

1.3 Planejamento experimental

Planejamento de um estudo estatístico → Coleta de dados → Planejamento experimental → Técnicas de amostragem

Planejamento de um estudo estatístico

O objetivo de todo estudo estatístico é coletar dados e então usá-los para tomar uma decisão. Qualquer decisão que seja tomada usando os resultados de um estudo estatístico será tão boa quanto o processo utilizado para obtenção desses dados. Se o processo tiver falhas, então a decisão resultante será questionável.

Instruções

Planejando um estudo estatístico

1. Identifique a variável (as variáveis) de interesse (foco) e a população do estudo.
2. Desenvolva um plano detalhado para a coleta de dados. Se usar uma amostra, tenha certeza de que a amostra representa a população.
3. Colete os dados.
4. Descreva os dados usando técnicas de estatística descritiva.
5. Interprete os dados e tome as decisões sobre a população usando estatística inferencial.
6. Identifique quaisquer erros possíveis.

Coleta de dados

Há várias maneiras de se coletar dados. Frequentemente, o foco do estudo determina a melhor maneira de fazer a coleta. A seguir, temos um breve resumo de quatro métodos de coleta de dados.

- *Faça um estudo observacional* Em um *estudo observacional*, um pesquisador observa e mede as características de interesse de parte de uma população, mas não muda as condições existentes. Por exemplo, foi realizado um estudo observacional no qual os pesquisadores observaram e registraram o comportamento oral com objetos não alimentícios de crianças acima de 3 anos de idade. (*Fonte: Pediatrics Magazine.*)

- **Realize um experimento** Ao realizar um experimento, um tratamento é aplicado em uma parte da população e as respostas são observadas. Outra parte da população pode ser usada como grupo de controle, no qual nenhum tratamento é aplicado. Em muitos casos, indivíduos (às vezes chamados de **unidades experimentais**) do grupo de controle recebem **placebos**, um tratamento não medicamentoso e que não causa danos, feito para parecer o tratamento real. As respostas do grupo de tratamento e do grupo de controle podem ser comparadas e estudadas. Por exemplo, foi realizado um experimento no qual diabéticos tomaram extrato de canela diariamente enquanto o grupo de controle não tomou nada. Depois de 40 dias, os diabéticos que tomaram o extrato de canela reduziram seu risco de problemas cardíacos, enquanto o grupo de controle não experimentou mudanças. (Fonte: *Diabetes Care*.)

Use uma simulação Uma **simulação** é o uso de um modelo matemático ou físico para reproduzir as condições de uma situação ou processo. A coleta de dados frequentemente envolve o uso de computadores. As simulações permitem que você estude situações que são impraticáveis ou mesmo perigosas para serem criadas na vida real, e frequentemente economizam tempo e dinheiro. Por exemplo, os fabricantes de automóveis usam simulações com bonecos para estudar os efeitos das batidas em humanos. Durante a leitura deste livro, você terá a oportunidade de usar *applets* que simulam os processos estatísticos em computador.

Use um levantamento ou pesquisa de mercado Um levantamento ou pesquisa de mercado é uma investigação de uma ou mais características de uma população. Mais frequentemente, essas pesquisas são conduzidas com *pessoas*, por meio de perguntas feitas a elas. Os tipos mais comuns de levantamento são realizados por meio de entrevistas, correio ou telefone. Ao planejar esse tipo de pesquisa, é importante escolher bem as perguntas para não obter resultados tendenciosos. Por exemplo, uma pesquisa é conduzida em uma amostra de médicos do sexo feminino para determinar se o argumento principal para a escolha profissional é a estabilidade financeira. Ao planejar uma pesquisa, seria aceitável fazer uma lista de razões e perguntar a cada indivíduo na amostra para selecionar sua primeira escolha.

Exemplo 1

Decidindo o método de coleta de dados

Considere os estudos estatísticos a seguir. Qual método de seleção de dados você usaria para coletar os dados para cada estudo? Explique seu raciocínio.

1. Um estudo do efeito da mudança dos padrões de voo no número de acidentes com aviões.
2. Um estudo dos efeitos da ingestão de farinha de aveia na redução de pressão arterial.
3. Um estudo sobre como alunos da quarta série resolvem um quebra-cabeça.
4. Um estudo sobre os índices de aprovação presidencial com os residentes nos Estados Unidos.

