

Capítulo 5

A Lógica da Amostragem do *Survey*

Neste livro, como em muitos outros, o termo *survey* tem sido usado com o sentido implícito de “*survey* por amostragem”, por oposição ao estudo de todos os componentes de uma população ou grupo. Tipicamente, métodos de *survey* são usados para estudar um segmento ou parcela — uma amostra — de uma população, para fazer estimativas sobre a natureza da população total da qual a amostra foi selecionada. Embora a prática de usar amostras seja aceita mais ou menos tacitamente, as razões para amostrar, em geral, não são conhecidas. Consideremos brevemente estas razões, antes de nos voltarmos para a lógica e as técnicas da amostragem.

Por que Fazer Amostragem?

Há duas razões para justificar a amostragem: tempo e custo. Uma entrevista de um *survey* domiciliar abrangente pode exigir de uma a três horas e de \$40 a \$100. A economia de estudar 2.000 pessoas em vez, digamos, 500.000 é óbvia. Portanto, amostrar muitas vezes torna possível um projeto, enquanto a recusa em amostrar acabaria com a pesquisa.

No entanto, a amostragem não deve ser vista como um mal necessário. Em geral não se reconhece que *surveys* por amostragem são muitas vezes mais precisos do que entrevistar todos os componentes de uma população. Este fato aparentemente bizarro resulta de diversas características da logística do *survey* por entrevista.

Primeiro, um grande projeto de entrevistas requer um número enorme de entrevistadores. Pesquisadores caracteristicamente tentam limitar seu pessoal aos melhores entrevistadores disponíveis, e um grande projeto provavelmente exige contratar qualquer pessoa disponível, resultando que a qualidade geral dos entrevistadores é mais baixa do que a habitual. A qualidade dos dados coletados se reduz pela menor qualidade dos entrevistadores. Além disso, uma pesquisa em menor escala permite procedimentos de acompanhamento mais severos, aumentando a taxa de entrevistas completas.

Segundo, já que entrevistar todos os membros de uma população grande requer uma fase de entrevistas longa, fica difícil e mesmo impossível especificar a *época* à qual se referem os dados. Se a pesquisa visar medir o nível de desemprego numa grande cidade, o índice de desemprego fornecido pelos dados do *survey* não se referem à cidade, seja do início, seja do fim das entrevistas. Em vez disso, tem que dar alguma data hipotética ao índice de desemprego, talvez o ponto médio do período de entrevistas. (Pedir aos entrevistados para responder sobre uma data especificada uniforme introduz o problema da recordação imprecisa.) O problema da atribuição do tempo é inerente a qualquer projeto de entrevistas não executado num só momento, e a gravidade do problema cresce com a duração do período de entrevistas. Se entrevistar tomar dez meses — com os índices de desemprego variando neste período —, o índice resultante não terá significado.

Finalmente, os requisitos administrativos de um grande *survey* são muito maiores do que os em geral encontradas pelos pesquisadores por *surveys*. A supervisão, os relatórios, o treinamento e assim por diante são muito mais difíceis num grande *survey*, e a qualidade dos dados coletados pode ser menor do que a obtida num trabalho menor e mais manejável. (Vale a pena observar que o Bureau do Censo nos EUA produz um *survey* por amostragem após o censo decenal para avaliar os dados colhidos na enumeração total.)

Dados por Amostragem São Realmente Exatos?

Apesar da discussão anterior, talvez você ainda sinta algum desconforto quanto à amostragem. Uma vez que é possível uma amostra representar erroneamente a população da qual é obtida, o pesquisador que usa métodos de amostragem enfrenta um perigo inevitável. No entanto, como mostraremos neste capítulo, procedimentos de amostragem bem-estabelecidos

podem tornar aceitável esse perigo. *Surveys* por amostragem podem permitir estimativas muito precisas sobre as populações que retratam. Mas o pesquisador de *surveys* por amostragem deve estar preparado para tolerar uma certa ambigüidade, uma vez que raramente é possível determinar o grau de precisão dos achados de uma amostra.

Pesquisadores de opinião política representam um grupo de pesquisadores por *surveys* a quem é dada a oportunidade de verificar a precisão dos achados das suas amostras. O dia da eleição é o do julgamento final para eles, e suas experiências são instrutivas em relação à questão mais geral da precisão dos *surveys* por amostragem.

A maioria dos críticos dos métodos de *survey* por amostragem lembra a votação do *Literary Digest* de 1936, que previu a vitória de Alfred M. Landon sobre Franklin D. Roosevelt por maioria absoluta. Pesquisando uma amostra de mais de 2 milhões de eleitores pelo correio, o *Digest* previu que Landon derrotaria Roosevelt em quase 15 pontos percentuais. A razão principal deste fracasso foi a *modura de amostragem* usada pelos pesquisadores (discutiremos este conceito depois). A amostra do *Digest* foi tirada de catálogos telefônicos e listas de registros de automóveis, um procedimento de amostragem que foi suficiente nas eleições de 1920, 1924, 1928 e 1932, mas que em 1936 não forneceu um corte representativo dos eleitores americanos. No rastro da Depressão, e no meio do New Deal, números sem precedentes de americanos pobres apareceram para votar, e estas pessoas pobres não estavam adequadamente representadas pelos catálogos telefônicos e listas de registros de automóveis.¹

Em 1936, George Gallup previu corretamente que Roosevelt ganharia um segundo mandato. Os procedimentos de amostragem de Gallup diferiam dos do *Literary Digest*. O Instituto Americano de Opinião Pública de Gallup foi pioneiro no uso da *amostragem por cotas* (descrita posteriormente neste capítulo), que garantia a todos os tipos de eleitores, ricos e pobres igualmente, adequada representação na amostra. Onde a pesquisa do *Digest* falhou em encontrar e interrogar os eleitores pobres — predominantemente Democratas —, a amostragem por cotas de Gallup o fez.

Doze anos depois, Gallup e a maioria dos outros pesquisadores de opinião política passaram pelo constrangimento de prever a vitória de Thomas Dewey sobre Harry Truman. Como observou Goodman Ace, acidamente: "Todos acreditam nas pesquisas de opinião pública. Todos, desde o homem da rua até o presidente Thomas E. Dewey."² Vários fatores conspiraram para

causar o fiasco da pesquisa de 48. Para começar, a maioria dos pesquisadores terminou sua pesquisa cedo demais, apesar de uma tendência constante em direção a Truman no transcorrer da campanha. Grande número dos eleitores que disseram não saber em quem votariam foram predominantemente para Truman. Mais importante, contudo — o fracasso de 48 apontou uma série de falhas inerentes à amostragem por cotas, o mesmo método que fora um avanço considerável sobre os métodos de amostragem do *Literary Digest*. Em 1948, diversos pesquisadores acadêmicos de *survey* estavam experimentando métodos de *amostragem probabilística*. De forma geral, eles tinham muito mais sucesso do que os amostradores por cotas, e a amostragem probabilística continua sendo o método mais respeitado por pesquisadores de *survey*.

Esta breve discussão conta uma história parcial das fases iniciais do *survey* por amostragem nos EUA, talvez à custa da fé do leitor nos métodos destes *survey*. Para pensar, consideremos o currículo mais recente da precisão da pesquisa de opinião política. Em novembro de 1988, George Bush recebeu 54% dos votos para presidente; Michael Dukakis, 46%. Mas os eleitores não tiveram que esperar os resultados das urnas para saber quem havia vencido. Várias pesquisas de boca de urna, entrevistando amostras de eleitores logo após saírem das seções eleitorais, mostraram as seguintes estimativas dos resultados, apresentados na Tabela 5-1.

TABELA 5-1

Resultados das pesquisas de boca de urna: novembro de 1988

	Bush	Dukakis
CBS News / New York Times	54%	46%
ABC News / Washington Post	54%	46%
NBC News / Wall Street Journal	54%	46%
Los Angeles Times	54%	46%

FONTE — Dados tirados de *Public Opinion*, p.25-26, jan./fev. 1989. Os dados do CBS foram refeitos para excluir os "respostas não".

No lugar dos 2 milhões de eleitores avaliados pelo *Literary Digest* em 1936, estas quatro pesquisas de boca de urna amostraram entre 6.000 e 20.000 eleitores para prever a votação dos mais de 90 milhões de americanos que foram às urnas. Os resultados destas pesquisas não foram surpreendentes, apenas confirmaram a tendência relatada nas pesquisas feitas na semana anterior à eleição.

O que as grandes pesquisas tinham a dizer, em geral baseadas em amostras de 2.000 ou menos entrevistados, está apresentado na Tabela 5-2.

TABELA 5-2

Pesquisas de opinião política antes da eleição de 1988

	Bush	Dukakis
CBS News / New York Times (4-5/11/88)	55%	45%
Pesquisa Gallup (3-5/11/88)	56%	44%
USA Today / CNN (3-6/11/88)	55%	45%
Pesquisa Harris (2-5/11/88)	53%	47%
ABC News / Washington Post (2-5/11/88)	55%	45%
NBC News / Wall Street Journal (1-5/11/88)	53%	47%

FONTE — Dados tirados de *Public Opinion*, p.39, nov./dez. 1988. Todos os dados foram refeitos para excluir os "respostas não".

Surveys por amostragem podem ser extremamente precisos. Ao mesmo tempo, devemos conceder que, ainda hoje, muitas vezes não o são. O resto do capítulo apresenta as razões e as regras para obter precisão na amostragem.

Dois Tipos de Métodos de Amostragem

É útil distinguir dois grandes tipos de métodos de amostragem: *amostragem probabilística* e *amostragem não probabilística*. O cerne deste capítulo se dedica à amostragem probabilística, por ser hoje o método mais respeitado e útil. Uma parte menor do capítulo considera os vários tipos de métodos de amostragem não probabilística.

Vamos iniciar com uma discussão sobre a lógica da amostragem de probabilidade, seguida de uma breve taxonomia de conceitos e terminologia de amostragem. Depois nos voltaremos para o conceito de distribuição amostral, que é a base para estimar a precisão dos achados do *survey* por amostragem. Estas discussões teóricas serão seguidas por considerações sobre *molduras de populações* e *de amostragem*, focalizando os problemas práticos de como determinar o grupo alvo do estudo e como começar a selecionar uma amostra. A seguir examinaremos os tipos básicos de desenhos de *surveys*: *amostras aleatórias simples*, *amostras sistemáticas*, *amostras estratificadas* e *amostras por conglomerados* [cluster]. Então

discutiremos e descreveremos brevemente a amostragem não probabilística.

O capítulo termina com breves considerações sobre alguns usos não-*survey* dos métodos de amostragem em áreas como a análise de conteúdo, a observação participante e a análise histórica. Esperamos que você se familiarize o suficiente com a *lógica* do *survey* por amostragem para poder aplicar este conhecimento numa ampla variedade de situações.

Lógica da Amostragem de Probabilidade

A história das pesquisas de opinião políticas mostra que *surveys* por amostragem podem ser bem precisos. Ao mesmo tempo, deve ficar igualmente claro que as amostras devem ser selecionadas cuidadosamente. Consideraremos rapidamente o porquê disso.

As Implicações da Homogeneidade e da Heterogeneidade

Se todos os membros de uma população fossem idênticos uns aos outros em todos os sentidos, não haveria necessidade de procedimentos cuidadosos de amostragem; qualquer amostra seria suficiente. De fato, neste caso extremo de *homogeneidade*, um caso bastaria como amostra para estudar as características de toda a população.

Antes de descartar essa idéia como impossível, lembre-se de que muito da amostragem científica é feita desta forma. Nas ciências físicas, em algumas situações, é seguro fazer tal suposição. O químico que testa certas propriedades do carbono, por exemplo, não precisa passar pela enumeração sofrida de todo o carbono no mundo, para depois selecionar uma amostra probabilística de moléculas de carbono para estudo. Do mesmo modo, o médico ou o clínico que examina o sangue de uma pessoa não precisa retirar todo o sangue dela, para depois selecionar uma amostra probabilística de células sanguíneas. Para fins práticos, qualquer amostra de sangue da pessoa é suficiente.

Mas, se há variação ou heterogeneidade na população estudada, o pesquisador deve usar procedimentos de amostragem mais controlados. Vale a pena assinalar a aplicabilidade maior deste princípio além da pesquisa social. As origens da moderna teoria de amostragem estão na pesquisa agrícola, especialmente no trabalho de R. A. Fisher, cujo nome continua associado a algumas das estatísticas de *survey* mais usadas.

Para nossos fins, é mais importante assinalar a heterogeneidade dos grupos sociais. Como as pessoas diferem de várias formas, cada população humana é composta de indivíduos variados. Uma amostra de indivíduos de uma população deve conter essencialmente a mesma variação existente na população, para permitir descrições úteis dela. A amostragem probabilística é um método eficiente para extrair uma amostra que reflita corretamente a variação existente na população como um todo.

Viés Consciente e Inconsciente de Amostragem

É claro que qualquer um pode selecionar uma amostra de *survey*, mesmo sem treinamento ou cuidado especial. Para selecionar uma amostra de 100 estudantes universitários, uma pessoa pode ir ao campus e entrevistar os alunos que ali encontrar andando. Este tipo de método de amostragem, muitas vezes usado por pesquisadores não treinados, tem problemas muito sérios.

Para começar, há o perigo dos vieses pessoais do pesquisador afetarem a amostra selecionada desta forma; portanto, ela não representaria realmente a população estudantil. Suponha, por exemplo, que você se sente intimidado por estudantes com ar *bippy* e acha que eles ridicularizariam seu trabalho. Portanto, você pode evitar, consciente ou inconscientemente, entrevistar tais pessoas. Por outro lado, você pode sentir que as atitudes de estudantes de aspecto "sério" não seriam relevantes para suas finalidades de pesquisa, evitando entrevistá-los. Mesmo que buscasse entrevistar um grupo "equilibrado" de estudantes, talvez você não soubesse quais as proporções adequadas dos tipos diversos de alunos para tal equilíbrio, ou talvez você não tenha como identificar os diferentes tipos, apenas vendo-os passar.

Mesmo fazendo um esforço de entrevistar cada décimo estudante que entrasse na biblioteca universitária, você não garantiria uma amostra representativa, uma vez que tipos diferentes de alunos visitam a biblioteca com frequências diferentes. Portanto, a amostra sobre-representaria aqueles estudantes que usam a biblioteca com mais frequência.

Representatividade e Probabilidade de Seleção

Amostras de *surveys* devem representar as populações das quais são retiradas, se devem fornecer estimativas úteis quanto às características daquela população. Não necessitam, contudo,

ser representativas em todos os aspectos; representatividade, no sentido que tem para a amostragem, limita-se às características relevantes para os interesses substantivos da pesquisa. (Este ponto ficará mais evidente nas discussões posteriores sobre estratificação.)

Um princípio básico da amostragem probabilística é: *uma amostra será representativa da população da qual foi selecionada se todos os membros da população tiverem oportunidade igual de serem selecionados para a amostra.*³ Amostras com esta qualidade são freqüentemente rotuladas de *amostras MIPSE* (método de igual probabilidade de seleção). Embora discutamos mais tarde variações deste princípio, trata-se do princípio primordial que fundamenta a amostragem probabilística.

Indo além deste princípio básico, devemos entender que amostras — mesmo as amostras MIPSE cuidadosamente selecionadas — raramente, talvez nunca, são *perfeitamente representativas* das populações das quais são retiradas. A despeito dessa limitação, a amostragem probabilística oferece duas vantagens especiais para os pesquisadores.

Primeiro, amostras probabilísticas, embora jamais perfeitamente representativas, são tipicamente *mais representativas* do que outros tipos de amostras, porque os vieses discutidos acima são evitados. Na prática, há maior possibilidade de uma amostra probabilística ser representativa da população da qual é extraída do que uma amostra não probabilística.

Segundo, e mais importante, a teoria da probabilidade permite estimar a precisão ou representatividade da amostra. É concebível que você, por meios totalmente aleatórios, selecione uma amostra que represente quase perfeitamente a população maior. No entanto, é muito mais provável que não o faça, e você não tem como avaliar a probabilidade de ter atingido a representatividade. Por outro lado, a amostra probabilística pode fornecer uma estimativa precisa do seu sucesso ou fracasso. Após apresentar uma breve taxonomia da terminologia da amostragem, examinaremos os meios pelos quais se pode estimar a representatividade da sua amostra.

Conceitos e Terminologia de Amostragem

As seguintes discussões da teoria e prática da amostragem usam vários termos técnicos. Para facilitar a compreensão das discussões, é importante defini-los. Na maior parte, usaremos termos habitualmente usados em manuais sobre amostragem e estatística.

Ao apresentar esta taxonomia sobre terminologia e conceitos de amostragem, quero reconhecer uma dívida para com Leslie Kish e seu excelente manual sobre o *survey* por amostragem.⁴ Embora eu tenha modificado algumas das convenções usadas por Kish, a sua apresentação é a mais importante fonte para a nossa discussão.

Elemento

Um *elemento* é a unidade sobre a qual a informação é coletada, e que serve de base para a análise. Tipicamente, na pesquisa de *survey*, os elementos são pessoas ou certos tipos de pessoas. Entretanto, outros tipos de unidades — famílias, clubes, corporações etc. — podem também constituir os elementos de um *survey*.

Universo

Um *universo* é a agregação teórica e hipotética de todos os elementos definidos num *survey*. Se o americano individual for o elemento do *survey*, então “americanos” serão o universo. Porém, um universo de *survey* é totalmente não especificado quanto a tempo e lugar e é essencialmente um termo inútil.

