

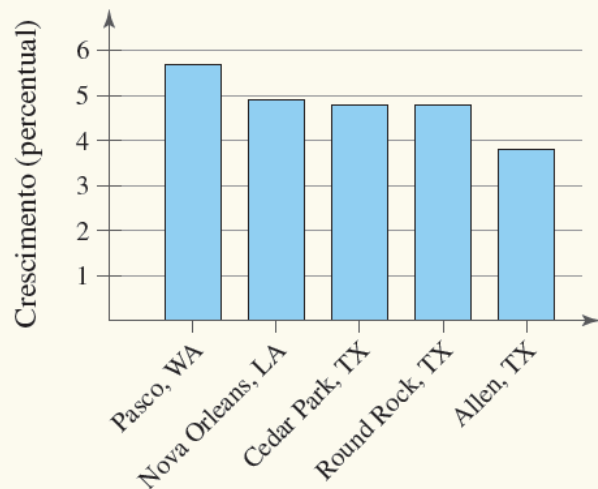
TTI109 – Estatística

Aula 01 – Introdução à Estatística

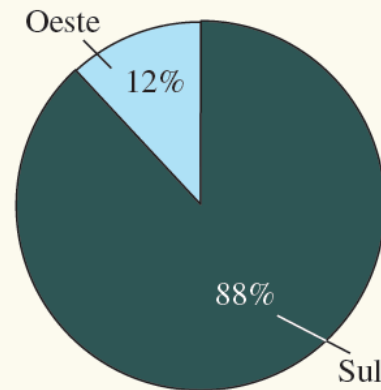


Uma visão geral da estatística

(a) Cidades americanas com maior crescimento populacional relativo.



(b) Localização das 25 cidades americanas com maior crescimento populacional relativo.



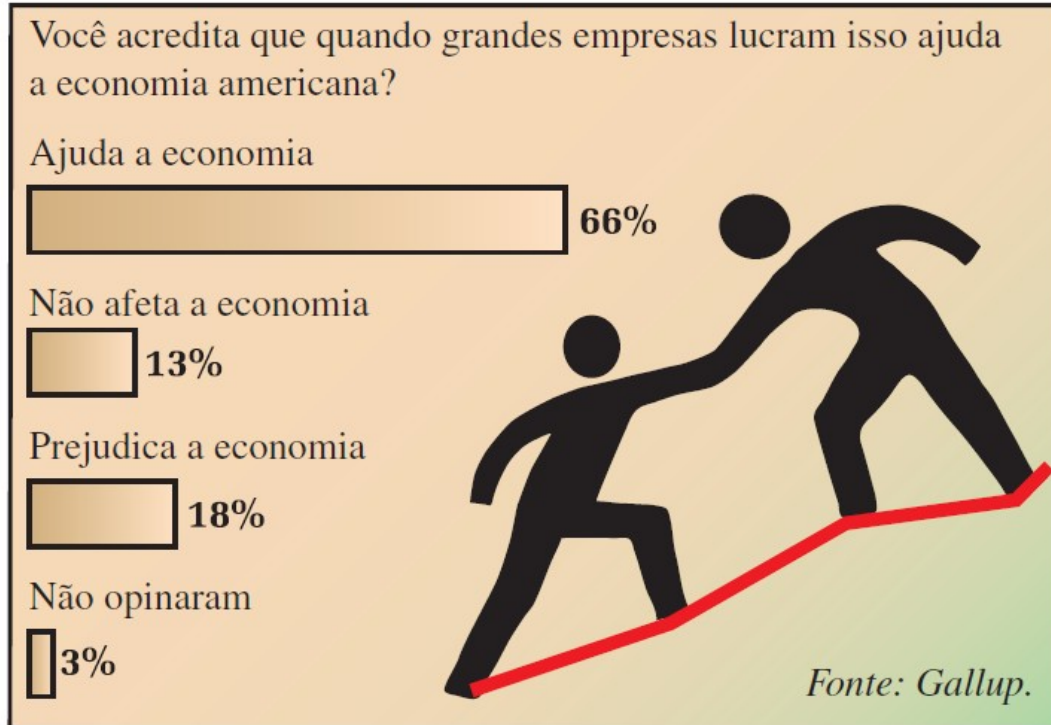
Coleta de dados

Uma visão geral da estatística

- “[Mulheres] que fumaram de um a 14 cigarros por dia tinham, aproximadamente, o dobro de risco de morte cardíaca repentina do que as não fumantes.” (*Fonte: American Heart Association.*)
- “O desperdício de alimentos [nos Estados Unidos], em relação ao suprimento disponível, tem aumentado progressivamente de aproximadamente 30% em 1974 para quase 40% em anos recentes.” (*Fonte: National Institute of Diabetes and Digestive and Kidney Diseases.*)
- “O percentual de estudantes de Detroit que alcançaram nível igual ou acima de *proficiência* [em leitura] foi de 7% [em um ano recente].” (*Fonte: U.S. Department of Education.*)

