem votar e inclua essas respostas nos resultados de sua pesquisa.

Uma boa pesquisa deve ter um painel de amostras atualizado e que liste todos os membros da população alvo, quando possível. Se obter uma lista como essa não for possível, então algum mecanismo deve ser empregado a fim de assegurar que todos os membros da população tenham a mesma oportunidade de serem selecionados para a pesquisa. Por exemplo, se uma pesquisa de casa em casa será realizada em uma cidade, é necessário que se utilize um mapa atualizado que inclua todas as casas da cidade em questão.

A Amostra Deve ser Selecionada Aleatoriamente

Uma característica importante de uma boa pesquisa é que a amostra da população alvo seja selecionada aleatoriamente. Aleatoriamente significa que todos os membros da população alvo devem ter as mesmas chances de serem incluídos na amostra. Ou seja, o processo usado para a seleção de sua amostra não pode ser parcial.

Suponha que você tenha um rebanho com 1.000 novilhos e precisa retirar uma amostra aleatória de 50 deles para fazer um exame para uma doença. Retirar os 50 primeiros novilhos que vierem em sua direção não se encaixaria na definição de amostra aleatória. Os primeiros novilhos que forem capazes de vir em sua direção, provavelmente são os que têm menos chances de apresentarem qualquer tipo de doença ou, talvez, sejam os mais velhos e mais amigáveis, que realmente são os mais suscetíveis a doenças. De qualquer forma, a pesquisa foi contaminada pelo viés. Como coletar uma amostra aleatória dos novilhos? Os animais provavelmente possuem etiquetas com um número de identificação, assim você deve conseguir uma lista com todos os números de identificação, selecione uma amostra aleatória desses números e localize os animais. Ou, se os animais ficam em gaiolas ou estábulos, numere esses locais e retire uma amostra aleatória desses números. Às vezes, para ser um estatístico é necessário usar muita criatividade de tal maneira a obter uma amostra verdadeiramente aleatória!

Para as pesquisas que envolvem pessoas, organizações de renome na área de pesquisa, como o Instituto Gallup, usam um procedimento de discagem digital e telefonam para os membros de sua amostra. Isso, é claro, acaba excluindo as pessoas que não possuem telefone e, por isso, essas pesquisas possuem um pequeno viés. Nesse caso, no entanto, a maioria das pessoas possui uma linha telefônica (mais de 95%, de acordo com o Instituto Gallup) e, portanto, o viés causado pelo fato de excluir as pessoas que não possuem telefone acaba não sendo um problema tão grande.

LEMBRE-SE

Uma boa pesquisa possui uma amostra aleatória de indivíduos da população alvo. Procure saber como a amostra foi selecionada, caso o processo não seja mencionado.

O Tamanho da Amostra Deve ser Grande o Bastante

Você já ouviu a expressão que diz,:"Quanto menos, melhor"? Com as pesquisas de opinião, a expressão é: "Menos de boa informação é melhor do que muito de má informação, mas, quanto mais boas informações você tiver, melhor" (não é muito prático, não é mesmo?)

Eis a ideia principal: se você tem um tamanho amostral grande e a amostra realmente representa a população alvo (e tiver sido selecionada aleatoriamente), você pode contar que a informação obtida será realmente muito precisa. A precisão depende do tamanho da amostra, e quanto maior é o tamanho amostral, maior será a precisão das informações.

Uma maneira rápida e prática de calcular a precisão de uma pesquisa é dividir o número 1 pela raiz quadrada do tamanho amostral. Por exemplo, a precisão de uma pesquisa com 1.000 pessoas (selecionadas aleatoriamente) fica dentro de $\frac{1}{\sqrt{1000}}$, que é 0,032 ou 32%. Essa porcentagem é a chamada margem de erro.

Tome cuidado com pesquisas que têm amostras grandes que não são selecionadas aleatoriamente. As pesquisas feitas pela Internet são os maiores exemplos. Uma empresa pode dizer que 50.000 acessaram seu site para responder à pesquisa, o que significa que os resultados obtidos baseiam-se em muita informação. No entanto, essa informação está enviesada, pois não representa as opiniões de ninguém além da opinião das pessoas que optaram por participar da pesquisa; ou seja, das pessoas que acessaram a Internet, visitaram o site e decidiram completar o questionário da pesquisa. Nesse caso, o pouco teria sido muito. A empresa deveria ter uma amostra menor, mas selecionada aleatoriamente.

A Insistência Minimiza a Falta de Respostas

Depois que o tamanho amostral tiver sido escolhido e os indivíduos da população alvo tiverem sido selecionados aleatoriamente, você tem que obter a informação de que precisa a partir das pessoas da amostra. Se você alguma vez já jogou fora uma pesquisa ou recusou-se a responder a algumas perguntas por telefone, você deve saber que não é nada fácil conseguir pessoas para participar de uma pesquisa. Se o pesquisador quiser minimizar o viés, a melhor maneira de lidar com a falta de respostas é "ficando na cola" das pessoas da amostra: insista uma, duas ou até três vezes, ofereça dinheiro, cupons, cartas respostas, chances de ganhar prêmios, entre outras coisas. Observe que se você oferecer mais do que algumas moedas de incentivo em agradecimento pela participação na pesquisa, você também poderá causar viés, pois, assim, as pessoas que realmente estiverem precisando de dinheiro estarão mais propensas a responder a pesquisa do que as que não estiverem.

Pense no que lhe faz responder a uma pesquisa. Se o incentivo dado pelo pesquisador não for atraente, talvez o assunto abordado seja de seu interesse. Infelizmente, é ai que o viés entra em cena. Se apenas as pessoas que se interessam pelo assunto respondem à pesquisa, apenas suas opiniões serão consideradas; pois as pessoas que não se importam com o assunto não responderão e, sendo assim, todos os votos de "eu não me importo" não serão contabilizados. O mesmo acontecerá com os votos das pessoas que realmente se importam com o assunto, mas não tiveram tempo para responder a pesquisa.

Por exemplo, suponha que 1.000 pessoas tenham recebido o questionário de uma pesquisa a respeito da alteração das regras dos parques para permitir a presença de cães. Quem irá responder? Muito provavelmente, os respondentes serão os que são totalmente a favor ou totalmente contra a presença de cães nos parques. Suponha que apenas 100 pessoas de cada lado responderam a pesquisa e 800 pessoas não devolveram o questionário. Isso significa que 800 opiniões não serão contabilizadas. Se nenhuma das 800 pessoas realmente se importasse com o assunto abordado, e se você pudesse contabilizar suas opiniões, os resultados seriam informados da seguinte forma: 800 ÷ 1.000 ou 80% "não opinou", 10% (100 ÷ 1.000) disseram ser a favor das mudanças e 10% (100 ÷ 1.000) foram contra as mudanças. No entanto, sem os votos dessas 800 pessoas, os pesquisadores só poderão dizer: "Das pessoas que responderam a pesquisa, 50% foram a favor e 50% foram contra o assunto abordado." Isso nos dá a impressão de um resultado muito diferente (e muito enviesado!) do esperado.

A taxa de resposta de uma pesquisa é a porcentagem encontrada ao dividir o número de respondentes pelo tamanho amostral total e multiplicar o resultado por 100%. Uma boa taxa de resposta deve ficar acima de 70%. Entretanto, a maioria das taxas de resposta encontra-se bem abaixo desse número, com exceção das que são realizados por organizações de renome, como o Instituo Gallup. Procure a taxa de resposta quando for examinar os resultados de uma pesquisa. Caso ela seja muito baixa (muito menor do que 70%) os resultados podem estar enviesados e devem ser ignorados.