População

População é a agregação teoricamente especificada de elementos do *survey*. Embora o termo vago “americanos” possa constituir o universo de um *survey*, delinear a população inclui definir o elemento “americano” (por exemplo, cidadania e residência) e o referencial de tempo para a pesquisa (americanos desde quando?) Traduzir o universo “nova-iorquinos adultos” para uma população trabalhável requer especificar a idade, definindo “adulto”, as fronteiras de Nova Iorque etc. Especificar o termo “estudante universitário” inclui levar em conta estudantes de horário integral e parcial, candidatos a diplomas ou não, estudantes dos cursos de graduação e pós-graduação e atributos similares.

Embora se tenha que começar especificando cuidadosamente a população, em geral a licença poética permite que você escreva seu relatório em termos do universo hipotético. Por facilidade de apresentação, mesmo o pesquisador mais consciencioso normalmente fala de “americanos” e não de “cidadãos residentes nos Estados Unidos da América a partir de 12 de novembro de

1989". A orientação principal neste assunto, como na maioria dos outros, é não confundir ou iludir os leitores.

População do survey

Uma *população do survey* é a agregação de elementos da qual é de fato extraída a amostra do *survey*. Lembre-se de que uma população é uma especificação teórica do universo. Na prática, raramente você pode garantir que todos os elementos que se enquadram nas definições teóricas estabelecidas têm efetivamente a chance de serem selecionados para a amostra. Mesmo as listas de elementos compiladas especificamente para fins de amostragem costumam ser algo incompletas. Alguns estudantes são sempre inadvertidamente omitidos das listagens de alunos. Há usuários de telefones que exigem que seus nomes e números não sejam incluídos nos catálogos. Portanto, a população de *survey* é a agregação dos elementos dos quais a amostra é selecionada.

Muitas vezes, pesquisadores decidem limitar suas populações de *survey* mais severamente do que o indicado nos exemplos acima. Firmas de pesquisa de opinião podem limitar suas "amostras nacionais" aos quarenta e oito estados adjacentes, omitindo, por razões práticas, o Alasca e o Havaí. Um pesquisador, ao selecionar uma amostra de professores de psicologia, pode limitar a população do *survey* aos professores de psicologia em departamentos de psicologia, omitindo os que pertencem a outros departamentos. (Num certo sentido, podemos dizer que estes pesquisadores redefiniram seus universos e populações, contanto que tenham tornado estas revisões claras para os leitores.)

Unidade de Amostra

Uma *unidade de amostra* é o elemento ou conjunto de elementos considerados para seleção em alguma etapa da amostragem. Numa amostra simples, de estágio único, as unidades de amostra são o mesmo que os elementos. Mas, em amostras mais complexas, pode-se usar níveis diferentes de unidades de amostra. Por exemplo, um pesquisador pode, primeiro, extrair uma amostra de quarteirões numa cidade, depois uma amostra de residências nesses quarteirões é, enfim, uma amostra de adultos nas residências selecionadas. As unidades de amostra, nestas três etapas de amostragem são, respectivamente, quarteirões, residências e adultos, sendo

que só os últimos são os elementos. De forma mais específica, os termos "unidades primárias de amostra," "unidades secundárias de amostra" e "unidades finais de amostra" devem ser usadas para designar as etapas sucessivas.

Moldura de Amostragem

Uma *moldura de amostragem* é a lista de unidades de amostra da qual a amostra, ou alguma etapa dela, é selecionada. Se uma amostra simples de estudantes for selecionada de uma listagem de estudantes, a listagem seria a moldura de amostragem. Se a unidade primária de amostra para uma amostra populacional complexa for o quarteirão, a lista dos quarteirões é a moldura de amostragem — na forma de livreto impresso, cartão arquivo da IBM, fita magnética ou arquivo em disquete.

No planejamento de amostra de etapa única, a moldura de amostragem é a lista dos elementos compondo a população de *survey*. Na prática, muitas vezes são as molduras de amostragem disponíveis que definem as populações de *survey*, e não o contrário. Frequentemente, você começa com um universo ou talvez uma população em mente para sua pesquisa e depois procura possíveis molduras de amostragem. As molduras disponíveis são examinadas e avaliadas, e você escolhe a moldura que representa a população de *survey* mais apropriada para seus fins.

A relação entre populações e molduras de amostragem é crítica, não tendo recebido atenção suficiente. Uma seção posterior deste capítulo aprofunda a questão.

Unidade de Observação

Uma *unidade de observação*, ou unidade de coleta de dados, é um elemento ou agregação de elementos de que se coleta informação. A unidade de análise e a unidade de observação muitas vezes são o mesmo — a pessoa individual —, mas isso não é necessário. Por exemplo, entrevistar chefes de família (as unidades de observação) para coletar informação sobre todos os seus membros (as unidades de análise).

A tarefa é mais simples quando a unidade de análise e a unidade de observação são o mesmo. Mas, frequentemente, isso não é possível ou viável, obrigando-o a usar alguma engenhosidade para obter dados relevantes de suas unidades de análise, sem de fato observá-las.

Variável

Uma *variável* é um conjunto de características mutuamente excluídas, como sexo, idade, emprego etc. Podemos descrever os elementos de uma população em termos das suas características individuais numa variável. Tipicamente, *surveys* visam descrever a distribuição das características de uma variável numa população. Assim, você pode descrever a distribuição etária de uma população examinando a frequência relativa das diferentes idades dos seus membros.

Repare que uma variável, por definição, deve ter *variação*; se todos os elementos na população têm a mesma característica, esta característica é uma constante na população e não parte de uma variável.

Parâmetro

Um *parâmetro* é a descrição sumária de uma variável numa população. A renda média de todas as famílias numa cidade e a distribuição etária da população da cidade são parâmetros. Parte importante da pesquisa por *surveys* envolve a estimativa dos parâmetros populacionais baseada em observações amostrais.

Estatística

Uma *estatística* é a descrição sumária de uma variável numa *survey* por amostragem. Portanto, a renda média computada a partir de um *survey* por amostragem e a distribuição etária daquela amostra são estatísticas. Usa-se estatísticas de amostras para fazer estimativas sobre parâmetros populacionais.

Erro amostral

O *erro amostral* é discutido com mais detalhes numa seção posterior deste capítulo. Métodos de amostragem probabilística raramente ou nunca fornecem estatísticas exatamente iguais aos parâmetros que buscam estimar. Entretanto, a teoria da probabilidade permite estimar o grau de erro a ser esperado num determinado desenho de amostragem.

Níveis de Confiança e Intervalos de Confiança

Estes termos também são discutidos com mais detalhes numa seção posterior. Computar o erro amostral permite expressar a precisão das suas estatísticas em termos do nível de confiança,

estabelecendo que essas estatísticas estão dentro de um intervalo especificado do parâmetro. Por exemplo, você pode dizer que está "95% confiante" que sua estatística (por exemplo, 50% favorecem o candidato X) está dentro de mais ou menos (+) 5 pontos percentuais do parâmetro populacional. Quando você expande o intervalo de confiança para uma dada estatística, sua "confiança" aumenta; você pode dizer que tem 99,9% de confiança de que sua estatística está dentro de + 7,5 pontos percentuais do parâmetro.

Teoria da Amostragem Probabilística e Distribuição Amostral

Esta seção examina a teoria básica da amostragem probabilística tal como é aplicada ao *survey* por amostragem; consideraremos a lógica da distribuição amostral e erro amostral em relação a uma *variável binomial*, ou seja, uma variável consistindo em duas características.

Teoria de Amostragem Probabilística

A finalidade última do *survey* por amostragem é selecionar um conjunto de elementos de uma população de tal forma que descrições destes elementos (estatísticas) descrevam com precisão a população total da qual foram selecionadas. A amostragem probabilística fornece um método para ampliar a possibilidade de alcançar esta meta, bem como métodos para estimar o grau de sucesso provável.

Seleção aleatória é a chave deste processo. Num processo de seleção aleatória, cada elemento tem chance igual de seleção, independente de qualquer outro evento do processo de seleção. Lançar uma moeda perfeita é o exemplo mais frequentemente citado; tirar uma cara ou uma coroa é independente de seleções anteriores de caras ou coroas. Outro exemplo é jogar um conjunto de dados perfeitos.

Estas imagens de seleção aleatória, contudo, raramente se aplicam diretamente aos métodos de *survey* por amostragem. É bem mais provável você usar tabelas de números aleatórios ou programas de computador que selecionam aleatoriamente as unidades de amostragem. A disponibilidade destes auxílios à pesquisa torna-os um bom começo para nossa discussão sobre a amostragem aleatória.

As razões para usar métodos de seleção aleatória, ou seja, tabelas de números aleatórios ou programas de computador,

são duas. Primeiro, este procedimento funciona como antídoto aos vieses conscientes ou inconscientes do pesquisador. O pesquisador que seleciona casos de forma intuitiva pode muito bem selecionar casos que apóiam expectativas ou hipóteses de pesquisa preexistentes. Segundo, a seleção aleatória dá acesso à teoria da probabilidade, que é a base para as estimativas sobre parâmetros populacionais e as estimativas de erro. Passemos agora ao exame deste segundo aspecto.

Distribuição Binomial da Amostragem

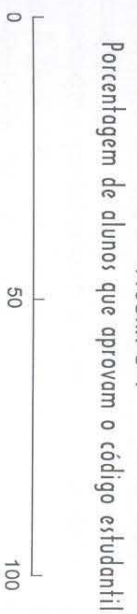
A forma mais clara de discutir o conceito da distribuição amostral é usar um exemplo simples de *survey*. Suponha que queremos estudar a população estudantil de uma universidade para medir a aprovação ou rejeição de um código de conduta estudantil proposto pela administração. A população do *survey* será a agregação de estudantes relacionados numa listagem de alunos (a moldura de amostragem). Os elementos serão os estudantes individuais da universidade. A variável considerada será atitudes quanto ao código, uma variável binomial — aprovar ou rejeitar. Selecionaremos uma amostra aleatória de estudantes para estimar todo o corpo estudantil.

A Figura 5-1 apresenta um eixo de ordenadas representando todos os valores possíveis deste parâmetro na população — de 0 até 100% de aprovação. O ponto médio do eixo — 50% — representa a situação na qual metade dos alunos aprova o código, enquanto a outra metade rejeita-o.

Vamos supor que demos um número a cada aluno da listagem, e que selecionamos 100 números de uma tabela de números aleatórios. Os 100 estudantes cujos números forem selecionados são entrevistados e perguntados quanto à sua atitude em relação ao código estudantil — se aprovam ou rejeitam. Suponhamos que esta operação nos dá 48 alunos que aprovam o código e 52 que rejeitam. Podemos representar esta estatística marcando a ordenada no ponto correspondente a 48%.

Agora suponha que selecionamos outra amostra de 100 alunos exatamente da mesma forma, que medimos sua aprovação/rejeição do código estudantil e que 51 estudantes desta amostra aprovem o código; esse resultado pode ser representado por outra marca no local apropriado da ordenada. Repetindo o processo, podemos descobrir que, numa terceira amostra, 52 alunos aprovam o código.

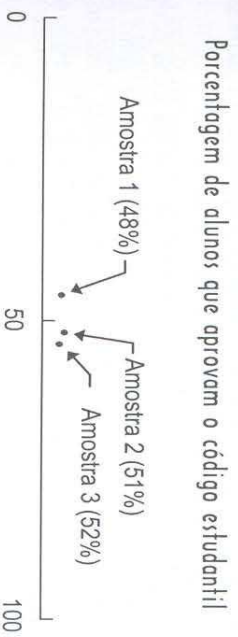
FIGURA 5-1



Porcentagem de alunos que aprovam o código estudantil

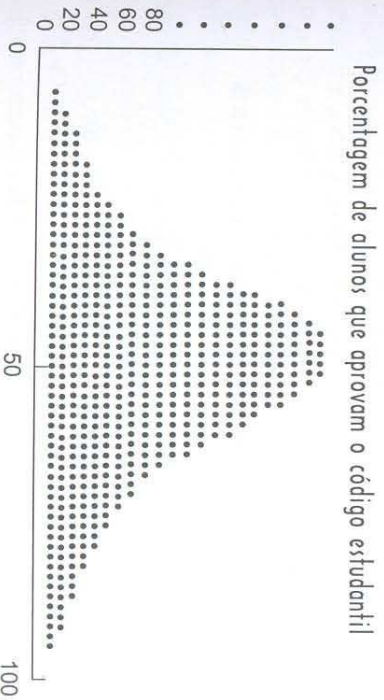
A Figura 5-2 mostra as três diferentes estatísticas representando as porcentagens de alunos em cada uma das três amostras aleatórias que aprovam o código estudantil. A regra básica da amostragem aleatória é que as amostras extraídas de uma população fornecem estimativas do parâmetro relativo à população total. Portanto, cada amostra aleatória dá uma estimativa da porcentagem de estudantes em todo o corpo estudantil que aprovam o código de estudantes. No entanto, infelizmente selecionamos três amostras e agora temos três estimativas distintas.

FIGURA 5-2



Porcentagem de alunos que aprovam o código estudantil

FIGURA 5-3



Porcentagem de alunos que aprovam o código estudantil

Para escapar deste dilema, vamos extrair mais e mais amostras de 100 alunos cada, perguntar a cada uma das amostras sobre sua aprovação/rejeição do código e marcar as novas estatísticas no nosso gráfico de resumo. Ao extrair muitas amostras, vamos descobrir que algumas das novas amostras repetem as estimativas dadas por amostras anteriores. Para explicar esta situação,

acrescentemos um eixo de abcissas à figura, representando o número de amostras que fornecem estimativas. A Figura 5-3 mostra o produto dos nossos novos esforços amostrais.

A distribuição das estatísticas amostrais, mostrada na Figura 5-3, é chamada *distribuição amostral*. Ao aumentar o número de amostras selecionadas e entrevistadas, também aumentamos a extensão das estimativas fornecidas pela operação de amostragem. Num sentido, agravamos nosso dilema de tentar adivinhar o parâmetro da população. No entanto, a teoria da probabilidade oferece certas regras importantes relativas à distribuição de amostras mostrada na Figura 5-3.

Primeiro, se muitas amostras aleatórias independentes são extraídas de uma população, as estatísticas amostrais fornecidas por estas amostras estarão *distribuídas em torno do parâmetro populacional* de uma forma conhecida. Embora haja uma ampla faixa de estimativas, a maioria está mais próxima dos 50% do que de outra parte do gráfico. Portanto, a teoria da probabilidade nos diz que o valor real está próximo aos 50%.

Segundo, a teoria da probabilidade nos dá uma fórmula para estimar *quão próximo* as estatísticas da amostragem estão aglomeradas em torno do valor real. Esta fórmula contém três fatores: o parâmetro, o tamanho da amostra e o *erro padrão* (uma medida do erro da amostragem).

$$\text{Fórmula: } S = \sqrt{\frac{PQ}{n}}$$

Símbolos: P, Q : Os parâmetros populacionais para o binomial; se 60% do corpo estudantil aprovam o código e 40% rejeitam, P e Q são 60% e 40%, ou 0,6 e 0,4.

Observe que $Q = 1 - P$ e $P = 1 - Q$.

n : O número de casos em cada amostra.

S : O erro padrão.

Suponha que o parâmetro populacional seja 50% de aprovação e 50% de rejeição. Recorde que estivemos selecionando amostras de 100 casos cada. Quando estes números são postos na fórmula, vemos que o erro padrão é igual a 0,05, ou 5%.

Em termos da teoria da probabilidade, o erro padrão é um dado valioso porque indica em que medida as estimativas amostrais estão distribuídas em torno do parâmetro populacional. Especificamente, a teoria da probabilidade indica que certas proporções das estimativas amostrais estão dentro de incrementos especificados de erros padrão do parâmetro populacional. Aproximadamente 34% (0,3413) das estimativas amostrais estão até um desvio padrão acima do parâmetro da

população, e outros 34% até um desvio padrão abaixo do parâmetro. No nosso exemplo, o erro padrão é 5%, então sabemos que 34% das amostras darão estimativas de aprovação estudantil entre 50% (parâmetro) e 55% (um erro padrão acima); outros 34%, das amostras mostrarão estimativas entre 50% e 45% (um desvio padrão abaixo do parâmetro). Reunindo-as, sabemos que aproximadamente dois terços (68%) das amostras resultarão em estimativas dentro (mais ou menos) de 5% do parâmetro.

Além disso, a teoria da probabilidade diz que cerca de 95% das amostras estarão dentro de (mais ou menos) dois erros padrão do valor verdadeiro, e 99,9%, das amostras dentro de (mais ou menos) três desvios padrão. Neste exemplo, portanto, sabemos que só uma amostra em mil resultaria numa estimativa abaixo de 35% de aprovação ou acima de 65% de aprovação.

A proporção de amostras caindo entre um, dois ou três erros padrão do parâmetro é constante para qualquer procedimento de amostragem aleatória como a que foi descrita, contanto que se seleccione grande número de amostras. Contudo, o tamanho do erro padrão é uma função do parâmetro populacional e do tamanho da amostra. Se voltarmos à fórmula, vemos que o erro padrão aumenta como função de um aumento na quantidade PQ . Veja, também, que esta quantidade atinge o máximo quando a população é dividida ao meio. Se $P = 0,5$, $PQ = 0,25$; se $P = 0,6$, $PQ = 0,24$; se $P = 0,8$, $PQ = 0,16$; se $P = 0,99$, $PQ = 0,0099$. Por extensão, se P for 0,0 ou 1,0 (isto é, se 0% ou 100% aprovarem o código estudantil), o erro padrão será 0. Se todos na população têm a mesma atitude (sem variação), então qualquer amostra dará exatamente a mesma estimativa.