Coleta de dados

Uma visão geral da estatística



Coleta de dados

Uma visão geral da estatística

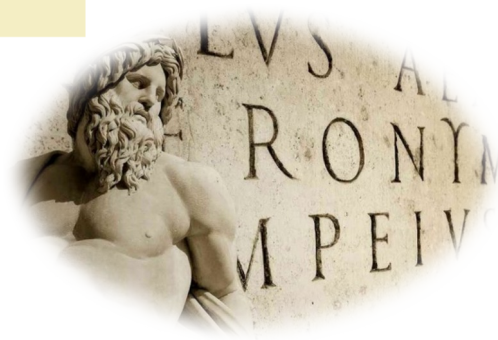
Definição

Dados consistem em informações provenientes de observações, contagens, medições ou respostas.

Definição

Estatística é a ciência que trata da coleta, organização, análise e interpretação dos dados para a tomada de decisões.

A palavra *estatística* é derivada da palavra latina *status*, que significa “estado”.



Uma visão geral da estatística

- Conjunto de dados:

Definição

Uma **população** é a coleção de *todos* os resultados, respostas, medições ou contagens que são de interesse.

Uma **amostra** é um subconjunto ou parte de uma população.



Uma amostra deve ser *representativa de uma população* de modo que seus dados possam ser usados para tirar conclusões sobre aquela população

Uma visão geral da estatística

Identificando conjuntos de dados

Em uma pesquisa recente, foi perguntado a 614 proprietários de pequenas empresas nos Estados Unidos se eles achavam que a presença de sua empresa no Facebook tinha valor. Duzentos e cinquenta e oito dos 614 responderam que sim. Identifique a população e a amostra. Descreva o conjunto de dados da amostra. (*Adaptado de: Manta.*)

População? Amostra?

Respostas de todos os proprietários de pequenas empresas nos Estados Unidos (população)

Respostas dos 614 proprietários de pequenas empresas na pesquisa (amostra)

Uma visão geral da estatística

- Conjunto de dados:

Definição

Um **parâmetro** é a descrição numérica de uma característica *populacional*.

Uma **estatística** é a descrição numérica de uma característica *amostral*.

↳ É importante notar que uma estatística amostral pode diferir de uma amostra para outra, enquanto um parâmetro populacional é constante para uma população.

Uma visão geral da estatística

Distinguindo entre um parâmetro e uma estatística

Determine se o valor numérico descreve um parâmetro populacional ou uma estatística amostral. Explique seu raciocínio.

Estatística

1. Uma pesquisa recente com aproximadamente 400.000 empregadores reportou que o salário médio inicial para um especialista em marketing é de **US\$ 53.400** por ano. (*Fonte: National Association of Colleges and Employers.*)
2. A nota média de matemática obtida no vestibular pelos calouros de uma universidade é **514**.
3. Em uma checagem aleatória de 400 lojas varejistas, o FDA (*Food and Drug Administration*) descobriu que **34%** das lojas não estavam estocando peixes na temperatura apropriada.

Parâmetro

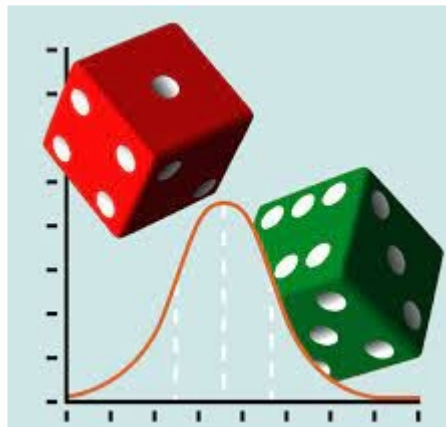
Uma visão geral da estatística

- Ramos da estatística:

Definição

Estatística descritiva é o ramo da estatística que envolve a organização, o resumo e a representação dos dados.

Estatística inferencial é o ramo da estatística que envolve o uso de uma amostra para chegar a conclusões sobre uma população. Uma disciplina no estudo da estatística inferencial é a probabilidade.



Uma visão geral da estatística

Estatística descritiva e inferencial

Determine qual parte do estudo representa o ramo descritivo da estatística. Que conclusões podem ser tomadas do estudo usando estatística inferencial?