Selecionar um tamanho amostral inicial pequeno e insistir persistentemente para que as pessoas respondam é melhor do que inicialmente selecionar uma amostra grande e acabar com uma taxa de resposta muito baixa. A insistência contínua reduz o viés.

Da próxima vez que você for convidado a participar de uma boa pesquisa (segundo os critérios listados neste capítulo), considere a possibilidade de responder. Assim, você estará fazendo sua parte para banir o viés do mundo!

O Tipo de Pesquisa Usado Deve ser Apropriado

As pesquisas podem ser feitas de várias formas: por correspondência, por telefone, pela Internet, de porta em porta e nas ruas (quando alguém segurando uma prancheta lhe para no meio da rua e pergunta se você tem tempo para responder a algumas perguntas). Um critério muito importante, mas que às vezes é deixado de lado, é se o tipo de pesquisa usado é apropriado para a situação.

Por exemplo, se a população alvo é formada por pessoas que possuem problemas de visão, enviar-lhes, por correio, um questionário escrito em uma fonte minúscula não é uma boa opção (sim, isso já aconteceu). Se você deseja realizar uma pesquisa com vítimas de violência doméstica, entrevistá-las em suas casas também não é apropriado.

Suponha que sua população alvo seja os moradores de rua de sua cidade. Como você os encontrará? Eles não possuem endereço ou telefone, por isso nenhum tipo de pesquisa é apropriado (essa é uma questão muito complexa e uma das quais os governos precisam enfrentar quando chega a hora do censo). Você pode sair e conversar pessoalmente com cada um, onde quer que eles estejam, mas encontrá-los não é tarefa fácil. Perguntar a abrigos locais, igrejas ou qualquer outro grupo que ajude os moradores de rua pode dar boas pistas de onde começar sua busca.

Ao examinar os resultados de uma pesquisa, procure saber o tipo de pesquisa usado e veja se ele foi apropriado à situação.

As Perguntas Devem ser Bem Escritas

A maneira como a pergunta de uma pesquisa é escrita pode afetar os resultados. Por exemplo, na época em que Bill Clinton era presidente dos Estados Unidos e aconteceu o escândalo com a estagiária Monica Lewinsky, uma enquete realizada pela CNN/Instituto Gallup, entre os dias 21 e 23 de agosto de 1998, pedia que os entrevistados julgassem a favorabilidade de Clinton, e cerca de 60% deram-lhe um resultado positivo (o tamanho amostral dessa pesquisa foi de 1.317 pessoas e a margem de erro foi de 3 pontos percentuais para mais ou para menos). Quando a CNN/Instituo Gallup reescreveu a pergunta pedindo para os entrevistados julgarem a favorabilidade de Bill Clinton "como pessoa", os resultados mudaram: Apenas 40% dos entrevistados lhe classificaram positivamente.

Na noite seguinte, a CNN/Instituto Gallup realizou outra pesquisa sobre o mesmo assunto. Eis algumas perguntas e respostas:

- Você aprova ou desaprova a maneira como o Presidente Clinton está desempenhando suas funções? (Sessenta e dois por cento aprovaram e 35% desaprovaram).
- Você tem uma opinião favorável ou desfavorável com relação a Clinton? (Cinquenta e quatro por cento disseram ser favorável, 43% disseram ser contra).
- Considerando tudo, você está satisfeito por ter Clinton como presidente? (Cinquenta e seis disseram que sim, 42% disseram que não).
- Se você pudesse votar novamente em um dos candidatos a presidente de 1996, em quem você votaria? (Quarenta e seis por cento disseram que votariam em Bill Clinton, 34% para Bob Dole, 13% para Ross Perot).

Todas essas perguntas eram a respeito do mesmo assunto: a opinião das pessoas em relação ao presidente Bill Clinton no momento em que ocorreu o escândalo com Monica Lewinsky. E ainda que elas sejam semelhantes, todas são escritas de maneira ligeiramente distinta e você pode observar como os resultados são diferentes. Portanto, a maneira de escrever realmente é importante.

Provavelmente, o maior problema com a formulação das perguntas é o uso de perguntas indutivas. Ou seja, perguntas escritas de maneira que você saiba o que o pesquisador quer que você responda. Isso leva a resultados enviesados, que dão muito importância a certas resposta por causa da maneira como a pergunta foi escrita, e não por ser realmente a opinião das pessoas.

Muitas pesquisas contêm perguntas indutivas (tanto intencionalmente quando por erro de projeto) para fazer com que você diga o que o entrevistador quer que você diga. Abaixo, veja alguns exemplos de perguntas indutivas semelhantes às que já encontrei na imprensa:

- Qual posição assemelha-se mais a sua maneira de pensar? A dos democratas, que são a favor de um plano fiscal realista e responsável para equilibrar o orçamento por um período razoável de tempo, não deixando de cumprir com suas responsabilidades com os americanos mais vulneráveis, ou a dos Republicanos, que propõe a execução de um equilíbrio orçamentário obrigatório, fazendo cortes severos na educação e na saúde.

- Você acha que o presidente deveria ter o poder de veto individual sobre determinados itens do orçamento para eliminar o desperdício?

- Você acha que os oficiais do Congresso e da Casa Branca deveriam dar um bom exemplo e desfazerem-se dos benefícios e dos privilégios especiais que atualmente possuem em favor de cada contribuinte?

Ao ver os resultados de uma pesquisa importante para você, peça uma cópia das perguntas que foram feitas e as analise para você ter certeza de que foram feitas de maneira neutra e que tentaram minimizar o viés.

O Momento da Pesquisa Também Deve ser Apropriado

O momento da pesquisa é fundamental. Os eventos atuais modelam a opinião das pessoas e, enquanto alguns pesquisadores tentam determinar o que a população realmente pensa sobre determinado assunto, outros tentam tirar proveito dessas situações, especialmente as negativas. Por exemplo, quando tiroteios aconteceram em escolas, a questão do controle de armas foi levantada pelas enquetes e pesquisas. É claro que logo depois da tragédia, mais pessoas são a favor do controle de armas do que eram antes, ou seja, os resultados aumentam vertiginosamente. Depois de um tempo, entretanto, as opiniões voltam a ser o que eram antes; enquanto isso, os pesquisadores projetam seus resultados como se eles refletissem a opinião pública real.

O momento para se realizar qualquer pesquisa, independente de seu assunto, também pode causar viés. Por exemplo, imagine que sua população alvo seja as pessoas que trabalham período integral. Se você realizar uma pesquisa por telefone, ligando para suas casas entre 9 da manhã e 5 da tarde, você terá resultados muito enviesados, pois esse período compreende o horário de expediente da maior parte das pessoas que trabalham período integral!

Verifique a data em que a pesquisa foi realizada e veja se você pode determinar algum evento relevante que possa ter temporariamente influenciado os resultados. Também procure saber o horário do dia ou da noite em que a pesquisa foi realizada: ela foi conduzida durante o horário mais conveniente para a população alvo?

O Treinamento das Pessoas que Realizam as Entrevistas

As pessoas que realizam as entrevistas para as pesquisas têm um trabalho duro pela frente. Elas têm que lidar com pessoas que desligam o telefone antes que elas possam dizer algo, com respostas do tipo "tire nosso nome da lista" e com as secretárias eletrônicas. E quando elas conseguem alguém para responder, seu trabalho fica ainda mais difícil. É aí que elas têm que coletar os dados da maneira mais precisa e imparcial possível.