O erro padrão é também uma função inversa do tamanho da amostra. À medida que o tamanho da amostra cresce, o erro padrão diminui, e as várias amostras estarão aglomeradas mais próximas ao valor verdadeiro. Outra regra é evidente a partir da fórmula: devido à raiz quadrada na fórmula, o erro padrão fica reduzido pela metade se o tamanho da amostra for *quadruplicado*. Em nosso exemplo, amostras de 100 produzem um erro padrão de 5%; para reduzir o erro padrão a 2,5%, teríamos que aumentar o tamanho da amostra para 400. (O apêndice B traz uma tabela para ajudá-lo a fazer essas estimativas.)

O que discutimos provém da teoria da probabilidade estabelecida, referente à seleção de grandes números de amostras aleatórias. Se o parâmetro populacional for conhecido, e muitas amostras aleatórias forem selecionadas, poderemos prever quantas amostras cairão dentro de intervalos especificados a partir do parâmetro. No entanto, estas condições não são típicas do *survey* por amostragem. Usualmente, o

pesquisador desconhece o parâmetro e conduz um *survey* por amostragem para estimá-lo. Ademais, os pesquisadores de *survey* normalmente não selecionam grandes números de amostras, em geral extraindo só uma. Apesar disso, a discussão acima sobre a teoria da probabilidade fornece uma base para inferências sobre a situação típica dos *surveys*.

Se a teoria da probabilidade diz que 68% das amostras estarão dentro de um erro padrão do parâmetro, o pesquisador infere que uma amostra aleatória tem uma probabilidade de 68% de estar dentro daquela faixa. Neste sentido, falamos de *níveis de confiança*: você tem "68% de confiança" que a estimativa da sua amostra está dentro de um erro padrão do parâmetro, ou tem "95% de confiança" que a estatística da amostra está dentro de dois erros padrão do parâmetro, e assim por diante. Compreensivelmente, a confiança aumenta quanto mais se estende a margem para erro. Você está virtualmente certo (99,9% de confiança) que a estimativa da amostra está dentro de três erros padrão do valor verdadeiro.

Embora confiante (em algum nível) de estar dentro de uma certa faixa do parâmetro, já assinalamos que raramente você conhece o parâmetro. Para resolver este dilema, você substitui o parâmetro na fórmula pela sua estimativa de amostragem; na falta do valor verdadeiro, você o substitui com a melhor estimativa disponível.

O resultado destas inferências e estimativas é que o pesquisador de *survey* torna-se capaz de estimar tanto um parâmetro populacional quanto o grau esperado de erro com base numa amostra extraída da população. Começando com a pergunta: "Qual porcentagem do corpo estudantil aprova o novo código?", você pode selecionar uma amostra aleatória de 100 alunos e entrevistá-los. Você pode relatar que sua melhor estimativa é que 50% do corpo estudantil aprovam o código, e tem 95% de confiança que entre 40 e 60% (mais ou menos dois erros padrão) aprovam. A faixa de 40 a 60% é chamada de *intervalo de confiança*.

Esta é a lógica básica da amostragem probabilística. A seleção aleatória permite ligar os achados do *survey* por amostragem ao corpo da teoria da probabilidade para estimar a precisão daqueles achados. Todos os enunciados de precisão na amostragem devem especificar tanto um nível de confiança quanto um intervalo de confiança. Você pode informar que está *x%* confiante de que o parâmetro populacional está entre dois valores especificados.

A discussão precedente considerou apenas um tipo de estatística — as porcentagens produzidas por uma variável binomial ou dicotômica. A mesma lógica se aplica ao exame de outras estatísticas, tais como renda média. Uma vez que as computações são mais complicadas, nesta introdução consideramos apenas os binomiais.

Advertimos que o uso da teoria da probabilidade, como discutido nesta seção, não está plenamente justificado em termos técnicos. A teoria da distribuição amostral, por exemplo, faz suposições que quase nunca se aplicam às situações de *surveys*. A fórmula usada para avaliar o número de amostras que estariam contidas dentro de incrementos especificados de erros padrão, por exemplo, supõe uma população infinita, um número infinito de amostras e amostragem com substituição. Ademais, na nossa discussão, o salto inferencial da distribuição de várias amostras para as características prováveis de uma amostra foi grosseiramente simplificado.

Essas advertências são feitas por razões de perspectiva. Pesquisadores, às vezes, parecem superestimar a precisão das estimativas produzidas pelo uso da teoria da probabilidade em conexão com a pesquisa por *survey*. Mencionaremos em outras partes deste capítulo e do livro que variações nas técnicas de amostragem e fatores de não-amostragem podem reduzir ainda mais a legitimidade de tais estimativas. Não obstante, os cálculos discutidos nesta seção podem ser de grande valor para compreender e avaliar dados. Embora não forneçam estimativas tão precisas quanto se pode supor, podem ser bastante válidos para fins práticos, e são inquestionavelmente mais válidos do que estimativas derivadas menos rigorosamente, baseadas em métodos de amostragem menos rigorosos. Você deve se familiarizar com a *lógica* básica subjacente aos cálculos, para poder reagir sensatamente tanto aos seus próprios dados quanto aos dados informados por outros.

Populações e Molduras de Amostragem

A seção anterior lidou com um modelo teórico para amostragem de *survey*. Se o consumidor de *surveys*, o estudante ou o pesquisador precisam conhecer essa teoria, devem igualmente saber avaliar as condições imperfeitas existentes no campo. Esta seção discute um aspecto das condições de campo que requer compromisso entre as condições e os supostos teóricos — a congruência ou a disparidade entre populações e molduras de amostragem.

Dito simplesmente, uma moldura de amostragem é a lista, ou fac-símile razoável, de elementos dos quais é selecionada uma amostra probabilística. A próxima seção trata dos métodos de selecionar amostras, mas primeiro precisamos considerar a própria moldura de amostragem. Amostras corretamente selecionadas fornecem informações apropriadas para descrever a população de elementos que compõem a moldura de amostragem — nada mais. É preciso enfatizar este ponto, dada a tendência tão comum entre pesquisadores de selecionar amostra de uma dada moldura de amostragem e então fazer afirmações sobre uma população semelhante, mas não idêntica, à população do *survey* definida pela moldura de amostragem. É o problema mais amplo da generalização social-científica, equivalente a estudar uma pequena Igreja Luterana em Dakota do Norte para descrever a religião nos EUA. No restante da seção examinaremos finalidades diversas para os *surveys* e discutiremos boas e más molduras de amostragem para aquelas finalidades.

Surveys de organizações frequentemente são o tipo mais simples de *survey*, do ponto de vista da amostragem, porque, tipicamente, organizações possuem listas de membros. Nestes casos, a lista de membros é uma excelente moldura de amostragem. Se uma amostra aleatória é selecionada de uma lista de membros, os dados colhidos desta amostra podem ser considerados como representativos de todos os membros — *caso todos os membros forem incluídos na lista*. Se alguns não tiverem sido incluídos, deve-se fazer um esforço para amostrá-los; do contrário, os achados do *survey* só podem ser considerados como representativos daqueles membros na lista.

Populações que oferecem boas listas organizacionais são escolas primárias e secundárias, estudantes universitários; membros de igrejas; operários de fábricas; membros de fraternidades e grêmios, de clubes sociais, de serviço ou políticos; e membros de associações profissionais.

Estes comentários se aplicam principalmente a organizações locais. Organizações estaduais ou nacionais frequentemente não têm uma única lista de membros facilmente disponível. Por exemplo, não há uma lista única de todos os membros da Igreja Episcopal. Mas um planejamento de amostragem um pouco mais complexo pode aproveitar as listas de membros de igrejas locais, amostrando primeiro igrejas e depois subamostrando as listas de membros dessas igrejas.

Outras listas de indivíduos podem ser especialmente relevantes às necessidades de pesquisa de *survey*. Por exemplo,

agências governamentais mantêm listas de eleitores registrados caso você queira fazer uma pesquisa pré-eleitoral de opinião ou um exame mais detalhado do comportamento eleitoral, mas você deve certificar-se de que a lista está atualizada. Listas semelhantes contêm nomes de donos de automóveis, de recebedores de benefícios sociais, de contribuintes de impostos, pessoas com licenças comerciais, profissionais licenciados etc. Embora algumas destas listas sejam difíceis de conseguir, podem fornecer excelentes molduras de amostragem. Como os elementos de amostragem num *survey* não precisam ser pessoas, há diversas listas com outros tipos de elementos: universidades, negócios de vários tipos, cidades, revistas acadêmicas, jornais, sindicatos, clubes políticos, associações profissionais etc.

Catálogos telefônicos são muito usados para pesquisas de opinião “rápidas e sujas”. Inegavelmente são fáceis de fazer e de baixo custo, o que sem dúvida responde pela sua popularidade, e se você quiser fazer afirmações sobre assinantes de telefones, o catálogo é uma moldura de amostragem *razoavelmente boa*.³ Infelizmente, com excessiva frequência, os catálogos telefônicos são usados como lista da população de uma cidade ou dos seus eleitores. Há muitos defeitos neste tipo de raciocínio, o principal sendo um viés de classe social. Gente pobre tem menos possibilidade de ter telefones; uma amostra de catálogo telefônico, portanto, provavelmente terá uma predisposição para as classes média e alta. (O Capítulo 10 examina um método especial para amostragem em relação a *surveys* telefônicos: *discagem aleatória de algarismos*.)

Catálogos de ruas e mapas de impostos são frequentemente usados como fontes fáceis para amostras de residências, mas podem sofrer das mesmas desvantagens dos catálogos telefônicos, tais como serem incompletos e/ou enviesados. Por exemplo, em regiões urbanas com zoneamento rigoroso, unidades residenciais ilegais têm pouca possibilidade de aparecerem em registros oficiais. Resulta que, não tendo estas unidades chance de seleção, os achados das amostras não são representativas dessas unidades, tipicamente mais pobres e superlotadas.

Revisão de Populações e Molduras de Amostragem

Uma vez que a literatura da pesquisa por *surveys* tem sido surpreendentemente descuidada com as questões de populações e molduras de amostragem, dou-lhes atenção especial. Para maior ênfase, listo, em revisão, as principais diretrizes a lembrar.

1. Os achados dos *surveys* por amostragem só podem ser considerados representativos da agregação de elementos que compõem a moldura de amostragem.
2. Muitas vezes, as molduras de amostragem não incluem de fato todos os elementos que seus nomes sugerem. (Diretórios de estudantes não incluem todos os estudantes; catálogos telefônicos não incluem todos os assinantes.) Omissões são quase inevitáveis. Portanto, uma primeira preocupação deve ser avaliar a extensão destas omissões e corrigi-las se possível. (Naturalmente, talvez você possa ignorar, sem prejuízo, um pequeno número de omissões de difícil correção.)
3. Para poder generalizar a população que compõe a moldura de amostragem, é necessário que todos os elementos tenham representação igual nela; cada elemento deve aparecer uma só vez. Do contrário, elementos que aparecem mais de uma vez terão maior probabilidade de seleção, e a amostra, no geral, sobre-representará estes elementos.

Outros assuntos mais práticos, relacionados às populações e molduras de amostragem, são tratados em outros capítulos. Por exemplo, o formato da moldura de amostragem — uma lista numa publicação, um arquivo em fichas 3x5, placas *addressograph*, cartões IBM, fitas magnéticas, disquetes — é muito importante. Tais considerações frequentemente ganham precedência sobre considerações científicas; uma lista mais “fácil” poderá ser escolhida contra uma lista mais “difícil”, mesmo que esta seja mais apropriada à população alvo. Não devemos ser dogmáticos a este respeito; os pesquisadores devem pesar cuidadosamente as vantagens e as desvantagens relativas destas alternativas. O mais importante é que você esteja ciente — e deve informar seu leitor — das deficiências de qualquer moldura de amostragem que escolher.

Tipos de Desenhos de Amostragem

Introdução

Talvez, a essa altura, você esteja apavorado com a importância e as dificuldades para organizar sua moldura de amostragem; é um sentimento correto e saudável. Após estabelecer a moldura de amostragem, você deve selecionar uma amostra de elementos para estudo.

Até agora, focalizamos a *amostragem aleatória simples*. De fato, as estatísticas usadas em geral pelos pesquisadores de *survey* supõem tal amostra. Mas, como logo veremos, há diversas opções para fazer a amostragem, e raramente se opta pela aleatória simples, por duas razões. Primeiro, a amostragem aleatória simples só é possível com as molduras de amostragem mais simples. Segundo, a amostragem aleatória simples pode não ser o método melhor (o mais preciso). Discutiremos agora a amostragem aleatória simples e outros tipos de desenhos de amostragem.

Amostragem Aleatória Simples

A amostragem aleatória simples [AAS] é o método de amostragem básico suposto pelos cálculos estatísticos do *survey*. A matemática da amostragem aleatória é particularmente complexa, e nos desviaremos dela, preferindo descrever os métodos de campo que empregam este método.

Uma vez estabelecida uma moldura de amostragem de acordo com a discussão anterior, você numerará cada elemento da lista, atribuindo um e só um número a cada um, sem saltar nenhum número. A seguir, usa uma tabela de números aleatórios para selecionar os elementos da amostra.

Se sua moldura de amostragem está em formato próprio para computadores, como um disquete, uma amostra aleatória simples pode ser extraída automaticamente, usando um programa de computador relativamente simples. Com efeito, o computador pode numerar os elementos da moldura de amostragem, gerar sua própria série de números aleatórios e imprimir a lista de elementos selecionados.

Amostragem Sistemática

A amostragem aleatória simples raramente é usada na prática. Em geral, não é o método de amostragem mais eficiente, e já vimos que pode ser trabalhoso quando feito manualmente. A AAS habitualmente requer uma lista de elementos; quando esta lista está disponível, geralmente os pesquisadores usam o método de *amostragem sistemática*, em vez da amostragem aleatória simples.

Na amostragem sistemática, cada k^o elemento na lista total (por exemplo, cada 100^o) é escolhido para inclusão na amostra. Se a lista tem 10.000 elementos, e você quer uma amostra de 1.000 elementos, seleciona cada décimo elemento para a amostra. Para garantir contra algum viés humano

possível ao usar este método, você seleciona o primeiro elemento aleatoriamente. No exemplo acima, você começa escolhendo um número aleatório entre 1 e 10; o elemento com este número é incluído na amostra, mais cada décimo elemento seguinte. Refere-se tecnicamente a este procedimento como uma "amostra sistemática de início aleatório".

Dois termos são usados frequentemente em conexão com a amostragem sistemática. O *intervalo de amostragem* é a distância padrão entre os elementos selecionados na amostra — 10 no exemplo acima. A *razão amostral* é a proporção de elementos selecionados da população — um décimo no exemplo acima. (A razão amostral é igual a $1/\text{intervalo de amostragem}$.)

Na prática, a amostragem sistemática é virtualmente idêntica à amostragem aleatória simples. De fato, se a lista de elementos for "randomizada" antes da amostragem, pode-se argumentar que uma amostragem sistemática retirada daquela lista é, na realidade, uma amostra aleatória simples. Os debates quanto aos méritos relativos da amostragem aleatória simples e da amostragem sistemática foram resolvidos largamente em favor do método mais simples, a sistemática. Empiricamente, os resultados são virtualmente idênticos. Além disso, como veremos, a amostragem sistemática, em algumas situações, é um pouco mais precisa do que a amostragem aleatória simples.

A amostragem sistemática traz um perigo: o arranjo dos elementos na lista pode desaconselhar uma amostra sistemática. Este perigo é em geral referido pelo termo *periodicidade*. Se a lista de elementos estiver organizada num padrão cíclico que coincide com o intervalo de amostragem, resultará uma amostra grosseiramente tendenciosa. Dois exemplos bastam.

Num estudo de soldados durante a II Guerra Mundial, os pesquisadores selecionaram cada décimo soldado das listas das unidades. Mas estas estavam organizadas sob a forma de uma tabela organizacional, listando primeiro os sargentos e depois cabos e praças, pelotão por pelotão, sendo que cada pelotão tinha dez membros. Portanto, cada décima pessoa na lista era um sargento de pelotão, e a amostra sistemática selecionada tinha apenas sargentos. Poderia acontecer, também, que nenhum sargento fosse selecionado.

Em outro exemplo, suponha que queremos selecionar uma amostra de apartamentos num prédio. Se a amostra for extraída de uma lista dos apartamentos organizados em ordem numérica (por exemplo, 101, 102, 103, 104, 201, 202, e assim por diante), há o risco de o intervalo de amostragem coincidir

com o número de apartamentos num andar ou seu múltiplo. Neste caso, a amostra poderia incluir apenas apartamentos do lado noroeste, ou só apartamentos perto dos elevadores. Se estes tipos de apartamentos tiverem alguma característica peculiar em comum (por exemplo, aluguel mais elevado), a amostra seria tendenciosa. Risco semelhante pode surgir numa amostragem sistemática de casas numa subdivisão organizada com o mesmo número de casas num quarteirão.

Ao considerar uma amostra sistemática a partir de uma lista, você deve examinar cuidadosamente a natureza da lista. Se os elementos estão arranjados em alguma ordem particular, certifique-se de que aquela ordem não enviesará a amostra a ser selecionada e tome medidas para contrabalançar qualquer possível viés (por exemplo, tirar uma amostra aleatória simples de porções cíclicas).

