

Introdução à Estatística e conceitos de amostragem

Fernando de Pol Mayer

Laboratório de Estatística e Geoinformação (LEG)
Departamento de Estatística (DEST)
Universidade Federal do Paraná (UFPR)



Este conteúdo está disponível por meio da Licença Creative Commons 4.0
(Atribuição/NãoComercial/Partilha Igual)

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?
Por que?

Computadores
Áreas

Amostragem

Tipos

Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?
Por que?
Computadores
Áreas

Amostragem

Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?

Por que?

Computadores

Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

- Etimologia da palavra: do latim *status* \Rightarrow estado
- Origem: coleta e apresentação de dados de interesse do Estado
 - Informações sobre populações e riquezas
 - Fins militares e tributários
- Conjunto de métodos especialmente apropriado ao tratamento de dados numéricos, afetados por uma multiplicidade de causas
- Estes métodos **fazem uso** da Matemática, e especialmente do cálculo de **probabilidades**

- Confúcio relatou levantamentos feitos na China há mais de 2000 anos AC
- No Egito antigo, os faraós fizeram uso sistemático de informações de caráter estatístico
- O mesmo aconteceu com antigas civilizações como Maias, Astecas e Incas
- Imperadores faziam levantamentos de suas propriedades conquistadas (imperadores romanos, Carlos Magno, Guilherme, o Conquistador) para se inteirar de suas riquezas
- Essa prática tem sido continuada nos tempos modernos, por meio de recenseamentos, como aqueles feitos pelo IBGE no Brasil

Como Ciência

- Permite organizar, descrever, analisar, e interpretar dados
- Utiliza-se da **Teoria da Probabilidade** para modelar a aleatoriedade e a incerteza associada aos fenômenos naturais, econômicos, sociais, . . .
- Auxilia a tirar **conclusões** sobre as características das fontes de onde os dados foram retirados, para melhor compreendê-los
- Indispensável para a **tomada de decisões** sob condições de *incerteza*, sob o menor **risco** possível

Como tecnologia

- Permite avaliar as incertezas e os seus efeitos no planejamento e interpretação de experiências e de observações de fenômenos da natureza e da sociedade
- Permite analisar e tirar conclusões de uma grande quantidade de informações
- A estatística tem sido utilizada para
 - Otimização de recursos econômicos
 - Aumento da qualidade e produtividade
 - Análise de decisões judiciais
 - Previsões (climáticas, econômicas, ...)

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?

Por que?

Computadores

Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?

Por que?

Computadores

Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

- Impossibilidade de estudar a população
- Aumento da capacidade de registro de dados que precisam ser compreendidos
- Expansão do conhecimento científico, das áreas de pesquisa e dos instrumentos de investigação
- Necessidade de compreensão dos fenômenos naturais e sociais, de otimização de recursos, planejamento de atividades, redução de riscos, de previsão de resultados para correta tomada de decisão

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

A Estatística pode ser pensada como a **ciência de aprendizagem a partir dos dados**

Vivemos na “*era da informação*”, e a Estatística possui as ferramentas necessárias para melhor compreender a **informação**

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?
Por que?

Computadores

Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?

Computadores

Áreas

Amostragem

Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

No passado, tratar um grande conjunto de dados era uma tarefa trabalhosa e cansativa

Com o avanço da tecnologia, os cálculos se tornaram rápidos e mecânicos, possibilitando a análise de um volume grande de informações em pouco tempo

No entanto, é necessário *conhecer* e *compreender* os conceitos básicos de Estatística para que possamos utiliza-la de forma adequada