Eis alguns problemas que podem surgir durante o processo de entrevista:

- O participante não entende a pergunta e precisa de mais informações. Quanto você pode falar para aquela pessoa e ainda permanecer neutro?
- A informação pode ser mal transcrita. Por exemplo, eu respondo que tenho 40 anos, mas a pessoa que está fazendo a pesquisa acidentalmente escreve 60.
- O entrevistador precisa tomar uma decisão. Por exemplo, suponha que a pergunta seja quantas pessoas estão morando em sua casa e o participante pergunte: "Devo contar meu primo Bob que está morando com a gente até conseguir um emprego?". Uma decisão tem que ser tomada.
- Os participantes podem dar informações falsas. Por exemplo, algumas pessoas odeiam tanto participar de pesquisas que vão além de simplesmente recusá-las. Pelo contrário, elas preenchem os questionários, mas fornecem respostas que podem levar ao erro. Por exemplo, uma mulher pode responder que tem 101 anos para a pergunta sobre sua idade.

Como os entrevistadores devem lidar com esses e outros vários desafios que ocorrem durante o processo de pesquisa? O segredo é ser claro e consistente sobre toda possibilidade que possa surgir e discutir como elas podem ser contornadas bem antes de começar a entrar em contato com os participantes. Isso significa que os entrevistadores precisam ser bem treinados.

DICA

Você também pode evitar problemas se conduzir um estudo piloto (um teste prático com apenas alguns participantes), que é gravado para permitir que os entrevistadores possam praticar e serem avaliados com relação à precisão e consistência durante a coleta de dados. Dessa forma, os pesquisadores podem antecipar problemas antes do início do processo de pesquisa e implantar medidas para melhor lidar com eles.

As pesquisas devem evitar perguntas ambíguas ou indutivas. O estudo piloto também pode rastrear as possíveis dificuldades relacionadas às perguntas ao avaliar como um pequeno grupo de participantes responde às questões apresentadas e quais perguntas eles fazem. E, para evitar a má transcrição da informação (ou seja, a digitação incorreta da informação), a pesquisa precisa ter todas as possíveis respostas marcadas de maneira clara. Por exemplo, se 1 representar a opinião de "totalmente contrário", então a posição de "totalmente a favor" também deve ser representada por um número (e não por outro tipo de código). Gravar as entrevistas para fazer uma verificação cruzada também pode ser útil.

Decida como lidar com os respondentes brincalhões, aqueles que estão simplesmente brincando e não fornecem respostas confiáveis. Uma sugestão é sinalizar a resposta daquela pessoa como sendo possivelmente descartável e, depois, telefonar mais tarde e tentar de novo.

A Pesquisa Deve Responder à Pergunta Original

Suponha que um pesquisador faça a seguinte declaração: "Quero descobrir os hábitos de compras dos consumidores". Até aí tudo bem, certo? Mas então, você vê a pesquisa e todas as perguntas são direcionadas a saber como as pessoas se sentem em relação ao ato de fazer compras ("O que você mais gosta/menos gosta com relação a fazer compras?", ou "De 1 a 10, qual nota você daria para o quanto você gosta de fazer compras?"). Essas questões não indagam o comportamento das pessoas na hora de fazer compras. Ainda que a posição em relação ao ato de fazer compras realmente influencie a maneira como as pessoas fazem suas compras, a medida real dos hábitos de compra é como os consumidores se comportam quando têm que fazer compras: o que costumam comprar, quanto gastam, onde compraram, com quem vão às compras, quando vão às compras, com que frequência vão às compras e assim por diante. Às vezes, os pesquisadores não se dão conta de que perderam o fio da meada, até mesmo depois que os resultados da pesquisas tenham ficado prontos. Depois do fato consumado (e muito tarde para se fazer qualquer coisa), eles veem que não conseguiram responder as perguntas de seu estudo usando os dados que coletaram, o que não é nada bom!

Antes de participar de uma pesquisa, pergunte ao entrevistador o que ele está tentando descobrir – qual é o propósito da pesquisa. Então, conforme você for lendo ou escutando as perguntas, se elas parecerem estar lhe levando a uma direção contrária, eu aconselho você a parar de responder e explicar o motivo, seja escrevendo ou pessoalmente.

Antes de projetar qualquer pesquisa, primeiro, escreva os objetivos da pesquisa. O que você quer saber? Depois, formule as perguntas para chegar a seus objetivos. Dessa forma, você terá certeza de que conseguirá ter suas perguntas respondidas.

Capítulo 21

Os Dez Erros Estatísticos mais Comuns

Neste Capítulo

- Reconhecendo os erros estatísticos mais comuns cometidos por pesquisadores e pela mídia
- Evitando erros ao fazer suas estatísticas

Este livro não fala apenas de como entender as estatísticas com as quais você se depara na mídia e em seu local de trabalho, mas fala também sobre como ir mais a fundo para examinar se essas estatísticas são realmente corretas, racionais e justas. Você tem que ser vigilante – e um pouco cético – para lidar com a explosão de informação do mundo de hoje, pois muitas das estatísticas com as quais você se depara são enganosas e estão incorretas, por acidente ou intencionalmente. Se você não avaliar criticamente a informação que está consumindo, quem fará isso por você? Neste capítulo, traço alguns erros estatísticos mais comuns cometidos por pesquisadores ou pela mídia e compartilho as formas de reconhecê-los e evitá-los.

Gráficos enganosos

Muitos gráficos e tabelas contêm informação ruim, enganosa e mentirosa ou, simplesmente, omitem informações importantes para que o leitor tome decisões relevantes a respeito do que está sendo apresentado. A Figura 21-1 mostra exemplos de quatro tipos de recursos visuais para a exibição de dados: gráficos pizza, gráficos de barras, gráficos de linhas e histogramas (observe que o histograma é, basicamente, um gráfico de barras para dados numéricos). Para cada tipo, tracei algumas maneiras mais comuns utilizadas para nos enganar (para mais informações sobre gráficos e tabelas, incluindo os gráficos e tabelas enganosos, veja o Capítulo 4).

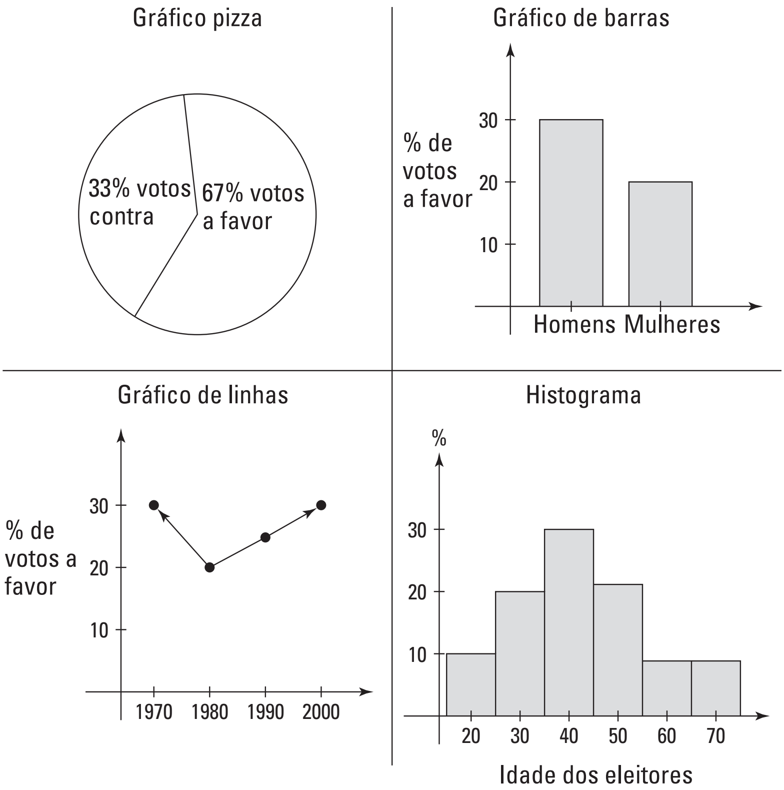


Figura 21-1: Quatro tipos de recursos visuais para a exibição dos dados.