Em resumo, a amostragem sistemática, em geral, é superior à amostragem aleatória simples, pelo menos em conveniência. Os problemas na ordenação dos elementos da moldura de amostragem geralmente podem ser remediados com facilidade.

Amostragem Estratificada

Nas duas seções precedentes, discutimos dois métodos alternativos de seleção de amostras a partir de uma lista. A *amostragem estratificada* não é uma alternativa a estes métodos, mas representa uma possível modificação no seu uso.

As amostragens aleatória simples e sistemática são importantes no sentido de garantirem um grau de representatividade e permitirem uma estimativa de erro. A amostragem estratificada é um método para obter maior grau de representatividade, reduzindo o provável erro amostral. Para compreender por que, voltemos rapidamente à teoria básica da distribuição amostral.

Lembre que o erro amostral é reduzido por dois fatores no desenho da amostra. Primeiro, uma amostra maior produz menor erro amostral do que uma amostra menor. Segundo, uma população homogênea produz amostras com menores erros amostrais do que uma população heterogênea. Se 99% da população concordam com determinado enunciado, é improvável que qualquer amostra probabilística interprete erroneamente a extensão da concordância. Se, por outro lado, a população está dividida meio a meio quanto ao enunciado, então o erro amostral será muito maior.

A amostragem estratificada baseia-se neste segundo fator na teoria de amostragem. Em vez de selecionar a amostra da

população total em geral, você garante que quantidades apropriadas de elementos sejam tirados de subconjuntos homogêneos da população. Por exemplo, estudando estudantes universitários, você pode organizar sua população por turmas, tirando de cada uma quantidades apropriadas. Numa amostra não estratificada, a representação por turma estaria sujeita ao mesmo erro amostral que outras variáveis. Na amostra estratificada por turma, o erro amostral quanto a esta variável reduz-se a zero.

Você pode querer usar um método de estratificação ainda mais complexo. Além de estratificar por turma, pode também estratificar por gênero, por média de notas etc. Assim, talvez você garanta que sua amostra tenha números apropriados de alunos da turma A com nota média 4,0; também números apropriados de alunas dessa turma com média 4,0 etc.

Portanto, a função última da estratificação é organizar a população em subconjuntos homogêneos (com heterogeneidade entre os subconjuntos) e selecionar o número apropriado de elementos de cada subconjunto. Já que os subconjuntos são homogêneos nas variáveis de estratificação, também podem ser homogêneos quanto a outras variáveis. Por exemplo, se idade estiver fortemente relacionada a turma, uma amostragem estratificada por turma também será mais representativa em termos de idade. Como aspirações ocupacionais estão relacionadas a sexo, uma amostra estratificada por sexo será mais representativa em termos das aspirações ocupacionais.

A escolha das variáveis de estratificação depende em geral de quais variáveis estão disponíveis. Sexo pode muitas vezes ser determinado através de uma lista de nomes. As listas universitárias são tipicamente organizadas por turma. As listas de professores podem indicar sua lotação por departamento. Arquivos de organizações governamentais podem estar organizados por região geográfica. Listas de registros de eleitores estão organizadas por seções e zonas eleitorais.

No entanto, ao selecionar as variáveis de estratificação das listas disponíveis, você deve se preocupar sobretudo com as que se relacionam, presumivelmente, com as variáveis que você quer representar com precisão. Como sexo está relacionado a muitas variáveis e muitas vezes está disponível para estratificação, acaba sendo usado freqüentemente. Escolaridade está relacionada a muitas variáveis, mas freqüentemente não está disponível para estratificação. Localização geográfica numa cidade, estado ou nação está relacionada a muitas coisas. Dentro de uma cidade, a estratificação por lugar geográfico, em geral, aumenta a representatividade em áreas como

classe social e grupo étnico. Dentro de uma nação, aumenta a representatividade em uma vasta gama de atitudes, bem como na classe social e na etnia.

Os métodos de estratificação na amostragem variam. Quando se trabalha com uma lista simples de todos os elementos na população, dois métodos predominam. Primeiro, pode-se agrupar os elementos da população em grupos discretos com base nas variáveis de estratificação usadas. Com base na proporção relativa de um grupo na população, seleciona-se — aleatória ou sistematicamente — um número de elementos daquele grupo na mesma proporção do tamanho de amostra desejado. Por exemplo, se alunos da turma A com média 4,0 são 1% da população estudantil e você quer uma amostra de 1.000 estudantes, selecione 10 estudantes dos alunos da turma A com média 4,0.

Como método alternativo, pode-se agrupar estudantes como descrito acima e pôr esses vários grupos juntos numa lista contínua, começando com todos os alunos da turma A com média 4,0 e terminando, digamos, com todas as alunas da turma Z com média 1,0 ou menos. Seleciona-se então uma amostra sistemática, com início aleatório, de toda a lista. Dada a organização de lista, uma amostra sistemática selecionaria números adequados (dentro de uma margem de erro de 1 ou 2) de cada subgrupo. (*Nota:* Uma amostra aleatória simples retirada de tal lista composta cancelaria a estratificação.)

O efeito da estratificação é garantir a representação adequada das variáveis de estratificação para a aumentar a representação de outras variáveis relacionadas a elas. Portanto, tomada como um todo, uma amostra estratificada tem possibilidade de ser mais representativa num certo número de variáveis do que no caso de uma amostra aleatória simples. Embora a amostra aleatória simples ainda seja vista como algo sagrado, deve estar claro que muitas vezes há coisa melhor.

Estratificação Implícita na Amostragem Sistemática

Mencionei antes que a amostragem sistemática pode, em certas condições, ser mais precisa do que a amostragem aleatória simples. É o caso sempre que o arranjo da lista cria uma "estratificação implícita". Como assinalai, se a lista de estudantes universitários está organizada por turma, uma amostra sistemática produzirá a estratificação por período, ao contrário de uma amostra aleatória simples. Outros arranjos típicos de elementos em listas podem ter o mesmo efeito.

Se uma lista de nomes que compõe uma moldura de amostragem está em ordem alfabética, ela é um pouco estratificada por origem étnica. Por exemplo, todos os McFavishes estão reunidos, bem como os Lees, Wongs, Yamamuras, Schmidts, Whitehalls, Weinsteins, Gonzaleses etc. Na medida em que qualquer destes subgrupos forma um subconjunto substancial da população total, estará adequadamente representado numa amostra sistemática retirada de uma lista alfabética.

Num estudo de estudantes da Universidade do Havai, depois de estratificados por turma, os estudantes foram organizados pelos números de identidade estudantil. No entanto, estes números eram seus números de Seguro Social. Os primeiros três dígitos do número de Seguro Social indicam o estado onde foi emitido. Portanto, dentro de cada turma, os estudantes estavam organizados pelo estado onde havia sido emitido seu número de Seguro Social, fornecendo uma estratificação aproximada por origem geográfica.

Uma lista ordenada de elementos pode ser mais útil do que uma lista não ordenada, aleatória. Ênfase este ponto para combater a crença infeliz de que listas devem ser randomizadas antes de ocorrer a amostragem sistemática. A lista só deve ser reordenada se o arranjo apresenta o problema da *periodicidade* já discutido.

Amostragem por Conglomerados em Múltiplas Etapas, Geral

As quatro seções precedentes lidaram com procedimentos razoavelmente simples de amostragem a partir de listas de elementos, que representam a situação ideal. Entretanto, infelizmente muita pesquisa social interessante requer selecionar amostras de populações que não podem ser facilmente listadas para fins de amostragem, tais como a população de uma cidade, de um estado ou da nação, todos os estudantes universitários nos Estados Unidos etc. Nestes casos, é necessário criar e executar um desenho de amostra mais complexo. Este desenho tipicamente envolve a amostragem inicial de *grupos* de elementos, conhecidos como *conglomerados* [clusters], seguida pela seleção de elementos de cada conglomerado selecionado. As variedades e procedimentos da *amostragem por conglomerados em múltiplas etapas* são dissecadas em detalhe nos exemplos de amostragem no Capítulo 6. Entretanto, vamos delinear-los aqui.

Pode-se fazer amostragem por conglomerados quando não é possível ou prático compilar uma lista exaustiva dos elementos da população alvo. Todos os membros de igrejas nos Estados Unidos exemplificam uma população assim.

Contudo, muitas vezes os elementos da população já estão agrupados em subpopulações, e uma lista destas subpopulações existe ou pode ser criada de forma prática. Por exemplo, membros de igrejas nos Estados Unidos pertencem a igrejas específicas, e é possível descobrir ou criar uma lista dessas igrejas. Seguindo um formato de amostragem por conglomerado, a lista de igrejas é amostrada numa das maneiras descritas nas seções anteriores (por exemplo, uma amostra estratificada sistemática). Em seguida, obtêm-se listas de membros de cada uma das igrejas selecionadas e amostra-se cada uma das listas, de forma a ter amostras de membros de igreja para estudo.⁶

Outra situação típica é a amostragem em áreas habitadas, como uma cidade. Embora não haja uma lista única da população de uma cidade, os cidadãos residem em quarteirões separados ou zonas de recenseamento. Portanto, é possível começar selecionando uma amostra de quarteirões, criar uma lista de pessoas morando nos quarteirões selecionados e, finalmente, subamostrar pessoas em cada quarteirão.

Num desenho mais complexo, você pode amostrar os quarteirões, listar as residências em cada quarteirão selecionado, amostrá-las, listar as pessoas residindo em cada uma e, finalmente, amostrar as pessoas em cada residência selecionada. Este desenho de amostragem em múltiplas etapas leva à seleção final de uma amostra de indivíduos sem requerer a listagem inicial de todos os indivíduos que compõem a população da cidade.

Portanto, também na amostragem por conglomerados em múltiplas etapas repetem-se dois passos básicos: listar e amostrar. A lista de unidades primárias de amostragem (igrejas ou quarteirões, por exemplo) é compilada e talvez estratificada para amostragem. Uma amostra dessas unidades é selecionada. As unidades primárias de amostragem selecionadas são, então, listadas e talvez estratificadas. A lista de unidades secundárias de amostragem é amostrada, e assim por diante. Os métodos de listagem e amostragem são apresentadas detalhadamente nos exemplos do Capítulo 6.

A amostragem por conglomerados é altamente recomendada pela eficiência, embora à custa da precisão. Enquanto uma amostra aleatória simples extraída de uma lista populacional está sujeita a um único erro amostral, uma amostra por conglomerados em duas etapas está sujeita a dois. Primeiro, a amostra inicial dos conglomerados representa a população de conglomerados somente dentro de uma faixa de erro amostral. Segundo, a amostra de elementos extraída de um conglomerado representa todos os elementos dele apenas dentro de uma faixa de erro amostral. Assim, por exemplo, há o risco de

selecionar uma amostra de quarteirões desproporcionalmente ricos, bem como uma amostra de residências desproporcionalmente ricas dentro destes quarteirões. A melhor solução ao problema está no número de conglomerados selecionados no início e no número de elementos selecionados em cada conglomerado.

Em geral, estamos restritos a um tamanho máximo de amostra; por exemplo, podemos ter o limite de 2.000 entrevistas numa cidade. Dada esta limitação maior, temos várias opções para desenhar a amostra por conglomerado. Num extremo, podemos escolher um conglomerado e selecionar 2.000 elementos dentro dele, em outro, escolher 2.000 conglomerados e selecionar um elemento em cada um. Naturalmente, nenhum é aconselhável, mas há muitas opções entre estes extremos. Felizmente, a lógica das distribuições amostrais provê uma diretriz geral a seguir.

Lembre que o erro amostral é reduzido por dois fatores: o aumento do tamanho da amostra e uma maior homogeneidade dos elementos a serem amostrados. Estes fatores operam em cada nível do desenho da amostra em múltiplas etapas. Uma amostra por conglomerados representará melhor todos os conglomerados se um grande número for selecionado e se todos eles forem bastante semelhantes. Uma amostra de elementos representará melhor todos os elementos de um conglomerado se um grande número de elementos for selecionado do conglomerado, e se todos os elementos forem bastante semelhantes.

Entretanto, dado um tamanho máximo da amostra, se o número de conglomerados aumentar, o número de elementos num conglomerado deve diminuir. Neste sentido, a representatividade dos conglomerados aumenta à custa da menor representação dos elementos de cada um, e vice-versa. Felizmente, podemos usar o fator da homogeneidade para aliviar esta dificuldade.

Os elementos de um conglomerado natural dentro de uma população são tipicamente mais homogêneos do que todos os elementos da população total. Os membros de uma igreja são mais semelhantes entre si do que todos os membros de igrejas; da mesma forma, os moradores de um quarteirão são mais semelhantes entre si do que todos os moradores de uma cidade inteira. Portanto, podem ser necessários relativamente poucos elementos para representar adequadamente um conglomerado natural, enquanto um número maior de conglomerados pode ser necessário para representar adequadamente a diversidade dos conglomerados. Vê-se isto com máxima clareza no caso extremo de conglomerados muito diferentes entre si, compostos de elementos

exatamente idênticos em cada um. Nesta situação, um número grande de conglomerados representaria adequadamente a variedade entre os conglomerados, bastando apenas um elemento de cada um para representar adequadamente todos os outros. Embora esta situação extrema nunca ocorra, ela está mais próxima da verdade mais vezes do que o seu oposto — conglomerados idênticos compostos de elementos muito divergentes.

Portanto, a diretriz geral para o desenho de conglomerados é maximizar o número de conglomerados selecionados, diminuindo o número de elementos de cada um. Mas esta diretriz científica deve ser contrastada com as restrições administrativas. A eficiência da amostragem por conglomerados funda-se na capacidade de minimizar a listagem de elementos da população. Selecionando inicialmente conglomerados, buscamos listar apenas os elementos dos conglomerados selecionados, em vez de todos os elementos da população inteira. Aumentar o número de conglomerados vai diretamente de encontro a este fator de eficiência. Um número menor de conglomerados pode ser listado mais rápida e economicamente do que um número maior. (Lembre-se de que todos os elementos de um conglomerado selecionado devem ser listados, mesmo que poucos vão ser escolhidos na amostra.)

O desenho final da amostra refletirá estas duas restrições. Provavelmente, você vai selecionar o maior número de conglomerados que seu orçamento permitir. Mas, para não deixar esta questão aberta demais, apresentamos uma regra simples. Pesquisadores populacionais convencionalmente selecionam cinco residências por zona ou quarteirão de recenseamento. Querendo entrevistar um total de 2.000 residências, você seleciona 400 quarteirões com cinco entrevistas de residências em cada um. Voltaremos a esta regra em exemplos posteriores de desenhos de amostra.

Antes de passarmos a procedimentos mais detalhados da amostragem por conglomerados, cabe lembrar que este método quase inevitavelmente envolve perda de precisão. A maneira como ocorre esta perda de precisão é complexa. Primeiro, um desenho de amostra em etapas múltiplas está sujeito a erros amostrais em cada uma das etapas. Como o tamanho da amostra é necessariamente menor, a cada etapa, do que o tamanho total da amostra, o erro amostral em cada etapa será maior do que no caso de uma amostra aleatória de elementos em etapa única. Segundo, o erro amostral é estimado com base na variância observada entre os elementos da amostra. Quando estes elementos são retirados

de conglomerados relativamente homogêneos, as estimativas do erro de amostragem serão otimistas demais, devendo ser corrigidas à luz do desenho de amostras por conglomerado. (Este problema será discutido em detalhes na consideração posterior sobre análise univariada.)

Amostragem de Aglomerados em Múltiplas Etapas, Estratificação

Até agora, discutimos a amostragem por conglomerados como se fosse selecionada uma amostra aleatória simples a cada etapa do desenho. Na prática, podemos usar as técnicas de estratificação discutidas antes para refinar e melhorar a amostra selecionada.

Exemplos posteriores detalharão métodos possíveis de estratificação; por ora, observemos que nossas opções básicas são essencialmente as mesmas que temos na amostragem de etapa única a partir de uma lista. Por exemplo, ao selecionar uma amostra nacional de igrejas, você pode, no começo, estratificar sua lista de igrejas por denominação, região geográfica, tamanho, localização urbana/rural, e talvez alguma medida da classe social. Informação do censo pode ser usada por pesquisadores de população para estratificar quarteirões de recenseamento em termos de composição étnica, classe social, valores das propriedades, qualidade de estruturas, natureza da propriedade das construções e tamanho.

Assim que as unidades primárias de amostragem (igrejas ou quarteirões, nos exemplos) estiverem agrupadas de acordo com as variáveis de estratificação relevantes disponíveis, pode-se usar tanto técnicas de amostragem aleatória simples ou a amostragem sistemática para extrair a amostra. Você pode selecionar um número especificado de unidades de cada grupo ou "estrato", ou arranjar os conglomerados estratificados numa lista contínua, amostrando sistematicamente esta lista. O erro amostral será tão mais reduzido, nesta etapa, quanto mais os conglomerados estiverem combinados em estratos homogêneos. Como antes, o objetivo primordial da estratificação é a homogeneidade.

Não há razão para não se poder estratificar em cada nível de amostragem. Os elementos listados de um aglomerado selecionado podem ser estratificados antes da próxima etapa de amostragem. No entanto, isso não é feito comumente. (Lembre a suposição sobre a homogeneidade relativa dentro dos aglomerados.)

Amostragem de Probabilidade Proporcional ao Tamanho (PPT)

Até agora, falei genericamente sobre a designação de elementos da amostra a conglomerados selecionados em termos de quantos conglomerados devem ser selecionados e quantos elementos deve haver em cada um. Esta seção discute em maior detalhe algumas opções.