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

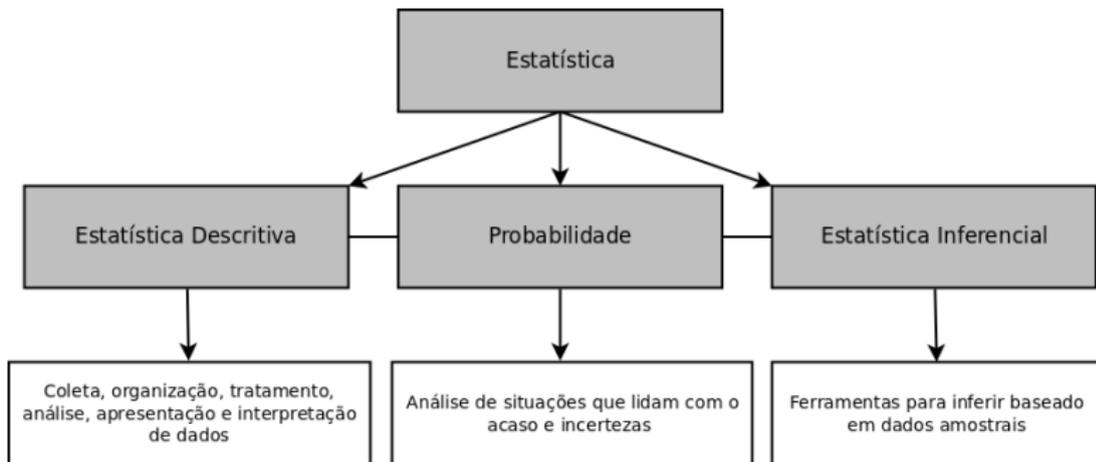
Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências



Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

Estatística Descritiva: etapa inicial de qualquer análise. É um conjunto de técnicas destinadas a descrever e resumir os dados, que auxiliam a descrever características de interesse.

⇒ “*Conheça seus dados*”

Probabilidade: é a ferramenta matemática utilizada pela Estatística para se estudar a **incerteza** oriunda de fenômenos **aleatórios**.

⇒ “*Qual a incerteza associada aos dados?*”

Estatística Inferencial: é um conjunto de técnicas que possibilita tirar conclusões sobre uma **população**, a partir de um subconjunto de valores (**amostra**).

⇒ “*Quais conclusões podemos tirar a partir destes dados?*”

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem

Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?

Por que?

Computadores

Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

Quando fazemos uma pesquisa, ou utilizamos algum mecanismo para obter informações, um dos objetivos principais é **coletar dados de uma pequena parte** de um *grande grupo* e aprender então alguma coisa sobre esse grupo maior

Introdução à
Estatística e
conceitos de
amostragem

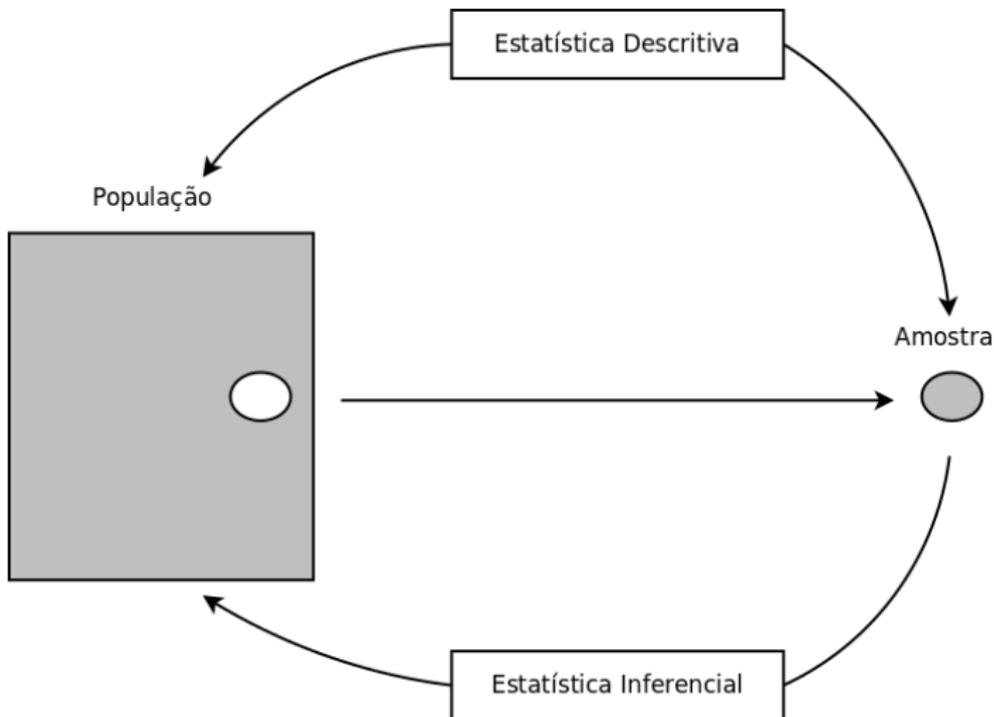
Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem

Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências



Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

“Astros do rock morrem jovens.”



Todos os astros do rock morrem jovens?

População: conjunto de indivíduos, objetos ou produtos que contém a característica que temos interesse. Exemplo:

- Característica: altura dos estudantes da UFPR
- População: **todos** os estudantes da UFPR

Observação

A população depende do interesse da pesquisa

Amostra: subconjunto da população, em geral com dimensão bem menor, que também possui a característica de interesse. Exemplo:

- Característica: altura dos estudantes da UFPR
- Amostra: 100 estudantes selecionados **ao acaso**

População → **censo** → **parâmetro**

Uma medida numérica que descreve alguma característica da população, usualmente representada por letras gregas: $\theta, \mu, \sigma, \dots$

Exemplo: média populacional = μ

População → **amostra** → **estatística**

Uma medida numérica que descreve alguma característica da amostra, usualmente denotada pela letra grega do respectivo parâmetro com um acento circunflexo: $\hat{\theta}, \hat{\mu}, \hat{\sigma}, \dots$, ou por letras do alfabeto comum: \bar{x}, s, \dots

Exemplo: média amostral = \bar{x}

- População: todos os alunos de uma única turma
- Característica: idade dos alunos

Censo: 22 21 24 23 20 22 21 25 24 24 23 19 25 24 23 23
20 21 23 20 23 22 23 23 25 25 20 23 24 20

Média populacional: $\mu = 22,5$ \leftarrow **Parâmetro**

Amostra de 5 alunos: 25 24 23 23 25

Média amostral: $\bar{x} = 24$ \leftarrow **Estatística**

Por que fazer amostragem?

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem

Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

- Parâmetros populacionais desconhecidos
- Impossibilidade de realização de um censo
- Mais barato, mais rápido

Por que fazer amostragem?

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem

Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

- Parâmetros populacionais desconhecidos
- Impossibilidade de realização de um censo
- Mais barato, mais rápido

Atenção!

Não existe nenhuma técnica estatística capaz de salvar uma amostra mal coletada!

Por que fazer amostragem?

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem

Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

- Parâmetros populacionais desconhecidos
- Impossibilidade de realização de um censo
- Mais barato, mais rápido

Atenção!

Não existe nenhuma técnica estatística capaz de salvar uma amostra mal coletada!

Em geral, uma amostra deve ser

- um subconjunto **representativo** da população
- **aleatória** (de alguma forma)

Existem diversas maneiras para se *retirar* uma **amostra** de uma *população* → **Teoria da Amostragem**

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?
Por que?

Computadores
Áreas

Amostragem

Tipos

Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

(A) Levantamentos amostrais

A amostra é obtida a partir de uma população bem definida, bem meio de processos bem definidos pelo pesquisador. Subdivide-se em dois grupos:

Probabilísticos Cada elemento da população possui a mesma probabilidade se ser selecionado para compor a amostra → mecanismos aleatórios de seleção

Não probabilísticos A seleção da amostra depende do julgamento do pesquisador. Há uma **escolha** deliberada dos elementos para compor a amostra → mecanismos não aleatórios de seleção

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas
Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais
Referências

(B) Planejamento de Experimentos

Aplica um **tratamento**, e passa a observar seu efeito entre o objeto de estudo. Requer, portanto, a interferência do pesquisador sobre a população, bem como o controle de fatores externos, com o intuito de medir o efeito desejado.