1. Uma grande amostra de homens com 48 anos de idade foi estudada durante 18 anos.

Estatística descritiva

Inferência

Estar casado está
associado a uma vida mais
longa para os homens



Ainda vivos aos 65 anos

Solteiros



Casados



Uma visão geral da estatística

2. Em uma amostra de analistas de Wall Street, a percentagem dos que previram incorretamente os lucros de empresas de alta tecnologia em um ano recente foi de 44%. (Fonte: *Bloomberg News*.)

Estatística descritiva

Inferência

O mercado de ações é difícil de ser previsto, até mesmo para os profissionais



Classificação dos dados

- Tipos de dados:

Definição

Dados qualitativos consistem em atributos, rótulos ou entradas não numéricas.

Dados quantitativos consistem em medidas numéricas ou contagens.

Preços de venda sugeridos para veículos Honda.

Modelo	Preço de venda sugerido
Accord Sedan	US\$ 21.680
Civic Hybrid	US\$ 24.200
Civic Sedan	US\$ 18.165
Crosstour	US\$ 27.230
CR-V	US\$ 22.795
Fit	US\$ 15.425
Odyssey	US\$ 28.675
Pilot	US\$ 29.520
Ridgeline	US\$ 29.450

Dados qualitativos

Dados quantitativos

Classificação dos dados

● Níveis de mensuração:

**Dados qualitativos,
nível ordinal**

Cinco maiores bilheterias de 2012.

1. Os Vingadores
2. Batman: O cavaleiro das trevas ressurgue
3. Jogos vorazes
4. 007 – Operação Skyfall
5. A saga crepúsculo: Amanhecer – Parte 2

Definição

Dados no **nível nominal de mensuração** são apenas qualitativos. Dados nesse nível são categorizados usando-se nomes, rótulos ou qualidades. Não é possível realizar cálculos matemáticos nesse nível.

Dados no **nível ordinal de mensuração** são qualitativos ou quantitativos. Dados nesse nível podem ser postos em ordem ou classificados, mas as diferenças entre as entradas de dados não têm sentido matemático.

Gêneros de filme.

Ação
Aventura
Comédia
Drama
Terror

**Dados qualitativos,
nível nominal**

Classificação dos dados

- Níveis de mensuração:

Dados quantitativos

Definição

Dados no **nível de mensuração intervalar** podem ser ordenados e é possível calcular diferenças que tenham sentido matemático entre as entradas de dados. No nível intervalar, um registro zero simplesmente representa uma posição em uma escala; a entrada não é um zero natural. Dados no **nível de mensuração de razão** são similares aos dados no nível intervalar, com a propriedade adicional de que, nesse nível, um registro zero é um zero natural. Uma razão de dois valores pode ser formada de modo que um dado possa ser expresso significativamente como um múltiplo de outro.

➡ Um *zero natural* é um zero que significa “*nenhum*”: R\$ 0,00 indica nenhum dinheiro, mas 0°C não indica ausência de calor e sim uma posição na escala Celsius.

Classificação dos dados

Classificando dados por nível

Dois conjuntos de dados são apresentados nas tabelas 1.5 e 1.6. Qual conjunto consiste em dados no nível intervalar? Qual conjunto consiste em dados no nível de razão? Explique seu raciocínio. (*Fonte: Major League Baseball.*)

Tabela 1.5 Anos das vitórias do New York Yankees na Série Mundial.

1923, 1927, 1928, 1932, 1936, 1937,
1938, 1939, 1941, 1943, 1947, 1949,
1950, 1951, 1952, 1953, 1956, 1958,
1961, 1962, 1977, 1978, 1996, 1998,
1999, 2000, 2009

**Dados quantitativos,
nível intervalar**

**Dados quantitativos,
nível de razão**

Tabela 1.6 Total de *home runs* da Liga Americana em 2012 (por time).

Baltimore	214	Minnesota	131
Boston	165	Nova York	245
Chicago	211	Oakland	195
Cleveland	136	Seattle	149
Detroit	163	Tampa Bay	175
Kansas City	131	Texas	200
Los Angeles	187	Toronto	198

Classificação dos dados

- Quadro resumo:

Nível de mensuração	Categorizar os dados	Ordenar os dados	Subtrair os dados	Determinar se um dado é múltiplo do outro
Nominal	Sim	Não	Não	Não
Ordinal	Sim	Sim	Não	Não
Intervalar	Sim	Sim	Sim	Não
Razão	Sim	Sim	Sim	Sim

Planejamento de um estudo estatístico

Instruções

Planejando um estudo estatístico

1. Identifique a(s) variável(is) de interesse (o foco) e a população do estudo.
2. Desenvolva um plano detalhado para a coleta de dados. Se usar uma amostra, certifique-se de que a amostra é representativa da 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.