Gráficos pizza

Os gráficos pizza são exatamente o que seu nome sugere: gráficos no formato de um círculo (ou pizza) divididos em fatias que representam a porcentagem de indivíduos que se encontram em cada grupo (de acordo com alguma variável categórica, tal como sexo, partido político ou estado empregatício).

Veja como provar um gráfico pizza e atestar sua qualidade:

- Confira se as porcentagens somam 100% ou chegam o mais próximo possível (qualquer erro de arredondamento deveria ser pequeno).
- Preste atenção nas fatias denominadas "outros" que sejam maiores que outras fatias. Isso significa que o gráfico pizza é muito vago.
- Cuidado com as distorções de ilusão de ótica causadas pelos gráficos pizza tridimensionais, nos quais a fatia mais próxima de você parece ser maior do que realmente é em virtude do ângulo em que é apresentada.
- Procure pelo número total de indivíduos que compõem o gráfico pizza, para que você possa determinar o tamanho da pizza antes

de ser dividida nas fatias que você está vendo. Caso o tamanho do conjunto de dados (o número de respondentes) seja muito pequeno, a informação não é confiável.

Gráficos de barras

O gráfico de barras é semelhante ao gráfico pizza, exceto pelo fato de que, ao invés de ser em formato de círculo e divido em fatias, o gráfico de barras representa cada grupo com um barra e a altura de suas barras é o que representa o número ou a porcentagem de indivíduos em cada grupo.

Ao examinar um gráfico de barras:

- Preste atenção nas unidades representadas pela altura das barras e no que os resultados representam em termos dessas unidades. Por exemplo, o número total de crimes versus a taxa de criminalidade, que é medida pelo número total de crimes per capita (ou seja, por pessoa).

- Avalie a adequabilidade da escala, ou o valor do espaço entre as unidades que expressam o número em cada grupo do gráfico de barra. Por exemplo, o número de reclamações de clientes para homens versus para mulheres poderia ser expresso por unidades de 1, 5, 10 ou 100. As escalas pequenas (por exemplo, ir de 1 a 500 de 10 em 10) fazem com que as diferenças pareçam maiores; as escalas grandes (por exemplo, ir de 1 a 500 de 100 em 100) fazem com que as diferenças pareçam menores.

Gráfico de linhas

O gráfico de linhas mostra como uma medida altera-se ao longo do tempo (por exemplo, os preços de ações, a média da renda familiar e a temperatura média).

Veja algumas questões com as quais você precisa ter cuidado ao analisar um gráfico de linhas:

- Observe tanto a escala do eixo vertical (quantidade) quanto a do eixo horizontal (linha do tempo); os resultados podem parecer mais ou menos drásticos do que realmente são graças a uma simples alteração dessas escalas.

- Leve em consideração as unidades retratadas pelo gráfico e verifique se foram ajustadas para uma comparação ao longo do tempo; por exemplo, os valores monetários estão sendo ajustados de acordo com a inflação?

- Tome cuidado com pessoas que tentam explicar uma tendência sem usar estatísticas adicionais que as fundamentem. Um gráfico de linhas geralmente mostra o que está acontecendo, o porquê de isso estar acontecendo é outra história!

- Fique atento às situações em que o eixo do tempo não está marcado com o mesmo espaçamento. Isso geralmente acontece quando há falta de dados. Por exemplo, o eixo do tempo talvez tenha espaços iguais entre os anos de 1971, 1972, 1975, 1976, 1978, quando, na verdade, deveria mostrar espaços vazios para os anos em que não há dados disponíveis.

Histogramas

O histograma é um gráfico que divide uma amostra em grupos segundo uma variável numérica (tal como idade, altura, peso ou renda) e mostra ou o número de indivíduos (frequência) ou a porcentagem de indivíduos (frequência relativa) que se encontra em cada grupo. Por exemplo, talvez 20% da amostra tivessem 20 anos ou menos; 30% tivessem entre 20 e 40 anos; 45% tivessem entre 40 e 60 anos; e 5% tivessem acima de 60 anos (sendo a idade a variável numérica para esse caso).

Eis alguns itens que devem ser verificados ao examinar um histograma:

- Observe a escala usada no eixo vertical (frequência/frequência relativa), especialmente porque os resultados podem estar exagerados ou diminuídos pelo uso de escalas inapropriadas.

- Veja se as unidades no eixo vertical estão mostrando frequência ou frequências relativas, leve essa informação em consideração quando for examinar as informações.

- Verifique a escala usada para os grupos da variável numérica nos eixos horizontais. Se os grupos se expressarem em intervalos pequenos (por exemplo, 0-2, 2-4 e assim por diante), os dados poderão parecer muito inconstantes. Mas, se os grupos estiverem expressos por grandes intervalos (por exemplo, 0-100, 100-200 e assim por diante), os dados poderão nos dar uma aparência mais suave do que a realidade.

Dados Enviesados

O viés, na estatística, é o resultado de um erro sistemático que ou superestima ou subestima o valor verdadeiro. Por exemplo, se usarmos uma régua para medir plantas, e essa régua estiver alguns centímetros mais curta do que deveria, todos os meus resultados serão enviesados, ou seja, todos estarão sistematicamente abaixo de seus valores verdadeiros.

Eis algumas fontes mais comuns de dados enviesados:

- Instrumentos de medição sistematicamente fora dos padrões. Por exemplo, o radar de um policial apontando que sua velocidade era de 120 km por hora, mas você sabe que estava a apenas 110 km por hora. Ou, então, uma balança que sempre aumenta 2 quilos.

- Participantes influenciados pelo processo de coleta de dados. Por exemplo, um questionário de pesquisa com a pergunta: "Você já discordou do governo?"; irá superestimar a porcentagem de pessoas que não estão satisfeitas com o governo.

- A amostra dos indivíduos não representa a população de interesse. Por exemplo, examinar os hábitos de estudo em uma universidade ao visitar apenas a biblioteca do campus colocará mais ênfase nos que mais estudam (veja mais na seção "Amostras Não Aleatórias", ainda neste capítulo).

- Pesquisadores não objetivos. Por exemplo, suponha que um grupo de pacientes esteja recebendo comprimidos de açúcar e que outro grupo esteja recebendo o medicamento verdadeiro. Os pesquisadores não podem saber quem faz parte de qual grupo, pois se souberem, podem projetar os resultados nos pacientes (como, por exemplo, perguntando se eles estão se sentindo melhor) ou prestar mais atenção nos que estão recendo o medicamento.