Solução

1. Por ser impossível criar essa situação, **use simulação.**
2. Neste estudo, você quer medir o efeito que um tratamento (ingestão de aveia) tem nos pacientes. Então, você deve **realizar um experimento.**
3. Como você quer observar e medir certas características de parte de uma população, você poderia fazer um **estudo observacional.**
4. Você poderia usar uma **pesquisa para perguntar** "Você aprova a maneira pela qual o presidente está lidando com o cargo?".

Tente você
1 Considere os estudos estatísticos a seguir. Qual método de coleta de dados você usaria para cada estudo?

1. Um estudo sobre os efeitos dos exercícios no alívio da depressão.
2. Um estudo do sucesso de graduandos de uma grande universidade para encontrar um emprego durante o primeiro ano da graduação.
 - a. Identifique o foco do estudo.
 - b. Identifique a população do estudo.
 - c. Escolha um método apropriado para a coleta de dados.

Foco: efeito dos exercícios no alívio da depressão.

População: coleção de todas as pessoas com depressão.

Experimento.

Foco: índices de sucessos de graduandos de uma grande universidade ao procurar emprego depois de um ano de formados.

População: coleção de todos os graduandos de uma grande universidade.

Pesquisa.

Planejamento experimental

Para produzir resultados significativos e não tendenciosos, os experimentos devem ser cuidadosamente planejados e executados. É importante saber quais passos devem ser realizados para que os resultados sejam válidos. Três elementos-chave de um experimento bem planejado são *controle, aleatorização e replicação*.

Em razão do fato de que os resultados podem ser arruinados por uma variedade de fatores, a capacidade de *controlar* esses fatores de influência é importante. Um desses fatores é *uma variável confounding*.

Definição

Uma **variável confounding** ocorre quando um pesquisador não pode dizer a diferença entre os efeitos de diferentes fatores em uma variável.

Por exemplo, para atrair mais consumidores, o dono de uma cafeteria faz um experimento reformando a loja e usando cores vibrantes. Ao mesmo tempo, um shopping center da região realiza sua grande inauguração. Se os negócios aumentarem na cafeteria, não podemos determinar se isso ocorreu por causa das novas cores ou por causa do novo shopping perto da cafeteria. Os efeitos das cores e do shopping center se confundem.

Outro fator que pode afetar os resultados experimentais é o *efeito placebo*. O **efeito placebo** ocorre quando um sujeito reage favoravelmente a um placebo quando, de fato, ele(a) não recebeu tratamento medicamentoso nenhum. Para ajudar a controlar ou minimizar o efeito placebo, uma técnica chamada **cega** pode ser usada.

Definição

A técnica **cega** é uma técnica na qual o sujeito não sabe se está recebendo tratamento ou placebo. Em um **experimento duplamente cego** (*double-blind*), nem o sujeito nem o pesquisador sabem se o sujeito está recebendo tratamento ou placebo. O pesquisador é informado depois que todos os dados forem coletados. Este tipo de planejamento experimental é o preferido pelos pesquisadores.

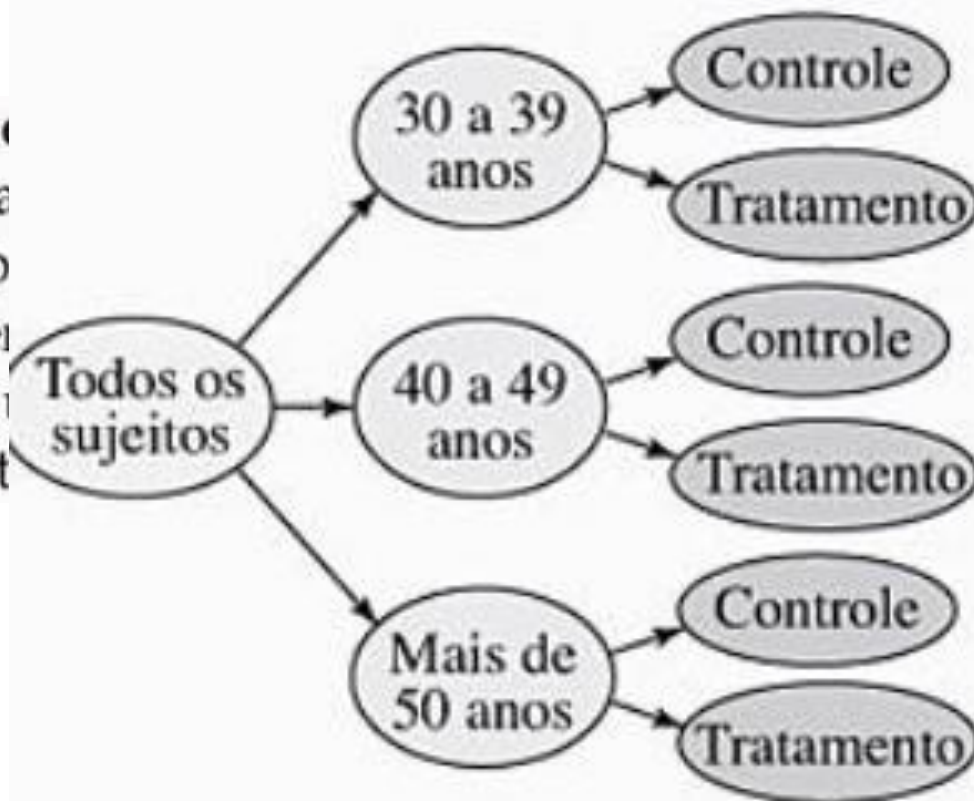
Outra técnica que pode ser usada para obter resultados imparciais é a *aleatorização*.