Para garantir a seleção de uma amostra representativa de elementos, você deve dar a cada elemento da população total uma oportunidade igual de ser selecionada. A forma mais simples de fazer isso numa amostra por conglomerados é dar a cada um a mesma chance de seleção e selecionar uma determinada proporção de elementos de cada conglomerado selecionado. Com uma população de 100.000 elementos agrupados em 1.000 conglomerados (de tamanhos variados), pretendendo amostrar 1.000 elementos, você seleciona um décimo dos conglomerados (100) com igual probabilidade e subseleciona um décimo dos elementos de cada conglomerado escolhido. Assim, cerca de 1.000 elementos são selecionados e cada elemento na população tem a mesma $(1/10 \times 1/10 = 1/100)$ probabilidade de seleção. Esta técnica de seleção de amostras, embora a mais clara e simples, não é a mais eficiente.

A maioria das amostragens por conglomerados envolve conglomerados de tamanhos muito diversos (em número de elementos). O pesquisador de religiões encontra igrejas muito grandes e muito pequenas. O pesquisador populacional acha quarteirões muito populosos e outros nem tanto. Além disso, os conglomerados pequenos muitas vezes são mais numerosos do que os grandes, embora os grandes possam responder por uma proporção maior da população total. Assim, alguns poucos grandes quarteirões podem conter uma proporção grande da população de uma cidade, enquanto o número maior de quarteirões menores de fato contém apenas uma pequena parcela da população.

Selecionar conglomerados com igual probabilidade, retirando uma proporção fixa de elementos dos conglomerados selecionados, resulta no seguinte: (1) Seleciona-se um número relativamente pequeno de conglomerados grandes na primeira etapa da amostragem. (2) Os elementos selecionados para representar todos os elementos dos grandes aglomerados são retirados de muito poucos desse conglomerados. Na situação extrema, toda a população da cidade residindo em dez grandes quarteirões poderia ser representada pelas pessoas vivendo em apenas um deles.

A seção sobre a amostragem por conglomerados discutiu a maior eficiência inerente à seleção de muitos conglomerados, com poucos elementos sendo retirados de cada um. Este princípio é posto em prática através do método de *amostragem de probabilidade proporcional ao tamanho* (PPT), que permite selecionar maior número de conglomerados, garante a representação dos elementos de conglomerados grandes e dá a cada elemento na população igual chance de seleção.

Na primeira etapa de amostragem, é dada a cada conglomerado oportunidade de seleção proporcional ao seu tamanho (em número de elementos). Grandes conglomerados têm maior chance de seleção do que os menores. Mas, na segunda etapa de amostragem, o mesmo *número* de elementos é escolhido de cada conglomerado selecionado. O efeito destes dois procedimentos é igualar as probabilidades últimas de seleção de todos os elementos, já que elementos em conglomerados maiores têm uma oportunidade de seleção menor *dentro* do seu aglomerado do que elementos em conglomerados menores. Por exemplo, um quarteirão com 100 residências terá dez vezes mais chance de seleção do que um quarteirão com apenas 10. Se ambos os quarteirões forem selecionados e o mesmo número de residências for selecionado de cada quarteirão, as casas no quarteirão maior terão dez vezes menos chance de seleção do que as casas no quarteirão menor. A fórmula abaixo indica a probabilidade de seleção de um elemento num desenho de amostragem PPT.

$$\left[\begin{array}{l} \text{probabilidade} \\ \text{do elemento} \end{array} \right] = \left[\begin{array}{l} \text{número de} \\ \text{aglomerados} \\ \text{selecionados} \end{array} \right] \times \left[\begin{array}{l} \text{parte do aglomerado} \\ \text{tamanho do aglomerado} \end{array} \right] \times \left[\begin{array}{l} \text{elementos} \\ \text{selecionados} \\ \text{por aglomerado} \\ \text{tamanho do} \\ \text{aglomerado} \end{array} \right]$$

Se 100 conglomerados são selecionados e 10 elementos escolhidos de cada um, a partir de uma população total de 100.000, a probabilidade geral de seleção de cada elemento será de 1.000/100.000, ou 1/100. Um conglomerado com 100 elementos tem uma probabilidade de seleção igual a 100 (conglomerados a serem selecionados) vezes 100/100.000 (tamanho do conglomerado/tamanho da população), ou 1/10; cada elemento tem uma chance de 10/100 (elementos por conglomerado/tamanho do conglomerado), ou 1/10, de ser selecionado *dentro* daquele conglomerado; a chance total de seleção do elemento, neste caso, é 1/10 vezes 1/10, ou 1/100. Um conglomerado com 10 elementos tem uma probabilidade de seleção de 100 vezes 10/100.000, ou 1/100, mas a chance de seleção de cada elemento dentro do conglomerado é 10/10, ou 1, fazendo com que a chance total de seleção seja igual a 1/100.

Independente do número de elementos num conglomerado, cada elemento tem, enfim, a mesma probabilidade de seleção. Pode-se ver isto mais claramente na fórmula, notando que o tamanho do conglomerado aparece no numerador e no denominador, cancelando-se mutuamente. Portanto, a probabilidade de seleção se torna o número de conglomerados a serem escolhidos vezes o número de elementos a serem escolhidos de cada conglomerado selecionado, dividido pelo tamanho da população. Isso, é claro, é o tamanho da amostra dividido pelo tamanho da população.

Costuma-se fazer duas modificações neste desenho de amostra PPT. Primeiro, podemos achar necessário representar conglomerados *muito* grandes na amostra; por exemplo, podemos querer garantir que todos os quarteirões (ou igrejas) com mais de, digamos, 1.000 elementos sejam incluídos na amostra. Neste caso, podemos selecionar todos estes conglomerados logo de início (com uma probabilidade de 1,0), devendo-se, então, atribuir aos elementos destes agrupamentos uma probabilidade igual à razão amostral geral. No exemplo anterior, 1/100 dos elementos residindo em cada um dos grandes quarteirões poderiam ser selecionados.

A segunda modificação é relativa a pequenos conglomerados. Se um número padrão de elementos for ser selecionado de cada conglomerado escolhido, conglomerados com menos elementos do que o número padrão representam um problema. Se a meta é dez residências em cada quarteirão selecionado, o que fazer com quarteirões com cinco residências? A solução habitual é combinar pequenos conglomerados, de forma que cada combinação tenha pelo menos o número padrão a ser selecionado. (Se os conglomerados estiverem estratificados, as combinações devem ser feitas dentro de cada estrato. Da mesma forma, pequenos conglomerados podem ser agregados a outros maiores, caso este procedimento garanta maior homogeneidade dentro do conglomerado resultante). O exemplo de uma amostra por conglomerados por área no Capítulo 6 demonstra a necessidade de dar esse passo, para garantir que se levem em conta blocos onde se acredita não haver residências.

A título de resumo, a amostragem por conglomerados é um método de amostragem difícil, mas importante, adequado sempre que for impossível compilar uma lista de todos os elementos da população em estudo. A discussão foi abstrata, mas o exemplo dado no Capítulo 6 oferece um quadro mais claro dos passos dados, de fato, numa seleção complexa de amostras.

Amostragem e Ponderação Desproporcionais

Em última análise, uma amostra probabilística é representativa de uma população se todos os elementos da população tiverem uma chance igual de seleção. Portanto, nas discussões precedentes, assinalamos que os vários procedimentos de amostragem resultam numa chance igual de seleção — mesmo que a probabilidade última de seleção seja o produto de várias probabilidades parciais.

Entretanto, mais genericamente, uma amostra probabilística é aquela em que cada elemento da população tem uma probabilidade *não-zero conhecida* de seleção, mesmo que elementos diferentes tenham probabilidades diferentes. Caso se tenha usado procedimentos probabilísticos controlados de amostragem, qualquer uma dessas amostras será representativa da população da qual é retirada se cada elemento da amostra receber um peso igual ao inverso da sua probabilidade de seleção. Portanto, no caso de todos os elementos da amostra terem a mesma chance de seleção, cada um recebe o mesmo peso: 1. (É chamada de amostra "autoponderada".)

Às vezes, é correto fazer amostragem e *ponderação* desproporcionais. Por exemplo, você pode amostrar subpopulações desproporcionalmente para garantir número suficiente de casos de cada subpopulação para análise.

Suponha que você está planejando um *survey* por amostragem de 1.000 residências numa cidade e que a cidade tem uma área decadente onde mora um quarto da população total, e você quer analisar detalhadamente essa área. Uma amostra representativa da cidade lhe daria aproximadamente 250 residências nessa região e 750 no resto da cidade. Mas, 250 casos podem não ser suficientes para a análise que você planeja. Uma solução é selecionar o mesmo número de residências (500) da região decadente que o do resto da cidade. Então, as residências da região decadente teriam uma chance desproporcionalmente maior de seleção em relação às de outras partes da cidade.

Contanto que você analise as duas amostras de áreas separadamente ou comparativamente, não precisa se preocupar com a amostragem diferencial. Mas, se quiser combinar as duas amostras para criar um quadro composto da cidade toda, leve a amostragem desproporcional em consideração através de um procedimento de ponderação.

Eis uma solução simples e intuitiva. Como a região decadente tem um quarto das residências da cidade, e a região

não-decadente, os demais três quartos, as 500 residências selecionadas da região não-decadente têm representação na cidade três vezes maior do que as 500 selecionadas da região decadente. Portanto, a solução de ponderação mais simples é contar cada residência selecionada na região não-decadente como *três* residências. Este procedimento dá a impressão de que você teria entrevistado 1.500 moradias na região não-decadente da cidade, que é o número que você teria selecionado e entrevistado se quisesse 500 moradias na região decadente e tivesse recursos para uma amostra total de 2.000.

Às vezes é possível fazer cópias dos arquivos de dados em certos casos — as residências da região não-decadente, neste exemplo. No entanto, alguns programas de análise de dados permitem especificar os pesos a serem atribuídos aos vários casos, e o computador os leva em consideração ao calcular as estatísticas na análise.

A amostragem desproporcional e os procedimentos de ponderação são às vezes usados em situações que envolvem os erros e as aproximações muitas vezes inerentes à amostragem complexa em múltiplas etapas. Isto será discutido em mais detalhes no exemplo de amostra por conglomerados por área no Capítulo 6, mas convém mencionar algumas das condições nas quais a ponderação é freqüentemente necessária.

1. Na amostragem estratificada por conglomerados, um determinado número de conglomerados pode ser selecionado de cada estrato, embora varie o tamanho dos diversos estratos. A ponderação diferencial pode ser usada para compensar aquelas variações.
2. Um certo conglomerado pode ser selecionado num desenho de amostra PPT com base no tamanho esperado ou estimado, mas uma investigação de campo posterior pode indicar que a estimativa inicial estava errada. Portanto, o conglomerado terá recebido uma chance de seleção desproporcionalmente alta ou baixa; a ponderação pode ser usada para compensar este erro.
3. Um desenho de amostra pode pedir a seleção de um décimo dos elementos de um conglomerado, mas este pode conter apenas cinquenta e dois elementos, apenas cinco dos quais sendo selecionados para estudo. Pode-se usar a ponderação para compensar o elemento 0,2 que, logicamente, não tinha como ser selecionado.

4. Dez elementos de um conglomerado podem ter sido selecionados, mas dois não tiveram como ser estudados (por exemplo, recusaram a entrevista). Pressupondo homogeneidade dentro do conglomerado, você pode atribuir peso 1,25 a cada um dos elementos estudados para compensar os dois que não o foram.

Todos estes usos de ponderação são ilustrados no exemplo final de amostragem do Capítulo 6. Entretanto, com a exceção do caso 4 acima, você pode derivar seu próprio procedimento de ponderação, determinando cuidadosamente a probabilidade de seleção, passo a passo, para cada elemento da amostra, e atribuindo um peso igual ao inverso daquela probabilidade. Três outros aspectos da ponderação devem ser discutidos antes de passarmos a outro tema.

Graus de Precisão na Ponderação

Em qualquer desenho complexo de amostra, há diversas opções em relação à ponderação, em conexão com a amostragem intencional e/ou inadvertidamente desproporcional. Podemos computar pesos para cada elemento até várias casas decimais, ou atribuir pesos aproximados para dar conta apenas das instâncias mais grosseiras de amostragem desproporcional. No caso anterior da cidade com a região suburbana sobre-amostrada, é improvável que a população daquela região seja exatamente um quarto da população da cidade; pode de fato ser 0,25001, 0,2600 ou 0,2816 da população total. No primeiro caso, provavelmente escolheremos aplicar a ponderação geral aproximada de casos descrita, se não houver outra amostragem desproporcional envolvida noutras etapas do desenho da amostra. Talvez também façamos o mesmo no segundo e terceiro casos. A precisão que buscamos na ponderação deve ser comensurável com a precisão que queremos nos nossos achados. Se nossos objetivos de pesquisa podem tolerar erros de poucos pontos percentuais, provavelmente não perderemos tempo e esforço fazendo ponderações exatas. Além disso, ao decidirmos o grau de precisão requerido, devemos levar em conta o grau de erro a ser esperado da distribuição normal de amostragem mais todos os diversos tipos de erro não-amostrais.

Em última análise, não há uma diretriz firme a seguir quando se determina a precisão a ser buscada na ponderação. Como em vários outros aspectos do desenho de *surveys*, desfrutamos de considerável latitude. Ao mesmo tempo, porém, devemos lembrar-nos da nossa decisão de ponderação ao informarmos

os resultados. Não devemos empregar um procedimento de ponderação aproximada e depois sugerirmos que os achados são precisos dentro de uma faixa minúscula de erro.

Métodos de Ponderação

Tendo delineado as preocupações científicas quanto a determinar o grau de precisão desejado na ponderação, devemos assinalar que a escolha muitas vezes é feita com base nos métodos disponíveis para ponderar. Os três métodos básicos para a ponderação são:

1. Para a ponderação aproximada de amostras extraídas de subpopulações, tabelas com pesos podem ser construídas das tabelas sem pesos, para cada uma das subamostras. No exemplo acima, você pode criar uma tabela simples de distribuições para a amostra suburbana e para a amostra não suburbana separadamente, triplicar o número de casos em cada célula da tabela não suburbana, somar os casos nas duas tabelas e computar as porcentagens para a tabela composta.
2. Para uma ponderação mais extensa e rápida, embora ainda aproximada, os conjuntos de dados para certos casos podem ser clonados. No exemplo anterior, podem ser feitas duas cópias adicionais para cada residência suburbana (totalizando três), o conjunto ampliado de casos não suburbanos podendo então ser combinado com os casos suburbanos, e o conjunto ampliado de dados pode ser analisado como se três vezes mais casas não suburbanas tivessem sido estudadas.

3. Muitos programas de computador para análise de dados são desenhados para atribuir um peso exato para cada caso no arquivo de dados original. Este método é o único apropriado para a ponderação refinada, porque é impossível clonar frações de casos, com algum significado.

Como em outras instâncias, também na ponderação as preocupações científicas em geral submetem-se a razões práticas. Se a análise for conduzida, por exemplo, através de cartões IBM, a ponderação deve, necessariamente, ser aproximada em vez de precisa.

Ponderação e Inferência Estatística

Advirto que os procedimentos de ponderação descritos têm efeitos sérios na maioria dos cálculos relacionados à inferência estatística. Se suas finalidades de pesquisa exigem inferências estatísticas precisas (por exemplo, estimativas de população), feitas com base em dados cuidadosamente ponderados, você deve consultar uma fonte especial sobre este assunto⁷ ou, melhor ainda, consultar um estatístico de amostragem *antes da amostragem ser desenhada*.

Panorama da Amostragem Probabilística

A longa e detalhada discussão precedente foi dedicada ao método chave de amostragem usado na pesquisa de *survey*: amostragem probabilística. Em cada uma das variações examinadas, vimos que os elementos são escolhidos de uma população, com base na seleção aleatória e com probabilidades não-zero conhecidas.

Dependendo da situação de campo, a amostragem probabilística pode ser muito simples ou extremamente difícil, demorada e cara. No entanto, independentemente da situação, continua sendo o método mais eficiente para a seleção de elementos para estudo por duas razões. Primeiro, a amostragem probabilística evita o VIÉS consciente ou inconsciente na seleção de elementos por parte do pesquisador. Se todos os elementos da população têm chance igual (ou desigual, com ponderação posterior) de seleção, há uma boa chance de a amostra assim selecionada representar bastante bem a população. Segundo, a amostragem probabilística permite estimativas de erro amostral. Embora nenhuma amostragem probabilística seja perfeitamente representativa em todos os aspectos, métodos controlados de seleção permitem estimar o grau de erro esperado.

Tendo discutido a amostragem de probabilidade em algum detalhe, examinemos agora, mais sucintamente, alguns métodos comuns de amostragem não-probabilística.

Amostragem Não-Probabilística

Apesar da superioridade dos métodos de amostragem probabilísticos na pesquisa de *survey*, às vezes usam-se métodos não-probabilísticos em seu lugar, em geral em situações em que a amostragem probabilística seria dispendiosa demais

e/ou quando a representatividade exata não é necessária. Os métodos principais da amostragem não-probabilística são os seguintes.

Amostragem Intencional ou por Julgamento

Ocasionalmente, pode-se selecionar a amostra baseado no próprio conhecimento da população e dos seus elementos, e da natureza das metas de pesquisa. Este método de amostragem é às vezes chamado *amostragem intencional* ou *por julgamento*. Principalmente no desenho inicial do questionário, você pode querer selecionar maior variedade de entrevistados para testar a aplicabilidade geral das perguntas. Embora os achados do *survey* não representem qualquer população significativa, o teste pode efetivamente descobrir algum defeito peculiar no seu instrumento de pesquisa. No entanto, deve-se referir a esta situação como um pré-teste e não um *survey* propriamente.