Estudo observacional: o pesquisador observa e mede as características de interesse de parte de uma população, mas não muda as condições existentes.

Experimento: um tratamento é aplicado em uma parte da população, chamada de **grupo de tratamento**. Outro grupo, o **grupo de controle**, não recebe tratamento. As respostas são comparadas e estudadas.

Planejamento de um estudo estatístico

Distinguindo entre um estudo observacional e um experimento

Determine se o estudo é observacional ou um experimento.

1. Pesquisadores estudam o efeito da complementação de vitamina D₃ em pacientes com deficiência de anticorpos ou com infecções frequentes do trato respiratório. Para realizar o estudo, 70 pacientes recebem 4.000 UI de vitamina D₃ diariamente por um ano. Outro grupo de 70 pacientes recebe um placebo diariamente por um ano. (Fonte: *British Medical Journal*.)
2. Pesquisadores conduzem um estudo para determinar o índice de aprovação pública nacional do presidente dos Estados Unidos. Para realizar o estudo, os pesquisadores ligaram para 1.500 residentes no país e perguntaram se eles aprovavam ou não o trabalho realizado pelo presidente. (Fonte: *Gallup*.)



Experimento



Estudo observacional

Coleta de Dados

Simulação: uso de um modelo matemático e físico para reproduzir as condições de uma situação ou processo.



Pesquisa: investigação de uma ou mais características de uma população – mais frequentemente, as pesquisas são conduzidas com pessoas, por meio de entrevistas.

Planejamento experimental

Elementos-chave: Controle, Aleatorização e Replicação

Falha no experimento:

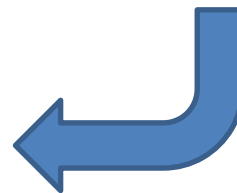
Definição

Uma **variável de confusão** ocorre quando um pesquisador não pode distinguir um ou mais fatores que causaram os efeitos provocados sobre a variável em estudo, gerando confusão.

Definição

O experimento cego (ou **cegamento**) é uma técnica na qual o indivíduo não sabe se está recebendo um tratamento ou um placebo. Em um **experimento duplo-cego**, nem o pesquisador nem os indivíduos sabem quem está recebendo um tratamento ou um placebo. O pesquisador é informado depois que todos os dados forem coletados. Esse tipo de planejamento experimental é preferido pelos pesquisadores.

Efeito placebo: ocorre quando um indivíduo reage favoravelmente a um tratamento quando, na verdade, recebeu um placebo.



*Para evitar o
efeito
placebo...*

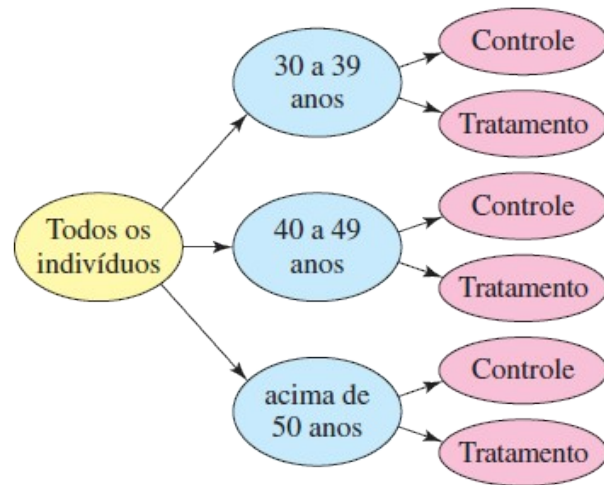
Planejamento experimental

Definição

Aleatorização é o processo de se designar indivíduos aleatoriamente para diferentes grupos de tratamento.

¬^ **Planejamento completamente aleatorizado:** os indivíduos são designados para diferentes grupos de tratamento por meio de seleção aleatória.

¬^ **Planejamento em blocos aleatorizados:** o pesquisador separa os indivíduos com características similares em blocos e, então, dentro de cada bloco, designa-os aleatoriamente para os grupos.



Planejamento experimental

Definição

Replicação é a repetição de um experimento sob condições iguais ou semelhantes.

✗ O tamanho da amostra não é grande o suficiente para validar os resultados. O experimento deve ser replicado, com amostra suficientemente maior, para melhorar a validade.