Definição

Aleatorização é o processo de distribuição aleatória dos sujeitos em diferentes grupos de tratamento.

Planejamento de blocos aleatórios

... diferentes grupos de



... designados para alguns experimentos com características semelhantes. Um planejamento de blocos aleatórios deve dividir os sujeitos aleatoriamente

Em um planejamento de blocos aleatórios, os sujeitos são divididos em diferentes grupos de tratamentos, onde cada grupo recebe um tratamento diferente. Um planejamento de blocos aleatórios é necessário quando os sujeitos têm características semelhantes. Um planejamento de blocos aleatórios deve dividir os sujeitos com características semelhantes em diferentes grupos de tratamentos.

Definição

Replicação é a repetição de um experimento usando um grande grupo de sujeitos.

Exemplo 2

Analisando um planejamento experimental

Uma empresa quer testar a eficácia de uma nova goma de mascar para ajudar as pessoas a pararem de fumar. Identifique um problema em potencial com o planejamento experimental dado e sugira uma maneira de melhorá-lo.

1. A empresa identifica dez adultos que são fumantes há bastante tempo. Cinco deles recebem a nova goma de mascar e os outros cinco recebem um placebo. Depois de dois meses, eles são avaliados e descobre-se que os cinco sujeitos que estão usando a nova goma pararam de fumar.

Solução

1. O tamanho da amostra usado não é grande o suficiente para validar os resultados. O experimento deve ser replicado para melhorar a validade.

2. A empresa identifica mil adultos que são fumantes há bastante tempo. Eles são divididos em blocos de acordo com o gênero. As mulheres recebem a nova goma e os homens recebem o placebo. Depois de dois meses, o grupo de mulheres tinha um número significativo de sujeitos que pararam de fumar.

Os grupos não são similares. A nova goma de mascar pode ter mais efeito nas mulheres do que nos homens ou vice-versa. Os sujeitos podem ser divididos em blocos de acordo com gênero, mas depois, dentro de cada bloco, eles precisam ser aleatoriamente designados para estar no grupo de tratamento ou de controle.

**Tente
você
2**

Usando as informações do Exemplo 2, suponha que a empresa identifique 240 adultos fumantes. Eles são designados aleatoriamente para estar no grupo de tratamento ou de controle. Cada sujeito também recebe um DVD sobre os perigos do cigarro. Depois de quatro meses, a maioria dos sujeitos no grupo de tratamento parou de fumar.

- a. Identifique um *problema em potencial* com o planejamento experimental.
- b. Como o planejamento poderia ser *melhorado*?

- a. Não há maneira de dizer por que as pessoas pararam de fumar. Elas poderiam ter deixado de fumar tanto pelo fato de terem mascado chiclete ou de terem assistido ao DVD.
- b. Dois experimentos poderiam ser feitos; um usando o chiclete e outro usando o DVD.