Exemplos: Estudo do efeito de um novo medicamento, experimentos agrônômicos

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?

Por que?

Computadores

Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

(C) Levantamentos Observacionais

Observa e mede características, mas **não modifica** o objeto de estudo. Os dados são coletados sem que o pesquisador tenha controle sobre as informações obtidas.

Exemplo: Verificar o valor das vendas de uma empresa em um certo período (não há como “selecionar” as vendas)

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?
Por que?

Computadores
Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?

Por que?

Computadores

Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

Para a escolha do método deve-se levar em conta:

- Tipo de pesquisa
- Acessibilidade e disponibilidade dos elementos da população
- Disponibilidade de tempo
- Recursos financeiros e humanos
- ...

Exemplos:

Amostragem por conveniência: elementos selecionados por serem imediatamente disponíveis.

Exemplo: Uma repórter entrevistando pessoas na rua

Amostragem por julgamento: uma pessoa experiente no assunto **escolhe** intencionalmente os elementos a serem amostrados.

Exemplo: Novo produto “testado” entre funcionários

Exemplos:

Amostragem por conveniência: elementos selecionados por serem imediatamente disponíveis.

Exemplo: Uma repórter entrevistando pessoas na rua

Amostragem por julgamento: uma pessoa experiente no assunto **escolhe** intencionalmente os elementos a serem amostrados.

Exemplo: Novo produto “testado” entre funcionários

Atenção

Na amostragem não probabilística, os elementos da população não tem a mesma probabilidade de serem selecionados, portanto **não há garantias da representatividade da população!**

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

Amostragem Aleatória Simples (AAS)

Todas as possíveis amostras de tamanho n tem a mesma chance de serem escolhidas (de uma população com N elementos)

Exemplos:

- Selecionar 10 estudantes de uma sala **por sorteio** e perguntar a idade
- Gerar uma amostra aleatória de 1000 números de matrícula de estudantes da UFPR (no computador!) e perguntar a idade

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

Amostragem Aleatória Simples (AAS)

- É o método mais simples para selecionarmos uma amostra probabilística de uma população
- Serve de base para outros procedimentos amostrais, planejamento de experimentos e estudos observacionais
- Utilizando-se um procedimento aleatório, sorteia-se um elemento da população. Repete-se o processo até que sejam sorteadas as n unidades na amostra.

Amostragem Aleatória Simples (AAS)

Com reposição: o mesmo elemento da população pode ser amostrado mais de uma vez. *A probabilidade de seleção **não** se altera.*

Sem reposição: cada elemento da população é amostrado uma única vez. *A probabilidade de seleção se altera.*

Atenção!

Na prática, em populações *infinitas* (muito grandes), a reposição ou não é irrelevante

Introdução à
Estatística e
conceitos de
amostragem

Estatística
O que é?
Por que?
Computadores
Áreas

Amostragem
Tipos
Métodos
Não probabi-
lísticos
Probabilísticos

Erros
amostrais

Referências

Amostragem Aleatória Simples (AAS)

Do ponto de vista da quantidade de informação contida na amostra, a amostragem *sem reposição* é mais adequada.

No entanto, a amostragem *com reposição* conduz a um tratamento teórico mais simples, pois ele implica que tenhamos **independência** entre as unidades selecionadas.

Portanto, na maioria dos casos quando nos referenciarmos a uma AAS, estamos nos referenciando a uma **amostragem aleatória simples com reposição**.

Amostragem sistemática

Utilizada quando os elementos estão dispostos de maneira organizada (ex.: fila, lista) e **aleatória**.

Escolhe um ponto de partida e seleciona-se cada k -ésimo elemento da população (ex.: o 50^o elemento)

Exemplo:

- Em uma fábrica de lâmpadas, a cada 100 peças produzidas, uma é retirada para teste

Amostragem estratificada

Indicada quando a população está dividida em grupos distintos, denominados **estratos**.

Dentro de cada estrato é realizada uma amostragem aleatória simples. O tamanho da amostra pode ou não ser proporcional ao tamanho do estrato.

