

Conceitos Básicos da Inferência Estatística

Prof. Eduardo Bezerra

Inferência Estatística

23 de março de 2018

Roteiro

- 1 População e amostra
- 2 Definição: Inferência Estatística
- 3 Definição: Planos Amostrais
- 4 Estatísticas e parâmetros

Roteiro

- 1 População e amostra
- 2 Definição: Inferência Estatística
- 3 Definição: Planos Amostrais
- 4 Estatísticas e parâmetros

População e amostra

População

População é a totalidade de elementos ou resultados (escores, pessoas, medidas, etc) que estão sob discussão e dos quais se deseja inferir alguma informação.

Amostra

Amostra é qualquer subconjunto de uma população.

População e amostra - exemplo 01

Considere uma escola de 5.000 alunos.

- Se quisermos fazer um estudo das estaturas (e.g., qual a estatura média?), podemos colher uma amostra de, digamos, 40 alunos e estudar o comportamento da **variável estatura** apenas nesses alunos.
- A variável estudada poderia ser inteligência, número de irmãos, número de cáries, notas em história ou renda familiar, dentre outras.

População e amostra - exemplo 02

- Para investigar a *honestidade* de uma moeda, podemos lançá-la 50 vezes e contar o número de caras obtidas.
- A população pode ser considerada como tendo a distribuição da v.a. X , que assume o valor 1, com probabilidade p , se ocorrer cara, e 0 em caso contrário, com probabilidade $1 - p$.
- Ou seja, a população pode ser considerada como tendo distribuição de Bernoulli com parâmetro p .
- A variável ficará completamente especificada quando conhecermos o valor de p .
- A amostra será uma sequência de 50 números, zeros ou uns.

Roteiro

- 1 População e amostra
- 2 Definição: Inferência Estatística
- 3 Definição: Planos Amostrais
- 4 Estatísticas e parâmetros

Inferência Estatística

- O objetivo da Inferência Estatística é produzir afirmações sobre dada característica de uma **população**, com base nos dados contidos em uma **amostra** dessa população.
- Essa característica pode ser representada por uma **variável aleatória**.

Inferência Estatística

Uma população pode ser representada por uma **variável aleatória**, que possui uma determinada **distribuição de probabilidades** com determinado conjunto de **parâmetros**.

- Se tivéssemos acesso à *função de probabilidade* (no caso discreto) ou à *função de densidade* (no caso contínuo) dessa variável aleatória, o problema de fazer afirmações sobre a população estaria resolvido.
- Ocorre que muitas vezes não sabemos nada sobre a variável ou essa informação é parcial. Exemplo: no caso das alturas dos alunos de uma escola, podemos presumir que elas sigam uma distribuição normal, mas, em geral, desconhecemos os **parâmetros** que a caracterizam (média e variância).

Inferência Estatística

Em geral, pode acontecer de

- a) termos uma ideia da forma da curva da função de probabilidades, mas desconhecemos os valores dos parâmetros dessa distribuição.
- b) conhecermos os valores dos parâmetros, mas desconhecemos a forma da curva.
- c) não conhecermos nem a forma da curva nem os respectivos valores dos parâmetros.

Em qualquer caso, o uso de uma amostra pode nos revelar informações sobre o comportamento da variável (população).

Problemas alvos da Inferência Estatística

Dois problemas básicos na Inferência Estatística são:

- a) estimação de parâmetros;
- b) testes de hipóteses sobre parâmetros.

Roteiro

- 1 População e amostra
- 2 Definição: Inferência Estatística
- 3 Definição: Planos Amostrais**
- 4 Estatísticas e parâmetros

Planos amostrais

Existem diversos métodos de amostragem, denominados **planos amostrais**. Alguns deles:

- **Amostragem aleatória sistemática**
- **Amostragem aleatória estratificada**
- **Amostragem aleatória simples**

Amostragem aleatória sistemática

Amostra aleatória sistemática

Os itens ou indivíduos da população são ordenados de alguma forma - alfabeticamente ou por meio de algum outro método. Um ponto de partida aleatório é sorteado, e então cada k -ésimo membro da população é selecionado para a amostra.

Amostragem aleatória estratificada

Amostra aleatória estratificada

A população é inicialmente dividida em subgrupos (estratos) e uma subamostra é selecionada a partir de cada estrato da população.

Amostragem aleatória simples (AAS)

Amostra aleatória simples

Dada uma população de tamanho N , uma **amostra aleatória simples** de tamanho n é um conjunto de n unidades da população, tal que qualquer outro conjunto, dos $\binom{N}{n}$ conjuntos diferentes de n unidades, teria igual probabilidade de ser selecionado.

Amostragem aleatória simples (AAS)

Amostragem Aleatória Simples (AAS):

- A amostra é escolhida de tal forma que cada elemento da população tem a mesma probabilidade de ser selecionado.
- Se a população tem tamanho N , cada elemento dessa população tem a mesma probabilidade igual a $1/N$ de entrar na amostra.

Para construir uma amostra desse tipo, é possível utilizar

- uma **tabela de números aleatórios** para sortear (com mesma probabilidade) os elementos da amostra.
- uma **função para geração de números aleatórios** em algum software estatístico.

Amostragem aleatória simples (AAS)

Amostragem aleatória simples

Uma amostra aleatória simples de tamanho n de uma v.a. X , com dada distribuição, é o conjunto de n variáveis aleatórias independentes X_1, X_2, \dots, X_n , cada uma com a mesma distribuição de X .

Ou seja, a amostra é a n -tupla ordenada (X_1, X_2, \dots, X_n) , onde X_i indica a observação do i -ésimo elemento sorteado.

Amostragem aleatória simples (AAS)

Uma AAS pode ser recolhida *com reposição* ou *sem reposição*.

- quando a amostra é recolhida **com reposição**, cada elemento eventualmente selecionado pode ser selecionado novamente.
- quando a amostra é recolhida **sem reposição**, não há independência entre os elementos, fato que tem impacto na fórmula do cálculo das estimativas feito a partir dessa amostra.

Roteiro

- 1 População e amostra
- 2 Definição: Inferência Estatística
- 3 Definição: Planos Amostrais
- 4 **Estatísticas e parâmetros**

Estatísticas e parâmetros

Parâmetro

Um *parâmetro* é uma medida usada para descrever uma característica de uma população. É uma medida numérica (valor fixo) que descreve uma característica de uma população.

Estatística

Uma *estatística* é uma característica de uma amostra. É ela própria uma variável aleatória.

As *distribuições amostrais* das estatísticas permitem fazer inferências sobre os parâmetros populacionais correspondentes.

Estatísticas e parâmetros - exemplos e notações

Exemplos de parâmetros:

- μ (média populacional),
- σ^2 (variância populacional),
- σ (desvio-padrão populacional),
- p (proporção populacional).

Exemplos de estatísticas:

- \bar{X} (média amostral),
- S^2 (variância amostral),
- S (desvio padrão amostral),
- \hat{p} (proporção amostral).

Estatísticas e parâmetros - exemplos

Considere a variável aleatória “idade dos interessados em xadrez”.

- A quantidade total de pessoas interessadas em xadrez é um parâmetro.
- A esperança desta variável aleatória, digamos 35 anos, também é um parâmetro.

Agora considere que tomemos uma AAS dessa população.

- a média amostral, i.e., o valor obtido quando tomamos os elementos da amostra e calculamos a média de suas idades, é uma estatística.