Para identificar os dados enviesados, examine a maneira como eles foram coletados. Pergunte sobre a seleção de participantes, como o estudo foi realizado, que perguntas foram usadas, que tratamentos (medicamentos, procedimentos, terapias, etc.) foram dados e quem sabia sobre eles, que instrumentos de medição foram usados e como foram calibrados, etc. Procure por erros sistemáticos ou favoritismo e, caso encontre muitos, ignore os resultados.

Sem Margem de Erro

A palavra "erro" tem uma conotação negativa, como se um erro fosse algo que pudesse sempre ser evitado. Na estatística, esse nem sempre é o caso. Por exemplo, o que os estatísticos chamam de erro amostral sempre ocorrerá quando alguém tenta estimar um valor populacional usando qualquer coisa que não seja a população inteira. O simples ato de selecionar uma amostra a partir de uma população significa que você irá deixar de fora certos indivíduos e, por sua vez, significa que você não irá conseguir o valor populacional exato e preciso. No entanto, não é preciso se preocupar. Lembre-se que a estatística pressupõe que você nunca precise dizer que está certo – você apenas precisa chegar perto. E, se uma amostra for grande e selecionada aleatoriamente, o erro amostral será pequeno.

Para avaliar o resultado estatístico, você precisa medir sua precisão – normalmente por meio da margem de erro. A margem de erro mostra quanto o pesquisador espera que seus resultados variem de uma amostra para a outra (para mais informações acerca da margem de erro, consulte o Capítulo 10). Quando um pesquisador ou a mídia deixa de mencionar a margem de erro, você fica sem saber a precisão dos resultados ou, pior ainda, você, simplesmente, supõe que tudo está correto, quando, em muitos casos, não é o que acontece. Os resultados de pesquisas mostrados na TV não costumavam mencionar a margem de erro, mas, ultimamente, podemos ver essa informação com mais frequência. Ainda assim, muitas revistas, jornais e pesquisas feitas na Internet omitem a margem de erro ou reportam uma margem de erro que acaba não tendo nenhum valor, já que os dados estão enviesados (veja o Capítulo 10).

Ao examinar resultados estatísticos nos quais um número está sendo estimado (por exemplo, a porcentagem de pessoas que apoia o governo) sempre procure pela margem de erro. Caso ela não tenha sido incluída, peça por ela! (Ou,se você receber outras informações pertinentes, calcule você mesmo a margem de erro usando as fórmulas apresentadas no Capítulo 10).

Amostras Não Aleatórias

Dos dados das pesquisas de opinião aos resultados dos estudos médicos, a maioria das estatísticas se baseia em dados coletados a partir de amostras de indivíduos, ao invés do uso de dados obtidos a partir de populações inteiras, em virtude dos custos e do tempo envolvidos. Além disso, não é necessário ter uma amostra muito grande para conseguir uma boa precisão, desde que a amostra coletada realmente represente a população de interesse. Por exemplo, uma pesquisa bem projetada e bem conduzida com 2.500 pessoas possui uma margem de erro de aproximadamente apenas 2% para mais ou para menos (veja o Capítulo 10). Para experimentos com grupos de tratamento e grupos de controle, os estatísticos gostariam de ver pelo menos 30 pessoas em cada grupo, de modo a de obter dados precisos.

Mas, como garantir que a amostra represente a população? A melhor maneira é selecionar os indivíduos da população aleatoriamente. Uma amostra aleatória é o subconjunto de uma população selecionado de tal maneira que cada membro da população tenha chances iguais de ser selecionado (como realizar um sorteio). Em uma amostra aleatória não existe nenhum tipo de favoritismo ou exclusão.

Muitas pesquisas e estudos não se baseiam em amostras aleatórias de indivíduos. Por exemplo, estudos médicos geralmente utilizam voluntários que, obviamente, candidatam-se voluntariamente – eles não são selecionados aleatoriamente. Não seria nada funcional ligar para as pessoas e dizer: “Você foi selecionado aleatoriamente para participar de nosso estudo sobre o sono. Você precisa vir até nosso laboratório e dormir aqui por quatro noites”. Nessa situação, o melhor a ser feito é estudar os voluntários e ver o quanto eles realmente representam a população alvo e, então, relatar os resultados. Você também pode procurar por tipos específicos de voluntários.

As pesquisas de opinião e as enquetes também precisam se basear em indivíduos selecionados aleatoriamente, mas, ao contrário do que acontece com os estudos médicos, nessa situação, fazer isso é muito mais fácil. Ainda assim, muitas pesquisas não usam amostras aleatórias. Por exemplos, aquelas enquetes feitas pela TV, pedindo aos telespectadores que telefonem e deem sua opinião não se baseiam em amostras aleatórias. Essas enquetes não dão aos indivíduos de toda a população as mesmas chances de serem escolhidos (na verdade, nesse caso, as pessoas mesmas se escolhem).

Antes de tomar qualquer decisão a respeito dos resultados estatísticos de uma pesquisa ou estudo, procure ver como a amostra de indivíduos foi selecionada. Se a amostra não foi selecionada aleatoriamente, considere os resultados como sendo um grão de areia.

Omissão do Tamanho Amostral

A quantidade de informação é sempre importante em se tratando da avaliação da precisão de uma estatística. Quanto mais informações são usadas para se obter uma estatística, mais precisa esta será – desde que essas informações não sejam enviesadas, é claro (veja a seção “Dados Enviesados”, já apresentada neste capítulo). O consumidor de informação estatística precisa avaliar a precisão da informação e, para isso, ele precisa saber como a informação foi coletada (consulte o Capítulo 16, para assuntos relacionados a pesquisas de opinião, e o Capítulo 17, para questões relacionadas a experimentos) e quanto de informação foi coletado (isso quer dizer que você precisa saber o tamanho da amostra).

Muitos gráficos e tabelas mostrados pela mídia não mencionam o tamanho amostral. Além disso, você também acaba descobrindo que as manchetes não são exatamente o que parecem ser quando os detalhes de um artigo revelam um tamanho amostral pequeno (reduzindo a confiabilidade dos resultados) ou quando não mencionam nenhuma informação sobre o tamanho da amostra (por exemplo, você provavelmente já viu o comercial que diz: “Quatro em cinco dentistas recomendam esta pasta de dente para seus pacientes”. Mas, e se eles realmente só tiverem entrevistado cinco dentistas?).

Sempre procure saber o tamanho amostral antes de tirar conclusões sobre a informação estatística. Quanto menor a amostra, menos confiável será a informação. Se o tamanho amostral não for mencionado, consiga uma cópia do relatório completo do estudo, entre em contato com o pesquisador ou com o jornalista que escreveu o artigo.

Correlações Mal interpretadas

A definição estatística de correlação é a força e a direção de uma relação entre duas variáveis numéricas. Em outras palavras, o valor pelo qual se espera que uma variável numérica (por exemplo, o peso) aumente ou diminua se outra variável numérica (por exemplo, a altura) aumentar ou diminuir. A correlação é um dos termos estatísticos mais mal compreendidos e incorretamente usados pelos pesquisadores, pela mídia e pelo público em geral. A seguir, veja três pontos muito importantes a respeito da correlação:

- **O termo correlação não pode ser usado para se referir a duas variáveis categóricas, tais como partido político e sexo.** A correlação é aplicada apenas a duas variáveis numéricas, como altura e peso. Portanto, se você ouvir alguém dizer que “parece que o padrão de voto está correlacionado à idade”, saberá que isso está incorreto. O padrão de voto e o sexo do eleitor podem estar associados, mas nunca correlacionados, de acordo com a definição estatística de correlação.