Técnicas de amostragem

Um **censo** é uma contagem ou medição de uma população *inteira*. A realização de um censo fornece informações completas, mas ela é frequentemente cara e difícil de realizar. Uma **amostragem** é uma contagem ou medição de *parte* de uma população e é mais comumente usada nos estudos estatísticos. Para coletar dados imparciais, o pesquisador deve ter certeza de que a amostra representa a população. Técnicas de amostragem apropriadas devem ser utilizadas para assegurar que as inferências sobre a população são válidas. Lembre-se de que quando um estudo é realizado com dados falhos, os resultados são questionáveis. Mesmo com os melhores métodos de amostragem, um erro de amostragem pode acontecer. Um erro de amostragem é a diferença entre os resultados da amostra e da população. Quando aprendemos sobre estatística inferencial, também aprendemos técnicas para controlar esses erros de amostragem.

Uma amostra aleatória é aquela na qual todos os membros de uma população têm chances iguais de serem selecionados. Uma amostra aleatória simples é aquela na qual toda amostra possível de mesmo tamanho tem a mesma chance de ser selecionada. Uma maneira de coletar uma amostra aleatória simples é designar um número diferente para cada membro da população e então usar uma tabela numérica aleatória, como a do Apêndice B. As respostas, contagens ou medições provenientes de membros da população cujos números correspondem àqueles gerados com o uso da tabela farão parte da amostra. Calculadoras e programas de computador também são utilizados para gerar números aleatórios (ver p. 29).

Tabela 1 Números aleatórios

92630	78240	19267	95457	53497	23894	37708	79862
79445	78735	71549	44843	26104	67318	00701	34986
59654	71966	27386	50004	05358	94031	29281	18544
31524	49587	76612	39789	13537	48086	59483	60680
06348	76938	90379	51392	55887	71015	09209	79157

(Parte da Tabela 1 encontrada no Apêndice B.)

Por exemplo, para usar uma amostra aleatória simples na contagem do número de pessoas que moram nas residências do Condado de West Ridge, você poderia designar um número diferente para cada residência, usar uma ferramenta tecnológica ou uma tabela de números aleatórios para gerar uma amostra de números e então contar o número de pessoas que vivem em cada uma das residências selecionadas.

Exemplo 3

Usando uma amostra aleatória simples

Há 731 estudantes que se inscreveram no curso de estatística em sua faculdade. Você deseja formar uma amostra de 8 estudantes para responder às questões de uma pesquisa. Selecione os estudantes que pertencerão à amostra aleatória simples.

Solução

Designue números de 1 a 731 para cada estudante do curso. Na tabela de números aleatórios, escolha um ponto de partida aleatoriamente e leia os dígitos em grupos de 3 (porque 731 é um número de 3 dígitos). Por exemplo, se você começar na terceira fileira da tabela, no começo da segunda coluna, você agruparia os números como a seguir:

719|66 2|738|6 50|004| 053|58 9|403|1 29|281| 185|44

Ignorando os números maiores do que 731, então os primeiros oito números são 719, 662, 650, 4, 53, 589, 403 e 129. Os estudantes que receberam esses números formarão a amostra. Para encontrar a amostra usando a TI-83/84, siga as instruções ao lado.

Existem muitas outras técnicas de amostragem comumente usadas. Cada uma tem vantagens e desvantagens.

- **Amostra estratificada** Quando é importante que uma amostra tenha membros de cada segmento da população, devemos usar uma amostra estratificada. Dependendo do foco do estudo, membros de uma população são divididos em dois ou mais grupos, chamados de estratos, que compartilham uma característica similar como idade, sexo, grupo étnico ou até mesmo preferência política. Uma amostra é então selecionada aleatoriamente de cada um dos estratos. O uso de uma amostra estratificada assegura que cada segmento da população será representado. Por exemplo, para coletar uma amostra estratificada do número de pessoas que moram em Condado de West Ridge, você poderia dividir as residências em níveis socioeconômicos e, então, selecionar aleatoriamente residências de cada nível.