Exemplos:

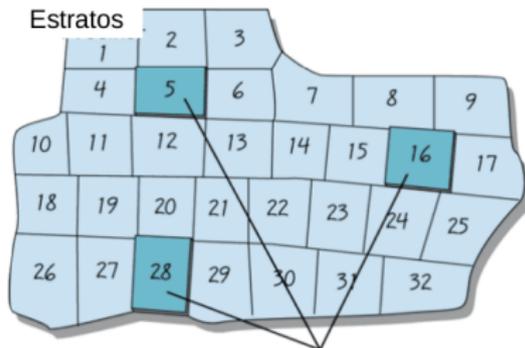
- Uma comunidade universitária com 8000 indivíduos está estratificada da seguinte forma

Estrato	População	Amostra
Professores	800	80
Funcionários	1200	120
Estudantes	6000	600

Amostragem por conglomerado

A área da população é dividida em seções (ou **conglomerados**, ex.: bairros, quarteirões). Os conglomerados são selecionados aleatoriamente. Dentro de um conglomerado, **todos** os elementos são amostrados.

Exemplo:



Todos os elementos
são amostrados

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?
Por que?

Computadores
Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

Erros amostrais

Diferença entre o resultado da amostra e o verdadeiro valor da população. Ocorre pois as amostras são **aleatórias!**

Erros não amostrais

Ocorre quando os dados amostrais são coletados **incorretamente**, devido a uma *amostra tendenciosa*, instrumento de medida defeituoso, anotações erradas, ...

Erros amostrais

Diferença entre o resultado da amostra e o verdadeiro valor da população. Ocorre pois as amostras são **aleatórias!**

Erros não amostrais

Ocorre quando os dados amostrais são coletados **incorretamente**, devido a uma *amostra tendenciosa*, instrumento de medida defeituoso, anotações erradas, ...

Atenção!

Os erros não amostrais não devem existir, ou devem ser minimizados

Não importa quão bem a amostra seja coletada, os **erros amostrais** sempre irão ocorrer

Cada vez que uma amostra aleatória for retirada de uma população, um resultado diferente será observado

Selecione uma amostra de tamanho $n = 5$ das idades dos estudantes de uma sala: 22 21 24 23 20 22 21 25 24 24 23 19 25 24 23 23 20 21 23 20 23 22 23 23 25 25 20 23 24 20

Repita 5 vezes (tente ser o mais aleatório possível!), calcule a média de cada amostra, e compare com a média populacional $\mu = 22,5$

Amostra	\bar{x}	$\epsilon = \bar{x} - \mu$
23 23 23 24 23	23.2	0.7
24 22 20 20 20	21.2	-1.3
21 20 19 22 25	21.4	-1.1
22 23 25 20 22	22.4	-0.1
21 20 22 24 20	21.4	-1.1

- O que isso nos diz a respeito das médias amostrais?
- O que isso nos diz a respeito da variabilidade das médias amostrais?
- E se fizemos uma “média das médias” de todas as amostras?

Voltaremos aqui mais tarde . . .

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?
Por que?

Computadores
Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

- 1 Estatística
 - O que é Estatística?
 - Por que estudar Estatística?
 - Estatística e o uso de computadores
 - Áreas da Estatística
- 2 Conceitos de amostragem
 - Tipos de Amostragem
 - Métodos de amostragem
 - Não probabilísticos
 - Probabilísticos
- 3 Erros amostrais
- 4 Referências

Introdução à
Estatística e
conceitos de
amostragem

Estatística

O que é?

Por que?

Computadores

Áreas

Amostragem

Tipos

Métodos

Não probabi-
lísticos

Probabilísticos

Erros

amostrais

Referências

- Bussab, WO; Morettin, PA. **Estatística básica**. São Paulo: Saraiva, 2002. 526 p. [Cap. 1 e 10]
- Magalhães, MN; Lima, ACP. **Noções de Probabilidade e Estatística**. São Paulo: EDUSP, 2008. [Cap. 1]