- **A correlação mede a força e a direção de uma relação linear entre duas variáveis numéricas.** Ou seja, se você coletar duas variáveis numéricas (como peso e altura) e colocar todos os pontos em um gráfico, quando existir uma correlação, você poderá desenhar uma linha reta ligando esses pontos (linhas ascendentes ou descendentes) e a linha se ajustará muito bem. Se uma linha não se ajustar, as variáveis não são correlatas. Entretanto, isso não significa que elas não estejam relacionadas. Elas podem ter algum outro tipo de relação; apenas não possuem uma relação linear. Por exemplo, as bactérias multiplicam-se em taxas exponenciais ao longo do tempo (o número de bactérias aumenta de forma cada vez mais rápida) e não em taxas lineares (que representaria um crescimento constante ao longo do tempo).

- **Correlação não necessariamente significa que exista uma relação de causa e efeito.** Por exemplo, suponha que alguém relate que as pessoas que bebem refrigerantes diet têm mais tumores cerebrais do que as que não bebem esse tipo de refrigerante. Se você for uma das pessoas que bebem refrigerantes diet, não entre em pânico ainda. Isso pode ser uma pegadinha da natureza que alguém acabou notando por acaso. No máximo, isso indica que mais pesquisas devam ser realizadas (além de apenas observações), para mostrar que o refrigerante diet realmente causa tumores no cérebro.

Variáveis Confusas

A variável confusa é uma variável que não foi incluída no estudo e que pode influenciar seus resultados, gerando um efeito confuso. Por exemplo, suponha que um pesquisador tente dizer que comer alga ajuda-lhe a ter uma vida mais longa e, ao examinar o estudo, você descobre que o pesquisador utilizou uma amostra de pessoas acima de 100 anos, que regularmente, incluem a alga em sua alimentação. Suponha, também que você leia as entrevistas dessas pessoas e descubra alguns de seus outros segredos para uma vida mais longa (além de comer alga): essas pessoas comem alimentos muito saudáveis, dormem pelo menos 8 horas por dia, bebem bastante água e praticam exercícios físicos todos os dias. Portanto, é a alga que lhes faz viver mais? Talvez, mas não é possível afirmar, porque as variáveis confusas (exercícios físicos, consumo de água, alimentação e hábitos de sono) também poderiam ser a causa de uma vida mais longa.

Um erro muito comum encontrado em pesquisas é a falta de controle das variáveis confusas, o que deixa os resultados abertos a avaliações mais detalhadas (e você deve estar entre os que fazem essas avaliações!). A melhor forma de controlar as variáveis confusas é realizando um experimento bem projetado, que envolva o estabelecimento de dois grupos o mais semelhantes possível, exceto pelo fato de que um dos grupos (denominado grupo de tratamento) irá receber o tratamento e o outro irá receber um placebo (um tratamento falso; esse grupo é denominado grupo controle). Você, então, irá comparar os resultados dos dois grupos, atribuindo quaisquer diferenças significativas ao tratamento (e nada mais, em um mundo ideal).

O estudo com as algas não foi um experimento. Foi um estudo observacional. Nos estudos observacionais, não existe nenhuma forma de controle das variáveis; as pessoas são apenas observadas e a informação registrada.

Sempre que você se deparar com resultados de um estudo que alega mostrar uma relação de causa e efeito ou diferenças significativas entre grupos, verifique se o estudo era um experimento e se as variáveis confusas foram controladas. Se a realização de um experimento não for possível em virtude de motivos éticos (por exemplo, demonstrar que o cigarro causa câncer de pulmão por meio de um estudo que force metade dos indivíduos a fumarem 10 maços de cigarro por dia durante 20 anos, enquanto a outra metade não fuma nada), você terá que confiar nas evidências fornecidas por muitos estudos observacionais que abranjam diversas situações, todas levando ao mesmo resultado.

Os estudos observacionais são ótimos quando se trata de enquetes e pesquisas de opinião, mas não para mostrar relações de causa e efeito, pois eles não controlam as variáveis confusas. Para essa finalidade, um experimento bem projetado fornecerá muito mais evidências.

Números mal calculados

Não é porque uma estatística apareceu na mídia significa que ela esteja correta. Na verdade, erros acontecem a todo o momento (por acidente ou propositalmente), portanto fique atento a eles. Eis algumas dicas para identificar números mal calculados:

- **Certifique-se de que a soma de tudo realmente seja o valor mencionado.** Com os gráficos pizza, por exemplo, veja se todas as porcentagens somam 100%.

- **Verifique duas vezes até mesmo as operações mais básicas.** Por exemplo, um gráfico diz que 83% dos americanos são favoráveis a determinada questão, mas a notícia diz que 7 em cada 8 americanos são a favor da questão. Será que é a mesma coisa? (Não, 7 dividido por 8 é 87,5%, 83% seria 5 em cada 6).

- **Procure saber a taxa de resposta de uma pesquisa; não se contente em apenas saber o número de participantes.** A taxa de resposta é o número de pessoas que responderam a pesquisa dividido pelo numero total de pessoas selecionadas e multiplicado por 100%. Se a taxa de resposta for muito abaixo de 70%, os resultados podem estar enviesados, pois você não sabe o que as pessoas que não responderam teriam dito.

- **Questione o tipo de estatística usada, para determinar sua adequabilidade.** Por exemplo, o número de crimes aumentou, mas o tamanho da população também. Os pesquisadores deveriam ter relatado a taxa de criminalidade (o número de crimes per capita).

Os resultados estatísticos se baseiam em fórmulas e cálculos que não pensam por si só - aqueles que usam esses cálculos é que deveriam pensar; mas, ou eles não usam o cérebro, ou querem (deliberadamente) evitar que você perceba a verdade. Sendo você um consumidor de informação (também conhecido como incrédulo convicto), deve tomar alguma atitude, e o melhor a ser feito, nesse caso, é questionar.

Relatos Seletivos dos Resultados

Outra situação nada boa é quando um pesquisador relata seu único resultado estatisticamente significativo (um resultado que muito provavelmente não aconteceria por acaso), mas omite as centenas de outros resultados aos quais ele chegou e não relatou, pois eram resultados não significativos. Se você soubesse de todos os outros testes, talvez estivesse se perguntando se esse resultado considerado é, realmente, significativo ou, simplesmente, foi em consequência do grande número de tentativas realizadas. Isso é o que os estatísticos gostam de chamar de data snooping ou data fishing.

Como se proteger dos resultados enganosos ocorrido graças ao data snooping? Descubra mais detalhes sobre o estudo; quantos testes foram feitos, quantos resultados não foram significativos e o que se considerou significativo. Em outras palavras, descubra tudo o que conseguir, para que, assim, você possa colocar esses resultados significativos dentro de uma perspectiva.

LEMBRE-SE

Para identificar números forjados e erros de omissão, a melhor coisa a ser feita é lembrar do ditado que diz que quando alguma coisa parece ser muito boa para ser verdade, ela provavelmente não é. Não acredite no primeiro resultado que lhe mostrarem, especialmente se ele alardear grandes descobertas. Espere para ver se outros conseguirão confirmar e repetir os mesmos resultados.

O Poder de um Testemunho

Ah, o testemunho – uma das maneiras mais poderosas de influenciar a opinião pública já criada. E uma das menos estatisticamente corretas. Um testemunho ou anedota é uma história baseada em uma experiência ou situação vivida por uma única pessoa. Por exemplo:

- A garçonete que ganhou na loteria.
- O gato que aprendeu a andar de bicicleta.
- A mulher que perdeu 45 quilos em dois dias com a milagrosa dieta da batata.
- A celebridade que diz ter usado uma coloração para o cabelo que se compra em qualquer farmácia e para a qual ela é a garota propaganda (com certeza!).