Grupo 1:
renda baixa



Grupo 2:
renda média



Grupo 3:
renda alta

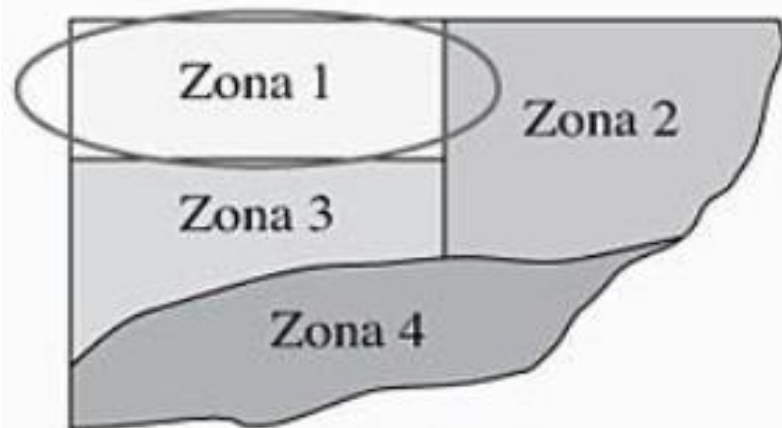
Amostra estratificada

Amostra por agrupamento Quando a população ocorre naturalmente, cada um tendo características por agrupamento pode ser a mais apropriada por agrupamento, divida a população em grupos e inclua todos os membros em um ou mais (múltiplos de *clusters* poderiam ser seções diferentes filiais de um banco. Por exemplo, para um estudo do número de pessoas que moram em West Ridge, divida as residências em grupos postais, então, selecione todas as residências

Importante

Para uma amostra estratificada, cada um dos estratos contém membros com certas características (por exemplo, um grupo de idade em particular). Em contraste, os *clusters* consistem de um agrupamento geográfico, e cada *cluster* deve consistir de membros com todas as características (por exemplo, todas as faixas etárias). Com amostras estratificadas, alguns dos membros de cada grupo são usados. Na amostra por agrupamento, todos os membros de um ou mais grupos são usados.

Zonas de códigos postais na região de West Ridge



Amostra por agrupamento

Amostra sistemática Uma amostra sistemática é aquela na qual é atribuído um número a cada membro da população. Os membros da população são ordenados de alguma maneira, um número inicial é selecionado aleatoriamente e, então, membros da amostra são selecionados em intervalos regulares a partir do número inicial. (Por exemplo, cada 3º, 5º ou 100º membro é selecionado.) Por exemplo, para coletar uma amostra sistemática do número de pessoas que moram em West Ridge, poderíamos designar um número diferente para cada residência, escolher aleatoriamente um número inicial, selecionar cada 100ª residência e contar o número de pessoas vivendo em cada uma. Uma vantagem da amostra sistemática é que ela é fácil de ser usada. No caso de qualquer padrão que aconteça regularmente nos dados, entretanto, esse tipo de amostragem deve ser evitado.



Amostra sistemática

Um tipo de amostra que frequentemente leva a estudos tendenciosos (portanto, não é recomendada) é a **amostra de conveniência**. Uma amostra de conveniência consiste somente de membros disponíveis de uma população.

Exemplo 4

Identificando as técnicas de amostragem

Você está realizando um estudo para determinar a opinião dos estudantes em sua escola sobre a pesquisa de células-tronco. Identifique a técnica de amostragem que você usaria se selecionasse as amostras listadas.

1. Você seleciona uma classe aleatoriamente e questiona cada aluno da classe.

Solução

1. Pelo fato de cada classe ser um subgrupo que ocorre naturalmente (um *cluster*) e você questiona cada aluno na classe, esta é uma amostra por agrupamento.

2. Você divide a população de estudantes com relação às graduações, seleciona aleatoriamente e questionar alguns de cada curso de graduação.

Como os estudantes são divididos em estratos (graduações) e uma amostra é selecionada de cada graduação, esta é uma amostra estratificada.

3. Você designa um número para cada aluno e gera números aleatoriamente. Então, você questiona cada estudante cujo número é selecionado aleatoriamente.

Cada amostra de mesmo tamanho tem chances iguais de ser selecionada e cada aluno tem chances iguais de ser selecionado, então esta é uma amostra aleatória simples.

**Tente
você**
4

Você quer determinar a opinião dos estudantes de sua escola sobre as células-tronco. Identifique a técnica de amostragem que está sendo usada se você selecionar as amostras listadas.

1. Você seleciona estudantes que estão em sua aula de estatística.
2. Você designa um número para cada estudante e, depois de escolher um número inicial, questiona cada 25º aluno.
 - a. Determine *como* a amostra é selecionada.
 - b. Identifique a *técnica de amostragem* correspondente.

(1) Uma amostra foi selecionada usando-se somente os estudantes disponíveis.

Amostragem de conveniência.

(2) A amostra foi selecionada numerando cada estudante na escola, escolhendo aleatoriamente um número inicial e selecionando estudantes em intervalos regulares a partir do número inicial.

Amostragem sistemática.