Algumas vezes, você quer estudar um pequeno subconjunto de uma população maior, onde muitos membros do subconjunto são facilmente identificados, mas cuja enumeração total seria quase impossível. Por exemplo, estudar a liderança de um movimento de protesto estudantil; muitos dos líderes são facilmente visíveis, mas não é viável definir e amostrar todos eles. Ao estudar todos ou uma amostra dos líderes mais visíveis, você pode coletar dados suficientes para seus propósitos.

Num desenho de amostra em etapas múltiplas, você pode querer comparar estudantes liberais e conservadores. Uma vez que, provavelmente, não é possível enumerar e amostrar todos os estudantes, você pode optar por amostrar as filiações dos Jovens Democratas e dos Jovens Republicanos. Embora tal desenho de amostra não forneça uma boa descrição dos estudantes liberais ou dos conservadores como um todo, pode ser suficiente para propósitos de comparação.

A *amostragem de zonas eleitorais selecionadas* para pesquisas de opinião política é um processo bem refinado de julgamento. Com base em resultados eleitorais anteriores em alguma área (cidade, estado, nação), você seleciona intencionalmente um grupo de distritos ou zonas eleitorais que, combinados, produzem resultados similares aos de toda a região. Então, em pesquisas posteriores, você só seleciona suas amostras naquelas zonas eleitorais. Naturalmente, a teoria diz que as zonas eleitorais selecionadas oferecem um corte de todo o eleitorado.

Cada vez que há uma eleição, permitindo avaliar a adequação do seu grupo de zonas eleitorais, você deve fazer revisões, adições ou retiradas. A meta é atualizar o grupo de zonas eleitorais, garantindo a boa representação de todas elas. Para ser efetiva, a amostragem de zonas eleitorais requer grande conhecimento político. Deve haver intimidade com a história política e social da área considerada, de modo a basear a seleção de zonas eleitorais num palpite *com conhecimento* quanto à representatividade. Ademais, este sistema de amostragem requer retroalimentação contínua para ser efetiva. Deve-se poder conduzir pesquisas frequentes e fazer validações eleitorais periódicas.

Amostragem por Cotas

Como disse antes, a amostragem por cotas começa com uma matriz descrevendo as características da população alvo. Por exemplo, você deve saber qual proporção da população é masculina e qual é feminina; para cada sexo, saber qual é a proporção nas várias categorias de idade, e assim por diante. Ao estabelecer uma amostra nacional por cotas, você deve saber qual proporção da população nacional é, digamos, urbana, do leste, masculina, abaixo de vinte e cinco anos, branca, da classe operária, e outros semelhantes, bem como todas as outras permutações nessa matriz.

Após a criação da matriz, com uma proporção relativa atribuída a cada célula, colhem-se dados de pessoas com todas as características de uma dada célula. Todas as pessoas nesta célula, então, recebem um peso apropriado à sua porção na população total. Quando todos os elementos da amostra estiverem ponderados assim, os dados como um todo devem resultar numa representação razoável da população total.

Há vários problemas inerentes à amostragem por cotas. Primeiro, a moldura das cotas (a proporção que as diferentes células representam) deve ser precisa, e muitas vezes é difícil obter informação atualizada para este fim. A falha do Gallup em prever a vitória de Truman em 1948 deveu-se em parte a este problema.

Segundo, pode haver vieses na seleção dos elementos da amostra de uma célula, mesmo que se estime com precisão sua proporção na população. Um entrevistador instruído a entrevistar cinco pessoas que se encaixam num dado conjunto complexo de características pode, ainda assim, evitar pessoas que moram nos andares de cima de um prédio com sete andares de escadas, ou pessoas em habitações muito precárias e/ou com cães ferozes.

Muitas vezes se usa uma variante da amostragem por cotas em conexão aos *grupos focais* na pesquisa de mercado. Nesta técnica, grupos de aproximadamente doze pessoas são, tipicamente, reunidos para discussões focalizadas sobre produtos ou comerciais. Recomenda-se reunir vários grupos de tipos diferentes de pessoas ou misturas específicas de pessoas.

Confiança em Sujeitos Disponíveis

Pesquisadores universitários freqüentemente fazem *surveys* com estudantes matriculados em grandes salas de aula. A facilidade e baixo custo deste método explica sua popularidade, mas estes *surveys* raramente produzem dados com algum valor genérico. Embora sirva à finalidade de pré-testar um questionário, este método de amostragem não deve ser usado para um estudo pretendendo descrever os alunos como um todo.

Parar pessoas numa esquina ou outro lugar quase nunca é um método de amostragem adequado, embora seja empregado com demasiada freqüência, como no caso de pesquisas de mercado em estacionamentos de supermercados. Este *survey* só se justifica se você quiser estudar as características de pessoas passando naquele ponto em particular em horas determinadas.

O termo "entrevista por interceptação em *shoppings*" às vezes é usado em referência a um *survey* no qual consumidores andando em *shoppings* são interceptados e solicitados a participar. Obviamente, é concebível que tais respondentes potenciais seriam apropriados a alguns *surveys* específicos — um estudo de consumidores daquele *shopping*, por exemplo. Entretanto, nestes casos, outro problema de amostragem é que a maior parte daqueles que são abordados se recusa a participar. Mas, num estudo num *shopping* em Chicago, Hornik e Ellis observaram que as taxas de resposta crescem quando os entrevistadores olhavam e tocavam os entrevistados ao pedir sua cooperação, especialmente se a entrevistadora era mulher.⁸

Mencionei esta pesquisa para indicar que, embora os métodos de amostragem probabilística convencional sejam quase sempre preferíveis na pesquisa de *survey*, há momentos em que outros métodos podem ser preferíveis ou necessários. Mas escolher um modelo de amostragem não-probabilístico jamais deve ser desculpa para a preguiça. Pesquisadores de *survey* devem encontrar formas de providenciar uma amostra representativa da população que querem estudar.

Usos Não-Survey dos Métodos de Amostragem

Nesta discussão sobre a lógica e os métodos de amostragem, focalizamos a pesquisa de *survey* — seleção de elementos de uma população para entrevista ou questionários auto-administrados. No entanto, a lógica básica da amostragem dá à discussão precedente um valor mais geral para o pesquisador social que usa outros métodos de coleta de dados. Fecharemos o capítulo com breves comentários sobre os usos não-*survey* dos métodos de amostragem.

Análise de Conteúdo

Um analista de conteúdo codifica e analisa documentos para fazer afirmações descritivas ou explicativas sobre a literatura composta pelos documentos, seu(s) autor(es) e/ou o meio social do qual fazem parte os documentos. Ele pode analisar os jornais de um país, os trabalhos de ficção de um romancista, a linguagem de projetos legislativos, e assim por diante.

Muitas vezes o volume de documentos para analisar é grande demais para uma cobertura completa. Neste caso, as técnicas de amostragem discutidas acima podem ser facilmente adaptadas. As unidades de amostragem podem ser palavras individuais, frases, parágrafos, artigos, livros etc. As unidades de amostragem podem ser estratificadas de maneira apropriada, e amostras aleatórias, sistemáticas ou mesmo por conglomerados ser selecionadas.

Experimentos de Laboratórios

Sujeitos para experiências de laboratórios muitas vezes são selecionados entre voluntários que respondem a um anúncio. Ocasionalmente, o projeto experimental requer emparelhamento dos sujeitos nos grupos experimental e de controle. Em alguns casos, estabelecem-se cotas para tipos diferentes de sujeitos. Sempre que o número de sujeitos em potencial exceder muito o número requerido pelo experimento, pode-se usar técnicas padrão de amostragem. Métodos de estratificação podem ser usados como auxiliares para satisfazer os requisitos das cotas.

Observação Participante

Diferentemente dos pesquisadores de *survey*, observadores participantes tipicamente buscam mergulhar na totalidade

do fenômeno sob estudo. Podem tentar observar todos os grandes eventos que estão ocorrendo, falar com o maior número de participantes possível, e assim por diante. No entanto, obviamente ninguém consegue observar tudo; alguma seletividade é inevitável. Se tal seletividade for incontrolada, há o risco de reunir um conjunto de observações tendenciosas, tal como um pesquisador de *survey* inepto ao selecionar uma amostra tendenciosa de entrevistados.

Embora não queira sugerir que todos os observadores participantes em todos os lugares em todas as situações de pesquisa devam buscar estabelecer métodos de amostragem rigorosos, estou convencido de que um observador participante bem versado na lógica e, talvez, em alguns dos métodos da amostragem de *survey* tem maior probabilidade de obter um conjunto representativo de observações. Ao avaliar o humor de estudantes reunidos para uma demonstração de protesto, um observador participante deve ter o cuidado de não falar apenas com os líderes da demonstração, devendo considerar a possível diferença entre estudantes que aparentemente parecem sozinhos e os que vão com amigos, conversar com estudantes em locais diversos da reunião e procurar conversar com os que chegam cedo e com os que chegam atrasados. Embora o observador não tenha condições de estratificar de forma rigorosa a amostra de entrevistados, ele pode ficar mais sensível a reparar diferenças relacionadas às variáveis citadas, refinando as observações de forma correspondente.

Resumo

Este longo capítulo buscou familiarizá-lo com as considerações mais importantes na lógica da amostragem de *survey* e com as técnicas mais comuns de amostragem. Esta discussão não basta para equipá-lo para cada situação de campo que você enfrentará, mas espero que uma compreensão da lógica da amostragem lhe permita chegar a julgamentos razoáveis e seguros por conta própria.

O capítulo seguinte descreve quatro amostras de *survey* com as quais estou intimamente familiarizado. Ao apresentar detalhes específicos destes desenhos de amostras um tanto diferentes e das muitas decisões tomadas na sua execução, espero dar-lhe alguma experiência concreta de amostragem. Repito que estes exemplos não podem ser exaustivos em relação às situações com as quais você se deparará, mas eles foram escolhidos para representar as situações de amostragem mais comuns encontradas na pesquisa de *survey*.

Notas

¹ Para uma discussão recente, ver SQUIRE, Peverill. Why the 1936 *Literary Digest* Poll Failed. *Public Opinion Quarterly*, v.52, p.125-133, primavera 1988.

² Citada na *Newsweek*, p.24, 8 de julho de 1968.

³ Logo veremos que o tamanho da amostra selecionada e as características reais da população maior afetam o grau de representatividade.

⁴ KISH, Leslie. *Survey Sampling*. New York: John Wiley & Sons, 1965.

⁵ Evidentemente, um guia telefônico não inclui novos assinantes nem quem prefere não ter o número listado. A amostragem se complica mais com a inclusão de listagens não residenciais, como lojas e escritórios.

⁶ Veja, por exemplo, GLOCK, Charles Y., RINGER, Benjamin B., BABBIE, Earl R. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967. Apêndice A.

⁷ Por exemplo, KISH. *Survey Sampling*.

⁸ HORNIK, Jacob, ELLIS, Shmuel. Strategies to Secure Compliance for a Mall Intercept Interview. *Public Opinion Quarterly*, v.52, p.539-551, inverno 1988.

Leituras Adicionais

BABBIE, Earl. *The Practice of Social Research*. Belmont, CA: Wadsworth, 1989. cap.7.

COCHRAN, William G. *Sampling Techniques*. New York: John Wiley & Sons, 1963.

FRANKEL, Martin R., FRANKEL, Lester R. Fifty Years of *Survey Sampling* in the United States. *Public Opinion Quarterly*, v.51, p.S127-S138, inverno 1987.

HANSEN, M. H., HURWITZ, W. N., MADOW, W. G. *Sample Survey Methods and Theory*. New York: John Wiley & Sons, 1953. 2 V.

KISH, Leslie. *Survey Sampling*. New York: John Wiley & Sons, 1965.

SCHEAFFER, Richard L., MENDENHALL, William, OTT, Lyman. *Elementary Survey Sampling*. North Scituate, MA: Duxbury Press, 1979.

Capítulo 6

Exemplos de Desenhos de Amostragem

O Capítulo 5 apresentou a lógica básica da amostragem de *survey* e delineou algumas opções de procedimentos possíveis. Este capítulo apresenta quatro estudos de caso de desenhos de amostras que representam diferentes situações e desenhos de amostragem.

O primeiro exemplo é uma amostragem estratificada sistemática de estudantes da Universidade do Havai. O segundo é uma amostragem por conglomerados de professores de escolas de medicina em todo o país, atribuindo igual probabilidade de seleção às unidades primárias de amostragem. O terceiro é uma amostra por conglomerados de mulheres de igrejas episcopais no norte da Califórnia, usando um projeto PPT (probabilidade proporcional ao tamanho) para a seleção de unidades primárias de amostragem. O último exemplo é uma amostra complexa de área, projetada para um *survey* de residências em Oakland, Califórnia.

Amostragem de Estudantes Universitários

A finalidade era fazer um *survey*, usando um instrumento auto-administrado, de uma interseção representativa de estudantes do campus principal da Universidade do Havai. As próximas seções descrevem os passos e as decisões envolvidas na seleção desta amostra.

População de Amostra e Moldura de Amostragem

A moldura de amostragem óbvia era o registro magnético em fita usado pela administração da universidade, com os nomes, os endereços atual e permanente, os números da Previdência Social dos alunos e outras informações como área de estudos, período, idade, sexo e assim por diante.

Mas a fita tinha arquivos de todas as pessoas que, por qualquer definição concebível, poderiam ser chamadas de estudantes. Como vários destes alunos pareciam inadequados para os fins do estudo, foi necessário definir a *população do survey* de forma mais restrita. A definição final incluía os 15.225 alunos de graduação matriculados no campus Manoá da universidade, de todas as faculdades e departamentos, em graduação e pós-graduação, e alunos americanos e estrangeiros. O programa usado para a amostragem limitou o estudo a alunos enquadrados nesta definição.

Estratificação

O programa de amostragem também permitia a estratificação de estudantes antes da seleção da amostra. Decidiu-se que a estratificação por período seria suficiente, embora os estudantes pudessem ter sido ainda mais estratificados dentro de cada período, por sexo, faculdade, área de estudo etc.

Seleção da Amostra

Uma vez agrupados por período (pelo programa de amostragem), extraiu-se uma amostra sistemática a partir da lista reordenada dos estudantes. O tamanho da amostra foi inicialmente definido em 1.100. Para extrair a amostra, o programa foi instruído a usar uma fração amostral de 1/14. Portanto, o programa gerou um número aleatório entre 1 e 14; o estudante com este número foi selecionado, bem como todo décimo quarto estudante daí por diante.

Depois de extrair a amostra desta forma, o computador foi instruído a imprimir o nome e o endereço de cada estudante em seis etiquetas postais auto-adesivas, que foram coladas nos envelopes usados para enviar os questionários.

Modificação da Amostra

A descrição precedente detalha o desenho inicial da amostra para o estudo de estudantes universitários. Antes de enviar os questionários, descobriu-se que custos inesperados na produção dos questionários tornavam impossível cobrir as despesas de

correio para 1.100 estudantes. Portanto, um terço das etiquetas postais foram selecionadas sistematicamente (com início aleatório) para exclusão da amostra. A amostra final foi reduzida para cerca de 770.

Esta modificação da amostra é mencionada para ilustrar a necessidade freqüente de mudar aspectos da pesquisa no meio do caminho. Como uma amostra sistemática de estudantes foi eliminada da amostra sistemática inicial, os 770 restantes ainda podiam ser considerados como representando razoavelmente a população do *survey*. Naturalmente, a redução no tamanho da amostra aumentou a extensão do erro amostral.

Amostragem dos Professores de Escolas de Medicina

Esta seção relata o desenho de amostra usado para selecionar uma amostra de professores de escolas de medicina, para um *survey* nacional estudando os efeitos da orientação científica sobre o cuidado humano de pacientes. O desenho da pesquisa demandava uma amostra de professores de escolas de medicina, nos departamentos de medicina e pediatria.

Em condições ideais, uma lista única de todos os professores nos dois departamentos teria sido obtida ou construída, e uma amostra selecionada a partir dela. Infelizmente, esta lista parecia não existir, de forma que se decidiu selecionar uma amostra por conglomerados em duas etapas. Na primeira, fez-se uma amostra de escolas de medicina; a seguir, professores foram escolhidos de cada uma destas escolas.

Desde o começo, o desenho da amostra foi dificultado pela indisponibilidade de dados. O desenho da pesquisa pedia um exame de professores em tempo integral e em tempo parcial. Embora, na época do estudo, houvesse cerca de 3.700 professores em tempo integral nos dois departamentos, nacionalmente, não havia dados bons disponíveis quanto aos números de professores em tempo parcial. Uma análise dos dados existentes sugeriu que o número total de membros em horário integral e parcial estava em torno de 12.000. Para fins do estudo, decidiu-se que uma amostra de 2.000 seria suficiente (uma fração amostral geral de 1/6).

A Seleção de Escolas de Medicina

Na época da pesquisa, 84 escolas de medicina com cursos de quatro anos de duração pertenciam à Associação Americana de Escolas de Medicina. Estas escolas compuseram a população

do *survey* das escolas, que foram então arranjadas em estratos geográficos e ordenadas por tamanho (número de alunos) dentro dos estratos.