Os testemunhos dão grandes notícias; quanto mais sensacionalista, melhor. Mas histórias sensacionalistas são exceções à regra. Elas não acontecem para muita gente.

Você pode estar achando que não está ao alcance dessas histórias sensacionalistas. Mas, e todas aquelas vezes em que você se deixou influenciar pela experiência de uma única pessoa? Seu vizinho adora provedor de Internet que ele assina e, por isso, você vai assinar também.

Seu amigo não teve sorte com determinada marca de carro, então você nem vai se dar ao trabalho de fazer um test-drive na concessionária. Seu pai conhece alguém que morreu em um acidente de carro porque ficou preso ao cinto de segurança, então ele decide nunca mais usar o cinto.

Ainda que não haja problema algum em se tomar algumas decisões baseando-se nesse tipo de histórias, algumas decisões mais importantes deveriam ter apoio em estatística real e em dados reais, que venham de estudos bem projetados e pesquisas cuidadosas.

Um testemunho é um conjunto de dados com tamanho amostral igual a um. Você não possui informação com a qual poderia comparar a história, não tem estatísticas para analisar e nem possíveis explicações ou informação para continuar – apenas uma única história. Não deixe que os testemunhos influenciem você. Ao invés disso, confie em estudos científicos e informações estatísticas baseadas em grandes amostras aleatórias de indivíduos que representem suas populações alvo (e não apenas uma única situação).

A melhor coisa a fazer quando alguém tentar lhe persuadir contando uma historia baseada em fatos reais, responda dizendo: "Mostre-me seus dados!"

Apêndice

Fontes

Este apêndice contém as fontes que usei em meus exemplos em todo o livro. Já que você é um detetive em potencial, talvez queira seguir essas pistas, a fim de conseguir mais informações que lhe levem a decisões mais informadas.

Capítulo 1

Todos os artigos de jornais mencionados foram retirados do *The Cincinnati Enquirer e do The Columbus Dispatch,* de janeiro de 26 de 2003.

Capítulo 2

Pesquisa sobre aquecimento de sobras no microondas. *USA Today*, 6 de setembro de 2001.

Website da goma de mascar Trident: `www.Tridentgum.com/consumer/htmlc0000.html`.

Número de crimes nos Estados Unidos de 1990-1998, retirado dos Relatóris de Crimes do FBI: `www.fbi.gov/ucr.htm`.

Website da loteria do Kansas: `www.kslottery.com`.

Tempo de consulta com pacientes evita processos por imperícia médica: *USA Today*, 19 de fevereiro de 1997.

Pesquisa de opinião de Ross Perot: *Guia de TV*, 21 de março de 1993; para mais informações, entre em contato com o United We Stand America em: `www.uwsa.com`.

Journal of the American Medical Association: `jama.ama-assn.org`.

New England Journal of Medicine: `http:content.nejm.org`.

The Lancet: `www.thelancet.com`.

British Medical Journal: `http://bmj.com`.

Instituto Gallup: `www.gallup.com`.

Capítulo 3

Instituto Gallup: www.gallup.com.

U.S. Bureau Census: www.census.gov.

Zinco para resfriador, a posição do travesseiro e o seu sono, sapatos de salto alto e o conforto dos pés: esses estudos podem ser encontrados em "Healthy Habits – that Aren´t", *Woman's Day*, 11 de fevereiro de 2003.

O cricrilar dos grilos e a temperatura: "Cricket thermometers", *Field & Stream*, julho de 1993, vol 98, número 3, p.21; para dados, veja: "The Songs of the Insects (1949) por George W. Pierce, Harvard University Press, PP. 12-21.

Crimes e polícia, Departamento de Justiça Americano: www.ojp.usdoj.gov/bjs/lawenf.htm.

Sorvete e homicídios (New York City): um bom artigo para começar a investigar sobre esse assunto e outras questões relacionadas é Spellman, B.A, & Mandel, D.R. (2003). Para mais sobre a psicologia do pensamento causal, veja Nadel, L. (Ed). *Encyclopedia of Cognitive Science* (Vol. 1, PP. 461-466)

Capítulo 4

Consumer Expenditure Survey, Bureau of Labor Statistics: www.bls.gov/cex.

Loterias:(Ohio)www.ohiolottery.com;(Florida)www.flalottery.com/lottery/edu/edu.html;(Michigan)www.michigan.gov/lottery;(NewYork)www.nylottery.org.

Divisão dos impostos: www.irs.gov/app/cgi-bin/slices.cgi.

Tendências populacionais/raciais/de trabalho, Departamento Americano do trabalho, Herman Report: Futurework: Trends and Challenges for Work in the 21st Century": www.dol.gov/asp/programs/history/herman/reports/futurework.

Gastos com transporte, Bureau of Transportation Statistics: www.bts.gov/publications/transportation_in_the_united_states/pdf/teconomy.pdf.

Estatísticas de nascimento, Departamento de Saúde Publica e Meio Ambiente do Colorado: www.cdphe.state.co.us/../cohid/birthdata.html.

Serviço de Receita Interna: www.irs.gov.

Estimativas de salários e empregos: www.wa.gov/esd/lmea/occdata/oeswage/Page2067.htm.

Estimativas populacionais por estado, US. Census Bureau: http://eire.census.gov/popest/data/states/tables/ST-EST2002

Capítulo 5

Dados da população Americana, U.S Census Bureau: www.census.gov/population/www.documentation/twps0038.pdf.

Salários dos jogadores da NBA:www.hoopsworld.com/article_21.shtml.

Fazendo amigos virtuais: Parks e Floyd (1996), *Journal of Communication* 46 (1), 80-97.

Renda familiar, U.S Census Bureau, *Money Income in the United States* 2001, PP. 26-27.

Capítulo 6

American Community Survey, Columbus OH 2001: www.census.gov/acs/ www/Products/Profiles/Single/2001/ACS/Narrative/155/NP15500US39180000049.htm.

U.S Census Bureau: www.census.gov.

Capítulo 7

Loteria Powerball de Connecticut: www.ctlottery.org/powerball.htm.

Capítulo 8

Nenhuma fonte foi utilizada nesse capítulo

Capítulo 9

Erros padrões das pesquisas de gastos de consumidores (Consumer expenditure survey): www.bls.gov/cex/2001/stnderror/age.pdf.

Bureau of Labor Statistics: www.bls.gov.

Notas do ACT e desvios padrões (2002): www.act.org/news/data/02/pdf/t6-7-8.pdf.

Tabela do ACT de 2002: www.act.org/news/data/02/pdf/data.pdf.

Capítulo 10

Instituto Gallup: www.gallup.com

Capítulo 11

Intervalos de confiança da renda familiar média: `www.census.gov/hhes/income/income01/statemhi.html`; relatório completo: `www.census.gov/prod/2002pubs/p60-218.pdf`.

Capítulo 12

Prevalência do uso de drogas entre adolescentes: "Monitoring the Future", financiado pelo Instituto Nacional de Abuso de Drogas e conduzido pela Universidade de Michigan: `http://monitoringthefuture.org/data/2002data-drugs`.

Capítulo 13

Nenhuma fonte foi usada nesse capítulo

Capítulo 14

Varizes: *Woman's Day*, 11 de fevereiro de 2003, página 28.