Analisando um planejamento experimental

Uma empresa quer testar a eficácia de uma nova goma de mascar desenvolvida 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 indivíduos que estão usando a nova goma pararam de fumar.
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, um número significativo de mulheres tinha parado de fumar.

Planejamento experimental

Definição

Replicação é a repetição de um experimento sob condições iguais ou semelhantes.

↯ Os grupos não são similares. A nova goma de mascar pode ter mais efeito nas mulheres do que nos homens ou vice-versa. Os indivíduos podem ser divididos em blocos de acordo com gênero, mas depois, dentro de cada bloco, eles precisam ser aleatoriamente designados para um dos grupos, tratamento ou controle.

Analisando um planejamento experimental

Uma empresa quer testar a eficácia de uma nova goma de mascar desenvolvida 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 indivíduos que estão usando a nova goma pararam de fumar.
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, um número significativo de mulheres tinha parado de fumar.

Técnicas de amostragem

Um *censo* é uma contagem ou medição de *toda a população*. A realização de um censo fornece informações completas, mas é frequentemente caro 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 não viesados, um pesquisador deve assegurar que a amostra é de fato *representativa da população*.



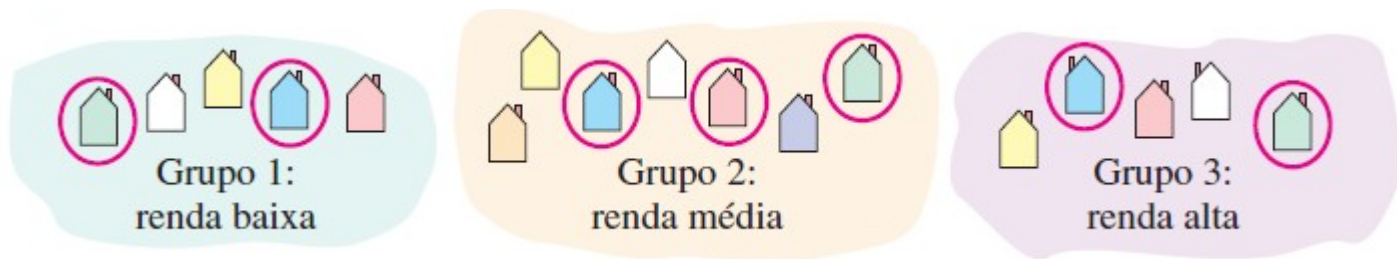
Amostragem aleatória: todos os elementos de uma população têm chances iguais de serem selecionados.

Amostragem aleatória simples: cada amostra possível de mesmo tamanho tem a mesma chance de ser selecionada.

Técnicas de amostragem

Amostragem estratificada: Elementos de uma população são divididos em dois ou mais subconjuntos, 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 está representado.



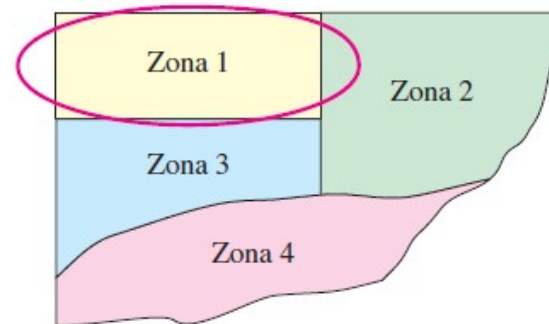
Técnicas de amostragem

Amostragem por conglomerado: Utilizada quando a população recai em *subgrupos que ocorrem naturalmente*, cada um tendo características similares.

- Dividir a população em grupos, chamados conglomerados, e selecionar todos os elementos em um ou mais (mas não em todos) conglomerados escolhidos.

- Ao usar esta forma de amostragem, é preciso assegurar que todos os conglomerados tenham características similares.

Amostragem por conglomerado — Zonas de códigos postais no Condado de West Ridge.



Técnicas de amostragem

Amostragem por conveniência (EVITAR!!!): consiste somente em membros da população que são fáceis de contatar...

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 ao selecionar as amostras listadas.

1. Você divide a população de estudantes com relação às graduações e, aleatoriamente, seleciona e questiona alguns estudantes em cada curso de graduação.
2. Você designa um número para cada estudante e gera números aleatórios. Então, você questiona cada estudante cujo número é selecionado aleatoriamente.
3. Você seleciona estudantes que são da sua turma de biologia.

✗ Amostragem estratificada

✓ Amostragem aleatória simples

└ Amostragem por conveniência (tendenciosa...)

TTI109 – Estatística

Aula 01 – Introdução à Estatística