A lista estratificada de escolas foi numerada de 1 a 84 e um número aleatório entre 1 e 6 selecionado (o intervalo amostral). A escola com o número escolhido e cada sexta escola daí por diante foram selecionadas na primeira etapa de amostragem. Foram enviadas cartas aos reitores das escolas selecionadas explicando a finalidade do *survey* e pedindo sua assistência na obtenção de uma lista de professores nos seus departamentos de medicina e pediatria.

Inicialmente, foram selecionadas quatorze escolas. Mas nem todos os reitores quiseram cooperar. À medida que as recusas foram chegando, uma escola alternativa para cada uma foi escolhida da lista, da seguinte maneira: uma escola adjacente à que recusava era escolhida jogando uma moeda.

Seleção de Professores

Logo que reitores concordaram a cooperar com o *survey*, foram compiladas em cada uma listas de todos os professores nos departamentos de medicina e pediatria. Todos estes professores foram incluídos na amostra final, recebendo questionários pelo correio.

Este não foi o melhor desenho de amostra possível. A amostra total de professores foi tirada de um número relativamente pequeno de escolas. Um desenho melhor teria selecionado mais escolas, com menos professores em cada uma. Por exemplo, um terço das escolas e metade dos professores em cada podiam ter sido estudados.

O desenho adotado foi limitado mais por questões administrativas do que científicas. Um estudo-piloto havia mostrado a dificuldade de obter aprovação e cooperação dos reitores. Mesmo quando eles concordavam em cooperar, muitas vezes demoravam a fornecer a lista dos professores. O principal gargalo na amostragem ocorreu neste ponto. Aumentar o número de escolas teria aumentado o tempo e os problemas envolvidos na seleção total da amostra. Apenas por isto, tomou-se a decisão de escolher um sexto das escolas e todo o corpo docente de cada uma.

Amostragem de Mulheres de Igrejas Episcopais

A finalidade deste estudo foi examinar as atitudes de mulheres membros de igrejas na diocese californiana da Igreja Episcopal. Desejava-se uma amostra representativa de todas

as mulheres da igreja na diocese. Como se esperava, não havia uma lista única destas mulheres, de modo que um desenho de amostra em etapas múltiplas foi criado. Na fase inicial de amostragem, as igrejas foram selecionadas; a seguir, mulheres foram selecionadas de cada igreja escolhida. Ao contrário da amostra de escolas de medicina, a de igrejas foi selecionada com *probabilidade proporcional ao tamanho* (PPT).

Seleção das Igrejas

A diocese em questão publica um relatório anual com uma lista das mais ou menos 100 igrejas componentes, com seus respectivos tamanhos em números de membros. A lista constituiu a moldura de amostragem para a primeira etapa.

Desejava-se cerca de 500 respondentes para o estudo, de forma que se decidiu escolher vinte e cinco igrejas com probabilidade proporcional ao tamanho, tirando vinte mulheres de cada igreja. Para isto, a lista de igrejas foi arranjada geograficamente, e foi feita uma tabela semelhante à lista parcial da Tabela 6-1.

TABELA 6-1

Forma usada na lista das igrejas

Igreja	Membros	Membros (acumulados)
Igreja A	3.000	3.000
Igreja B	5.000	8.000
Igreja C	1.000	9.000

Ao lado de cada igreja na tabela, pôs-se seu número de membros, para computar a frequência acumulada ao longo da lista. O total final chegou a cerca de 200.000. O objetivo, neste passo, foi selecionar uma amostra de vinte e cinco igrejas, de tal forma que cada igreja teria uma chance de seleção proporcional ao seu número de membros. Para isto, os totais cumulativos foram usados para criar, para cada igreja, faixas de números iguais aos números de membros daquela igreja. A igreja A na Tabela 6-1 recebeu os números de 1 a 3.000; a igreja B recebeu os números 3.001 a 8.000; a igreja C recebeu 8001 a 9000; e assim por diante.

Selecionando vinte e cinco números entre 1 e 200.000, seria possível selecionar 25 igrejas para a pesquisa. Os 25 números foram selecionados numa amostra sistemática como se segue. O intervalo amostral foi estabelecido em 8.000 (200.000/25), e um início aleatório foi escolhido entre 1 e

8.000. Suponha que o número aleatório escolhido tenha sido 4.538. Como este número estava dentro da faixa de números atribuídos à igreja B (3.001 a 8.000), a igreja B foi selecionada.

Incrementos de 8.000 (o intervalo amostral) foram somados ao início aleatório, e cada igreja em cuja faixa aparecia um dos números resultantes era selecionada para a amostra de igrejas. Deve estar claro que, assim, cada igreja na diocese tinha uma chance de seleção diretamente proporcional ao número de membros. Uma igreja com 4.000 membros tinha duas vezes mais chance de seleção do que uma igreja com 2.000 membros e dez vezes mais do que uma igreja de 400 membros.

Selecionando as Mulheres das Igrejas

Depois de selecionar a amostra de igrejas, obtiveram-se listas das mulheres de cada igreja. Vale notar aqui que, na prática, as listas variavam enormemente em forma e conteúdo. Em vários casos, foram fornecidas listas de todos os membros (homens e mulheres), tornando necessário separar as mulheres antes de amostrar. A forma das listas variava desde listas datilografadas até cartões 3x5 impressos a partir de placas de *addressograph*.

A cada lista que chegava de uma igreja selecionada, computava-se um intervalo amostral para ela, com base no número de mulheres membros e no número desejado (20). Se uma igreja tinha 2.000 mulheres, o intervalo amostral era estabelecido em 100. Um número aleatório era escolhido e incrementado no valor do intervalo amostral, para selecionar a amostra de mulheres daquela igreja. Este procedimento foi repetido para cada igreja.

Observe que este desenho de amostra só dá a cada mulher na diocese uma chance igual de seleção se metade dos membros de cada igreja forem mulheres (ou se uma proporção constante dos membros são mulheres), porque as igrejas tinham chance igual de seleção baseada no número total de membros (não estavam disponíveis números só de mulheres). Dadas as metas do estudo, as leves desigualdades de seleção foram consideradas insignificantes.

Um desenho de amostra mais sofisticado para a segunda etapa teria resolvido este possível problema. Como se deu a cada igreja uma chance de seleção baseada num número presumido de mulheres (supondo-se, por exemplo, 1.000 mulheres numa igreja com 2.000 pessoas), o intervalo amostral podia ter sido computado com base naquela suposição e não no número

real de mulheres listadas. Se tivesse sido suposto, na primeira etapa de amostragem, que uma igreja tinha 1.000 mulheres (de um total de 2.000 membros), o intervalo amostral podia ter sido estabelecido em 50 (1.000/20). Este intervalo podia ter sido usado na seleção de respondentes, independentemente do número real de mulheres listadas naquela igreja. Se, de fato, 1.000 mulheres estavam listadas, a igreja delas teria a chance apropriada de seleção e 20 mulheres seriam selecionadas. Se 1.200 mulheres estavam listadas, a igreja tinha uma chance muito pequena de seleção, mas este desequilíbrio teria sido remediado através da seleção de 24 mulheres pelo uso do intervalo de amostragem preestabelecido. Se, por outro lado, somente 800 mulheres estavam listadas, apenas 16 teriam sido selecionadas.

Amostragem de Casas em Oakland

Este exemplo final representa um dos desenhos de amostra mais complexos da pesquisa de *survey*: uma amostra de área por conglomerados. A finalidade deste estudo, feito em 1966, foi coletar dados relevantes ao estudo da pobreza nas regiões mais pobres de Oakland, Califórnia, usando o restante da cidade para fins de comparação. Uma vez que se ia usar os achados do *survey*, em parte, para apoiar solicitações de recursos federais para Oakland, era essencial que os dados coletados descrevessem precisamente a cidade.

Para fins do estudo, a cidade foi dividida em sete áreas: quatro eram oficialmente designadas áreas pobres, e as três restantes eram vistas tradicionalmente como seções distintas da cidade. O tamanho total da amostra foi estabelecido em 3.500 residências, com base em cálculos cuja complexidade excede o escopo deste livro. Entretanto, note-se que a escolha do tamanho da amostra começou com discussões de políticas referentes à "chance" que as autoridades municipais estavam dispostas a correr de que o *survey*, pelo erro amostral normal, subestimasse os níveis de pobreza e desemprego a ponto de não qualificar a demanda de financiamento federal. Se estes níveis fossem, de fato, altos o bastante para garantir o financiamento, uma amostra perfeitamente precisa demonstraria isto. Mas, com uma amostra pequena, a faixa de erro amostral abria a possibilidade de subestimação. Naturalmente, com o aumento do tamanho da amostra, a chance de subestimação reduzia-se. Portanto, desta maneira estabeleceu-se o tamanho final da amostra em 3.500 residências. (*Nota*: É desta forma que deve ser estabelecido o tamanho das amostras.)

Como o estudo pedia uma comparação das sete regiões da cidade, era importante descrevê-las com igual grau de precisão. Portanto, a amostra foi projetada de forma a selecionar 500 residências de cada região, embora as sete áreas diferissem muito no número total de residências. O restante desta discussão é dedicado aos procedimentos de seleção da amostra usados numa das áreas, já que os procedimentos foram essencialmente idênticos em todas.

Considerações Gerais

Como dissemos, planejamos uma amostragem de área por conglomerados. No primeiro estágio de amostragem, uma amostra estratificada de *quarteirões de recenseamento*¹ devia ser selecionada com probabilidade proporcional ao tamanho. Os quarteirões selecionados seriam listados da seguinte forma: enumeradores visitariam fisicamente cada quarteirão selecionado e preparariam listas de todas as residências ali achadas; então, uma amostra sistemática de cinco residências seria selecionada de cada quarteirão.

Como dissemos, cinco residências por quarteirão é uma convenção comum na amostragem de área por conglomerados. Tanto a teoria de amostragem quanto a experiência de *survey* sugerem que uma amostra de cinco residências representa o ponto de retornos decrescentes na descrição de quarteirões. Cinco residências permitem descrever um quarteirão de forma razoavelmente precisa, dada a homogeneidade tipicamente encontrada entre residências num único quarteirão. Embora, é claro, seis residências dessem uma descrição mais precisa, a redução correspondente na precisão da amostra dos quarteirões, selecionados para descrever a população de todos os quarteirões, anularia esta vantagem. (Supondo que o tamanho total da amostra é fixo.) Num exemplo extremo, você poderia limitar sua amostra a todas as residências encontradas num único quarteirão; ela permitiria uma descrição bem precisa do quarteirão, mas o próprio quarteirão não permitiria descrever muito precisamente todos os quarteirões e, por extensão, todas as residências da cidade.

Selecionar 500 residências de cada região da cidade significava selecionar 100 quarteirões, com cinco residências em cada um. Para aumentar a representatividade da amostra de residências, decidiu-se organizar os quarteirões em estratos relativamente homogêneos (grupos de quarteirões) e selecionar dois quarteirões de cada estrato. Portanto, a tarefa inicial em cada

região foi criar cinquenta estratos; a seguir, dois quarteirões foram selecionados de cada estrato, e, finalmente, cinco casas em cada quarteirão — totalizando 500 residências em cada região.

A Moldura de Amostragem da Primeira Etapa

Foram necessários dois tipos de dados para descrever os quarteirões de recenseamento. O mais importante era saber quantas residências havia aproximadamente em cada um, de forma que se pudesse selecionar uma amostra PPT. Para fins de estratificação, também era preciso conhecer algumas características relevantes dos quarteirões, como sua composição racial e seus níveis socioeconômicos.

As estatísticas dos quarteirões de recenseamento de 1960 dos Estados Unidos ofereciam os dois tipos de dados. Este arquivo indicava o tamanho de todos os quarteirões em 1960 (em número de unidades domiciliares) e também fornecia descrições variadas, como (1) porcentagem de não-brancos, (2) porcentagem de ocupantes pagando aluguel, (3) porcentagem em deterioração, e (4) valor das estruturas (seja em aluguel ou em avaliação). Infelizmente, o estudo foi feito seis anos depois da coleta de dados do censo, de forma que antecipou-se que várias das contagens de casas estariam desatualizadas.

Antes de começar a amostragem, gastaram-se dois meses examinando os mapas de planejamento de cada área. Unidades domiciliares construídas ou demolidas desde o censo de 1960 foram anotadas, e estes dados usados para corrigir o número estimado de residências por quarteirão. Enfim, preparou-se um registro de dados para cada quarteirão, mostrando seu tamanho atual esperado (em número de residências) e as diversas características compiladas no censo de 1960 (que não seriam atualizadas).

As seções seguintes descrevem os procedimentos de seleção da amostra usadas na Área de Pobreza D, que continha 9.938 residências estimadas na época da amostragem. Já que a meta de amostragem para a área era de 500 casas, a fração total de amostragem foi arredondada para 1/20.

Seleção de Grandes Quarteirões

Cada área da cidade tinha algum quarteirão de recenseamento com um grande número de residências. Alguns eram grandes em tamanho geográfico; outros tinham vários prédios grandes de apartamentos. Como a presença destes conjuntos

habitacionais podia prejudicar as técnicas de estratificação planejadas, e pelo fato destes quarteirões serem considerados muito importantes para os propósitos do estudo, decidiu-se que cada quarteirão com 200 ou mais residências seria automaticamente incluído na amostra. Estes quarteirões seriam listados, e um vigésimo (taxa de amostragem total para a região) das residências listadas selecionada para entrevistas.

Na Área de Pobreza D, os quarteirões maiores tinham um total de 702 residências. Portanto, trinta e cinco residências foram selecionadas daqueles quarteirões, deixando 465 para serem selecionadas do restante da área.

Lidando com Quarteirões Pequenos e Quarteirões Zero

Vários quarteirões na Área de Pobreza D (e em outros locais) tinham poucas residências, estimando-se que alguns não tinham nenhuma. Estes quarteirões apresentavam problemas especiais. Primeiro, um quarteirão com tamanho esperado igual a zero não teria chance de seleção. Se as estimativas estivessem incorretas e o quarteirão, de fato, tivesse residências, elas não teriam qualquer chance de seleção. Segundo, uma vez que o desenho de amostra básico pedia a seleção de cinco residências de cada quarteirão selecionado, aqueles com menos de cinco também representaram um problema. Além disso, se as estimativas de tamanho estivessem incorretas, alguns quarteirões que se acreditava ter mais do que cinco casas podiam, de fato, ter menos.

Para resolver estes problemas, cada quarteirão zero e cada quarteirão com um tamanho esperado de menos que dez residências foi "anexado" a um quarteirão adjacente. O número de residências esperado num pequeno quarteirão, se houvesse algum, era somado ao número do quarteirão adjacente, e o par de quarteirões tratado como um quarteirão único para fins de amostragem. O par tinha uma chance de seleção proporcional ao seu tamanho combinado; se o par fosse selecionado na amostra de quarteirões, ambos eram listados e amostrados como se fossem um único quarteirão.

Estratificação

Uma vez que trinta e cinco residências seriam selecionadas de grandes quarteirões, 465 precisavam ser selecionadas do restante da área, com seu total estimado de 9.236 residências. Com cinco residências a serem selecionadas de cada quarteirão da amostra, noventa e três quarteirões seriam necessários. Como dois

quarteirões seriam selecionados de cada estrato, a tarefa neste ponto era criar quarenta e sete (arredondado de 46,5) estratos.

Cada estrato (grupo de quarteirões) tinha que ter duas características. Primeiro, os quarteirões em cada estrato deviam ser tão semelhantes entre si quanto possível, em termos de composição racial, status socioeconômico (SSE) etc., para garantir que todos os tipos de quarteirões fossem selecionados na proporção adequada na amostra total. Segundo, cada estrato devia conter um total de aproximadamente 200 residências. Já que dez casas deviam ser selecionadas de cada estrato (dois quarteirões, cinco residências de cada), um tamanho de estrato de 200 produziria uma fração amostral total de 1/20, a fração estabelecida para toda a área. Naturalmente, todos os quarteirões tinham números especificados de residências, de forma que não foi possível criar estratos com exatamente 200 casas. (Uma discussão posterior tratará da correção estatística para tais variações.)

Ao criar homogeneidade entre os quarteirões agrupados num estrato, foi possível empregar as características de quarteirões fornecidas no censo de 1960. Para evitar confusão nas descrições empíricas seguintes, observe que a criação de homogeneidade foi feita de forma essencialmente *ad hoc* e arbitrária. Embora possa haver a tentação de criar um formato estratificado em bases teóricas desde o início (por exemplo, todos os quarteirões com mais de 75% de residentes não-brancos, mais de 50% de unidades alugadas e menos de \$150 em renda mensal média seriam combinadas para formar um estrato), esta abordagem não seria necessariamente apropriada à natureza da região ou área específica sendo amostrada. Em vez disso, cada grupo de quarteirões (representando uma região da cidade ou subconjunto dela) era examinado para determinar as variações das suas características, e um sistema de estratificação foi desenvolvido para se adequar àquelas características particulares. Portanto, áreas diferentes da cidade foram estratificadas de forma diferente; ademais, deve-se reconhecer que o formato de estratificação específico a uma região era apenas uma de várias, talvez igualmente apropriadas, possibilidades.