O sono do bebê: *Woman's Day*, 11 de fevereiro de 2003, p. 120 "And So, to Bed", por Loraine Stern.

Prevalência do uso de drogas entre adolescentes: "Monitoring the Future", financiado pelo Instituto Nacional de Abuso de Drogas e conduzido pela Universidade de Michigan. `http://monitoringthefuture.org/data/2002data-drugs`.

Instituto Gallup: `www.gallup.com`.

Instituto de Segurança na Estrada (testes de colisão): `www.hwysafety.org/default.htm`.

Relatório de Consumo: `www.consumerreports.org/main/home.jsp`.

Instituto Good Housekeeping (selo de aprovação): http://magazines.ivillage.com/goodhousekeeping.

Capítulo 15

Dr. Ruth: *Family Circle*, 1/2/97 página 102. "Full Circle"

Adderall: Woman's Day, 11 de fevereiro de 2003. Propaganda da página seguinte 110 (Shire U.S. Inc).

Capítulo 16

Instituto Gallup: www.gallup.com.

Guia dos seguidores de Enquetes do Instituto Gallup: www.gallup.com/Publications/pollwatcher.asp

The Harris Poll: http://vr.harrispollonline.com/register.

Zogby International: www.zogby.com.

Associação Médica Americana: www.ama-assn.org.

National Rifle Association (Associação Nacional do Rifle): www.nra.org.

U.S. Census Bureau: www.census.gov.

CBS celebrity activism survey: www.cbsnews.com/stories/2003/03/06/eveningnews/main543046.shtml.

CBS Internet dating survey: www.cbsnews.com/stories/2003/02/17/opinion/polls/main540870.shtml.

CBS pain survey: www.cbsnews.com/stories/2003/01/28/opinion/polls/main538259.shtml

Health surfing on the Web: www.cnn.com/2000/HEALTH/08/18/watercooler.ap/.

Otimismo do investidor: www.gallup.com/poll/releases/pr030224.asp.

Os piores carros do milênio: http://cartalk.car.com/About/worst-Cars/results5.html.

Pegando carona com motoristas alcolizados: www.reuters.com/newArticle.jhtml?type=healthNews&storyID=2345367.

Cuidado com a saúde infantil: http://my.webmd.com/content/article/60/67060.htm?lastselectedguid=(5FE84E90-BC77-4056-A91C-9531713CA348).

Crime Victimization Survey (Pesquisa Nacional de Vitimização): www.ojp.usdoj.gov/bjs/pub/pdf/cvus01.pdf.

Câncer de Mama: http://my.webmd.com/content/article/36/1728_50583/htm?lastselectedguid={5FE84E90-BC77-4056-A91C-9531713CA348}.

Uso de telefone celular: www.consumerreports.org/main/detailv2.jsp?CONTENT%3C%3Ecnt_id= 23371%FOLDER%3C%3Efolder_id=23051&bmUID=1047262402709.

Crime virtual: www.gocsi.com/press/20020407.html.

Assédio sexual: http://womenissues.about.com/library/blsexharassmentstats.htm.

Contando mentiras: Journal of Alpplied Social Psychology, 1997, v. 27, 12pp. 1048-1062, por Ester Backbier, Johan Hoogstraten e Katharina Meeerum Terwogt-Kouwenhoven.

Capítulo 17

Instituto Nacional do Câncer nos Estados Unidos: http://cancer.gov.

Maconha e Quimioterapia: The New York Times, 16 de setembro de 1975.

Informação sobre ensaios clínicos, Instituto Nacional de Saúde dos Estados Unidos: www.Clinicaltrials.gov.

Placebo HIV: The Manhattan Mercury (Manhattan, KS), 21 de setembro de 1997.

Estudo relacionando mães mais velhas a expectativas de vida mais longas: Kansas City Star, 11 de setembro de 1997.

Prevalência do uso de drogas entre adolescentes: "Monitoring the Future", financiado pelo Instituto Nacional de Abuso de Drogas e conduzido pela Universidade de Michigan: http://monitoringthefuture.org/data/2002data-drugs.

Capítulo 18

Institutos Nacionais da Saúde dos Estados Unidos (NIH): http://www.nih.gov.

O Hábito de Assistir Televisão e Outros Hábitos Sedentários em Relação ao Risco de Obesidade e Diabetes Mellitus Tipo 2 em Mulheres: Journal of the American Medical Association (2003) Volume 289, pp. 1785-1791, (NIH, Comunicado de impressa)

A expressão da raiva inversamente relacionada a ataques cardíacos e derrames: Psychosomatic Medicine (2003), 65(1), pp. 100-110; (NIH, Comunicado de impressa).

Consumo de leve a moderado de bebidas alcoólicas e a redução de ataques cardíacos em homens: www.nih.gov/news/pr/jan2003/niaaa-08.htm, (NIH, Comunicado de impressa).

O Tratamento precoce detém o glaucoma: www.nih.gov/news/pr/oct2002/nei-14.htm, (NIH, Comunicado de impressa).

Usar o zinco para combater o resfriado não é uma boa idéia: "Healthy Habits – that Aren't", Woman's Day, 11 de fevereiro de 2003.

Grilos e a temperatura:

Artigo do Garden Gate, Edição Número 5: www.gardengatemagazine.com/tips/25tip13.html.

Conversor do cricrilar dos grilos: Departamento de Previsão Metrológica, www.srh.noaa.gov/elp/wxcalc/cricketconvert.html.

Dados do cricrilar dos grilos: existem muitos conjuntos de dados diferentes para esse exemplo; a seguir, veja a fonte que a maioria dos estatísticos utiliza: Pierce, George W., The songs of the Insects, (1949), Havard University Press, PP. 12-21. Observação: eu usei apenas um subconjunto desses dados para meu exemplo (apenas por motivos de ilustração).

Para mais sobre grilos e a temperatura (e para ouvir o cricrilar dos grilos enquanto você lê): www.dartmouth.edu/~genchem/0102/spring/6win/cricket.html.

O uso da aspirina previne o desenvolvimento de pólipos em paciente com câncer de cólon: Grupo de Leucemia e Câncer do Instituto Nacional do Câncer, liderado por Electra Paskett, The Ohio State University: www.osu.edu/researchnews/archive/aspirin.htm.

Sorvete e homicídios (New York City): um bom artigo para começar a investigar sobre esse assunto e outras questões relacionadas é Spellman, B.A, & Mandel, D.R. (2003). Para mais sobre a psicologia do pensamento causal, veja Nadel, L. (Ed). Encyclopedia of Cognitive Science (Vol. 1, PP. 461-466)

Capítulo 19

FDA (U.S. Food and Drug Administration), órgão responsável de pela administração de alimentos e medicamentos nos Estados Unidos: www.fda.gov.

Gerenciamento de qualidade total (TQM): www.6sigma.us.

W.Edwards Deming: Out of the Crisis (Centro para Estudos de Engenharia Avançada, Instituto de Tecnologia de Massachusetts, 1986): www.deming.org.

Máquina para envase de tubos de pasta de dente: www.packagingdigest.com/articles/199710/52.html.

Perguntas mais frequente com relação ao envase de bisnagas: www.keyinternational.com/FAQ_Tube_Filling.html#Q4.

Tube Council (Associação da indústria de envase de bisnagas): www.tube.org.

Índice Remissivo

Símbolos

• A •

• B •

• C •

• D •

• E •

• F •

• G •

• H •

• I •

• N •

• O •

• P •