Entretanto, como regra geral, as variáveis de estratificação disponíveis foram consideradas numa determinada ordem para cada área da cidade: crescimento recente, composição racial, porcentagem de casas próprias/alugadas, valor das propriedades e deterioração. Sempre que esta ordem de variáveis de estratificação fosse inapropriada para uma dada área, as variáveis inapropriadas eram ignoradas ou consideradas num ponto diferente da estratificação.²

Portanto, em cada uma das sete regiões da cidade, iniciamos a estratificação com base no crescimento recente. Quando os registros de dados atualizados foram criados para cada quarteirão de recenseamento, foi feita uma anotação quanto ao número de unidades residenciais construídas após o censo de 1960. Como parecia razoável supor que quarteirões contendo novas construções difeririam de outros quarteirões, este dado foi usado como primeira variável de estratificação. Por exemplo, todos os quarteirões na Área de Pobreza D contendo alguma nova unidade foram separados para uma estratificação separada. Juntos, estes quarteirões tinham 1.254 residências estimadas. Com 200 casas por estrato, os quarteirões "com crescimento" foram agrupados em seis estratos.

A segunda variável de estratificação usada foi composição racial. Todos os quarteirões com crescimento foram ordenados em termos da porcentagem de residentes não-brancos vivendo em cada um. Alguns quarteirões tinham 80% ou mais de residentes não-brancos. Juntos, estes quarteirões tinham 247 residências. Uma vez que este número estava relativamente próximo da meta de 200 casas por estrato, estes quarteirões foram designados como o primeiro estrato.

Quarteirões contendo entre 30 e 79% de residentes não-brancos tiveram um total de 385 residências; decidiu-se criar dois estratos destes quarteirões. A variável de estratificação seguinte foi a porcentagem de residências alugadas (em contraposição às próprias). Nos estratos aqui considerados, os quarteirões com 36% ou mais de residências alugadas representavam um total de 214; as 171 restantes estavam em quarteirões com menos de 36% de residências alugadas. Estes dois grupos de quarteirões foram designados como o segundo e o terceiro estrato. (Nota: Não há nada intrinsecamente significativo na escolha do corte em 36%. A meta era criar estratos de tamanho aproximadamente igual, e o ponto de corte que preenchia esta condição foi usado.)

Quarteirões com menos de 30% de residentes não-brancos tinham 622 residências, pedindo três estratos. Examinando estes quarteirões em termos de todas as variáveis de estratificação, descobriu-se que mais de 10% das residências num grupo de quarteirões foram classificadas pelo censo como "em deterioração". Estes quarteirões tinham 214 residências. Como 10% de deterioração é bastante alto no contexto de todos os quarteirões, este grupo foi designado como sendo o quarto estrato.

Os quarteirões restantes tinham 408 residências, pedindo dois estratos. Quando se examinou a ocupação em termos de alugueis, descobriu-se que os quarteirões com 86% ou mais de alugueis tinham 201 residências; estes quarteirões foram

designados como o quinto estrato. Os quarteirões contendo as 207 casas restantes foram designados como o sexto estrato.

Em vista da complexidade deste procedimento, apresentamos na Tabela 6-2 um resumo esquemático da estratificação dos quarteirões com crescimento na Área de Pobreza D. Os grupos de quarteirões designados como estratos são indicados pela anotação S1, S2, e assim por diante. Lembre que a meta da estratificação era criar grupos relativamente homogêneos de quarteirões. Todos os quarteirões no estrato 5 têm em comum: (1) estão localizados na Área de Pobreza D em Oakland; (2) têm menos de 200 residências; (3) experimentaram um crescimento em número de residências desde 1960; (4) têm menos de 30% de residentes não-brancos; (5) têm menos de 10% das suas estruturas em deterioração; e (6) têm mais de 85% de alugueis. A homogeneidade extrema deste grupo de quarteirões é evidente.

Esta discussão está limitada à criação de seis estratos na Área de Pobreza D. Quarenta e um outros estratos foram criados desta maneira na mesma região, e o processo inteiro foi replicado mais seis vezes para as outras regiões da cidade.

TABELA 6-2

Estratificação de quarteirões com crescimento na Área de Pobreza D

Todos os quarteirões com crescimento (1.254 resid.)	S1
80% ou mais não-brancos (247 resid.)	
30-79% não-brancos (385 resid.)	S2
36% ou mais alugadas (214 resid.)	
35% ou mais alugadas (171 resid.)	S3
29% ou menos não-brancos (622 resid.)	
10% ou mais em deterioração (214 resid.)	S4
9% ou menos em deterioração (408 resid.)	
86% ou mais alugadas (201 resid.)	S5
85% ou menos alugadas (207 resid.)	S6

NOTA - resid. = residências. Usei este termo mais comum em lugar do termo tecnicamente correto *unidade habitacional*. Uma unidade habitacional é um quarto ou conjunto de quartos para uso residencial por uma pessoa ou família; uma residência é a pessoa ou grupo de pessoas residindo numa unidade habitacional.

Seleção de Quarteirões dentro dos Estratos

Dois quarteirões foram selecionados de cada estrato, com probabilidade proporcional ao seu tamanho. Usaremos o estrato 1 do exemplo anterior, que tinha 247 residências, para ilustrar o procedimento usado na seleção de quarteirões.

Para começar, a ordem dos quarteirões no estrato 1 foi randomizada. A seguir, a Tabela 6-3 foi criada a partir das estimativas de residências para cada quarteirão. As primeiras duas colunas na tabela identificam os quarteirões pelo número do trajeto do censo e número do quarteirão (dentro de cada trajeto ou trato). A terceira coluna apresenta os tamanhos estimados dos quarteirões; a quarta coluna dá o total cumulativo nos seis quarteirões. A coluna final apresenta a faixa de números assinalados a cada quarteirão com base no seu tamanho.

Como dois quarteirões deviam ser selecionados, o total cumulativo para o estrato (247) foi dividido por 2 (resultando em 123,5). Um número aleatório entre 1 e 123,5 foi selecionado. O quarteirão 27/5, com uma faixa de 13-42, continha este número aleatório, de forma que foi selecionado na amostra de quarteirões. O número aleatório foi, então, adicionado a 123,5. O quarteirão 27/14, com uma faixa de 149-196, continha este novo número (158,5), de forma que foi o segundo quarteirão selecionado na amostra.

TABELA 6-3

Procedimento de seleção dos quarteirões

Traço (trajeto) do censo	Quarteirão do censo	Número de residências	Total cumulativo	Faixa cumulativa
27	18	12	12	1-12
27	5	30	42	13-42 selecionado
27	23	26	68	43-68
28	4	80	148	69-148
27	14	48	196	149-196 selecionado
28	2	51	247	197-247

NOTA - Total resid./2=123,5; nº aleatório entre 1 e 123,5=35; nº aleatório + (total resid./2)=158,5.

Um programa de computador foi projetado para realizar a maior parte dos passos descritos acima. Os quarteirões foram divididos por estratos. O computador lia os relatórios de dados de um estrato, calculava e imprimia os totais cumulativos para aquele estrato, dividia o total por 2 e imprimia aquele número, e finalmente gerava e imprimia um número aleatório entre 1 e metade do total. Depois determinava quais quarteirões deviam ser incluídos na amostra.

Seleção de Residências dentro de um Quarteirão

O procedimento descrito resultou na seleção de cerca de 700 quarteirões de recenseamento em toda a cidade de

Oakland. Cinco residências (em geral) iam ser entrevistadas em cada quarteirão. Para atingir esta meta, mapas foram preparados de forma a identificar claramente cada um dos quarteirões selecionados, e enumeradores foram mandados para preparar uma lista de todas as residências em cada quarteirão. As listas, preparadas pelo uso de formulários padronizados, assemelhavam-se ao exemplo apresentado na Tabela 6-4.

TABELA 6-4

Amostra de folha de listagem

Número de residências	Nome da rua	Número da rua	Número do apartamento ou outra identificação
01	Walnut St.	2301	
02	Walnut St.	2303A	Duplex
03	Walnut St.	2303B	Duplex
04	Tenth Av.	(102?)	Sem número; casa marrom com sebe
05	Tenth Av.	104	Apart. 101
06	Tenth Av.	104	Apart. 102
07	Tenth Av.	104	Apart. 103
*			
*			
*			
47	Ninth Av.	103	
48	Ninth Av.	101	

Para preparar a listagem, o enumerador ia a uma esquina designada do quarteirão e contornava-o até voltar ao ponto de origem, anotando cada residência no formulário à medida que aparecia. Cada residência, seja uma casa com uma única família, um meio duplex ou um apartamento, era anotada separadamente, recebendo um número em ordem seriada contínua.

O processo de listagem forneceu uma nova estimativa do número de residências em cada quarteirão. (Como os enumeradores podiam cometer erros na listagem, este número ainda assim deve ser considerado uma *estimativa*.) Entretanto, lembre que cada quarteirão era selecionado com base em uma estimativa anterior do seu tamanho. Para levar em conta discrepâncias entre estas duas estimativas, empregamos uma técnica mencionada antes na nossa discussão sobre a amostragem de mulheres da igreja episcopal.

O intervalo amostral usado na seleção de residências de um dado quarteirão foi computado com base na estimativa anterior do seu tamanho e nas cinco residências que se pretendia

selecionar. Se o tamanho do quarteirão tivesse sido estimado em cinquenta residências, o intervalo amostral era estabelecido em dez. Este intervalo foi usado na seleção de residências, mesmo se o número de casas de fato listadas fosse maior ou menor. Se o quarteirão tivesse sessenta residências em vez das cinquenta estimadas, seis eram selecionadas na amostra; se tivesse só quarenta, quatro eram selecionadas. (Uma correção mais precisa também foi empregada, o que discutiremos em breve.)

A seleção de residências foi feita escolhendo um número aleatório entre 1 e o intervalo amostral. O número aleatório foi então somado ao intervalo amostral, e as residências listadas ao lado dos números resultantes na lista foram escolhidas para entrevista.

Quando as estimativas iniciais do tamanho do quarteirão eram muito imprecisas e os procedimentos mencionados produziam apenas uma ou duas residências na amostra (ou mais do que dez), um procedimento diferente foi usado. Nestes casos, um número arbitrário de residências (não mais do que dez) era selecionado, e se fazia uma anotação para que aquelas entrevistadas tivessem peso separado durante a análise.

Isso completa nossa discussão sobre os procedimentos usados na seleção de cerca de 3.500 residências na cidade de Oakland, Califórnia, em 1966. O processo todo, incluindo a atualização do tamanho dos quarteirões de recenseamento e a listagem dos quarteirões de amostra, levou cerca de cinco meses, exigindo aproximadamente vinte pessoas no pico do trabalho.

Ponderação das Residências da Amostra

No desenho da amostra aleatória simples, cada elemento na população tem a mesma probabilidade de seleção. Como resultado, a amostra agregada pode ser tomada como representativa da população da qual foi extraída. Se 2.000 respondentes forem selecionados de uma população de 2 milhões, então considerá-se que cada entrevistado estaria "representando" mais 999 pessoas não selecionadas. Para estimar o número de pessoas na população com uma determinada característica, multiplicamos por 1.000 o número de pessoas com aquela característica na nossa amostra. Este peso é o inverso da probabilidade dos respondentes serem selecionados na amostra. Quando todos têm a mesma probabilidade de seleção, o peso é irrelevante, exceto para estimar números na população.

Quando os respondentes têm probabilidades diferentes de seleção, a ponderação é mais importante, sendo relevante até mesmo na computação das porcentagens. O estudo

Oakland é um exemplo da necessidade de e dos métodos para ponderar elementos de amostra num desenho complexo de amostra. Entretanto, independentemente da complexidade do projeto, o princípio básico se mantém: o peso do respondente é o inverso da sua probabilidade de seleção na amostra.

Ao computar a probabilidade geral de seleção de um respondente, devemos lembrar-nos de que se temos várias etapas de amostragem, com probabilidades diferentes de seleção em cada uma, estas probabilidades devem ser multiplicadas para determinar a probabilidade geral. Se os respondentes pertencem a um grupo (igreja, quarteirão) com a chance de seleção de um décimo, e eles têm uma chance de seleção de um décimo *dentro* daquele grupo, sua probabilidade geral de seleção é de $1/100$.

No cálculo dos pesos das residências na amostra de Oakland, devemos considerar duas probabilidades separadas de seleção: a probabilidade de um quarteirão ser selecionado e a de uma residência ser selecionada dentro daquele quarteirão. Estas probabilidades devem ser calculadas como se segue.

Probabilidade do Quarteirão. Cada quarteirão tem a probabilidade de seleção igual ao seu tamanho estimado (TEQ), dividido pelo tamanho do estrato (TE), vezes 2 (dois quarteirões selecionados por estrato). Escrevemos esta equação como $2TEQ/TE$. Note que esta fórmula considera tanto a amostragem PPT quanto a variação dos tamanhos dos estratos. No nosso exemplo anterior, o quarteirão 27/5 tinha um tamanho estimado de 30 residências num estrato contendo 247; portanto, sua probabilidade de seleção era $(30 \times 2)/247 = 0,2429$.

Probabilidade da Residência. Dentro de cada quarteirão, cada residência selecionada tinha uma probabilidade de seleção igual ao número selecionado no quarteirão (n) dividido pelo número de fato (NFQ) listado para todo o quarteirão, escrito como n/NFQ . Caso se observasse na listagem que o quarteirão 27/5 tinha 34 residências (em vez de 30), e fossem selecionadas cinco delas, cada uma teria a probabilidade de seleção igual a $5/24 = 0,1471$.

Probabilidade Geral. Multiplicando as fórmulas separadas para obter a probabilidade geral, temos o seguinte: $(2TEQ)(n)/(TE)(NFQ)$. No exemplo acima, a probabilidade geral de seleção é $(2 \times 30 \times 5)/(247 \times 34) = 0,0357$, ou cerca de $1/28$. Observe que esta probabilidade é menor do que a fração de amostragem visada de $1/20$ na Área de Pobreza D. A razão para esta discrepância é que o estrato 1 tinha 247 residências em vez da meta de 200, e o quarteirão 27/5 tinha 34 residências em vez das 30 estimadas. Portanto, cada uma das cinco residências

selecionadas no quarteirão 27/5 tinha menor probabilidade de seleção do que na intenção original.

No entanto, note que, se o tamanho estimado do quarteirão e o tamanho real fossem idênticos, a fórmula de probabilidade seria reduzida de $(2TEQ)(n)/(TE)(NQ)$ para $2n/(TE)$, ou duas vezes o número de residências selecionadas dividido pelo número de residências no estrato. Se o tamanho do estrato (TE) fosse 200, a fração de amostragem visada de 1/20 seria atingida, uma vez que cinco residências teriam sido escolhidas no quarteirão.

Ponderação das Residências. Todas as diferenças em probabilidades de seleção foram levadas em consideração na alocação de pesos igual ao inverso da probabilidade geral de seleção de uma residência. No primeiro exemplo acima, cada residência entrevistada no quarteirão 27/5 teria recebido um peso de 28. O suposto é que cada uma daquelas residências representaria a si própria e outras 27 na cidade de Oakland.

Ponderação Adicional por Não-Resposta. Um comentário final. *Surveys* deste tipo nunca conseguem entrevistar todas as residências inicialmente selecionadas na amostra. Algumas pessoas na amostra se recusam a ser entrevistadas e outras não estão disponíveis.

Neste estudo, foi atribuído peso adicional a residências para dar conta da não-resposta. Supôs-se que as residências em que não se conseguiu fazer entrevistas eram mais como as casas entrevistadas no mesmo quarteirão do que como qualquer outra estimativa possível. Portanto, cada entrevista completada recebia um peso igual ao número de residências selecionadas naquele quarteirão dividido pelo número de entrevistas de fato completadas. Se quatro entrevistas em cinco eram completadas, cada entrevista completa recebia um peso adicional de $5/4 = 1,25$. Naturalmente, se todas as residências selecionadas eram entrevistadas com sucesso, o peso adicional era de 1, podendo ser ignorado.

Resumo

Os quatro exemplos de desenhos de amostra foram apresentados de forma a dar um quadro mais realista das situações de amostragem com as quais você pode se deparar na prática. Embora estes exemplos não esgotem as possibilidades de variação em condições de campo, objetivos do estudo e técnicas de amostragem, eles ilustram as situações mais típicas. Espero

que eles tenham ilustrado a lógica básica da amostragem de *survey*, ajudando-o a improvisar sabiamente quando se deparar com um problema novo.

Notas

¹ Em sua maior parte, um quarteirão de recenseamento é o mesmo que um quarteirão de cidade — uma área mais ou menos retangular, limitada por quatro ruas, com casas alinhadas ao longo das ruas. Entretanto, em alguns lugares as ruas são irregulares, e blocos de formato estranho recebem o nome de quarteirões.

² Por exemplo, na época do estudo, uma região da cidade de Oakland era quase totalmente branca. Portanto, fazia pouco sentido estratificar os quarteirões naquela área em termos da composição racial, embora fosse possível agrupar os quarteirões com *quaisquer* residentes não-brancos. Por outro lado, se um pequeno grupo de quarteirões na região tivesse, digamos, 20% ou mais de residentes não-brancos, estes poderiam ter sido colocados em um estrato separado.

Leituras Adicionais

- ALMOND, Gabriel, VERBA, Sidney. *The Civic Culture*. Princeton, NJ: Princeton University Press, 1963.
- BABBIE, Earl R. *Science and Morality in Medicine*. Berkeley: University of California Press, 1970.
- GLOCK, Charles Y., RINGER, Benjamin B., BABBIE, Earl R. *To Comfort and to Challenge*. Berkeley: University of California Press, 1967.
- GLOCK, Charles Y., STARK, Rodney. *Christian Beliefs and Anti-Semitism*. New York: Harper & Row, 1966.
- STOUFFER, Samuel A. *Communism, Conformity, and Civil Liberties*. New York: John Wiley & Sons, 1966.