Unidade V

quarta-feira, 13 de outubro de 2021 14:23

Unidade V - Noções de amostragem e estimação

- População. Censo e amostragem.
- Amostra aleatória. Estimador e estimativa.
- Intervalos de confiança para a proporção e média.

CONCEITOS BÁSICOS:

UnB – Universidade de Brasília
FGA – Faculdade UnB Gama
Graduação – ciclo básico

Probabilidade e Estatística aplicada à Engenharia

Profa. Marília Miranda

- Estimação de parâmetros: conceitos

INTRODUÇÃO

- Relembrando...

Métodos Estatísticos

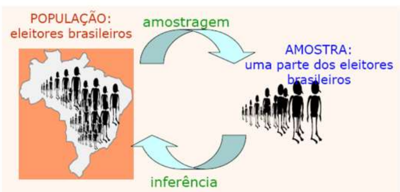
Estatística Descritiva

Inferência Estatística

INTRODUÇÃO

2 → Pi (proporção) → Mi (média)

- Inferência Estatística:
  - *Estimação de parâmetros*: generalizar resultados de uma amostra para a população de onde ela foi extraída;
  - *Testes de Hipóteses*: testar hipóteses com base em amostras.



ESTIMAÇÃO DE PARÂMETROS

- Estudaremos o problema de avaliar certas

PROPORÇÃO:

UnB – Universidade de Brasília
FGA – Faculdade UnB Gama
Graduação – ciclo básico

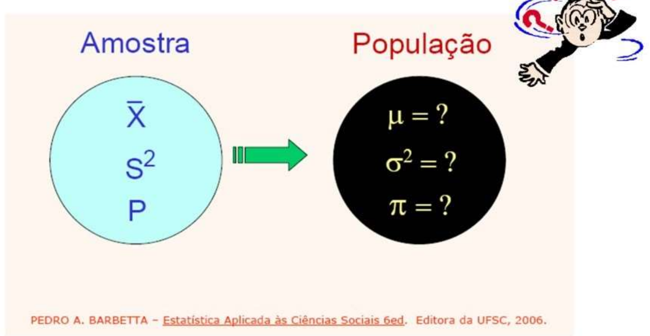
Probabilidade e Estatística aplicada à Engenharia

Profa. Marília Miranda

- Estimação de uma proporção

INTRODUÇÃO

- Com base em uma amostra, como estimar parâmetros populacionais?



ESTIMAÇÃO DE UMA PROPORÇÃO

- Vamos considerar o caso em que o tamanho da amostra é razoavelmente grande e o atributo em observação não seja muito raro ou quase certo, de tal forma que seja válida a aproximação para a distribuição Normal;
- Suponha que a população de onde foi extraída essa amostra seja muito grande, não necessitando considerar o seu tamanho nos cálculos;

1° Cálculo

→ Erro padrão

Variabilidade em torno de P.

ESTIMAÇÃO DE UMA PROPORÇÃO

MÉDIA:

UnB – Universidade de Brasília
FGA – Faculdade UnB Gama
Graduação – ciclo básico

Probabilidade e Estatística aplicada à Engenharia

Profa. Marília Miranda

-Estimação de uma média

ESTIMAÇÃO DE UMA MÉDIA

- Em linhas gerais...
  - Passo 1: cálculo do erro padrão $S_{\bar{x}}$
  - Passo 2: cálculo da margem de erro E, dado um nível de confiança
  - Passo 3: obtenção do intervalo de confiança para o parâmetro μ

IC(μ;nível de confiança)

ESTIMAÇÃO DE UMA MÉDIA

- Quando a variável em estudo é quantitativa, normalmente se tem interesse no parâmetro μ (média);
- Tendo uma amostra aleatória simples da população de interesse, podemos ter uma estimativa de μ através do cálculo da média dos valores da amostra:

$$\boxed{\bar{X} = \frac{\sum X}{n}}$$

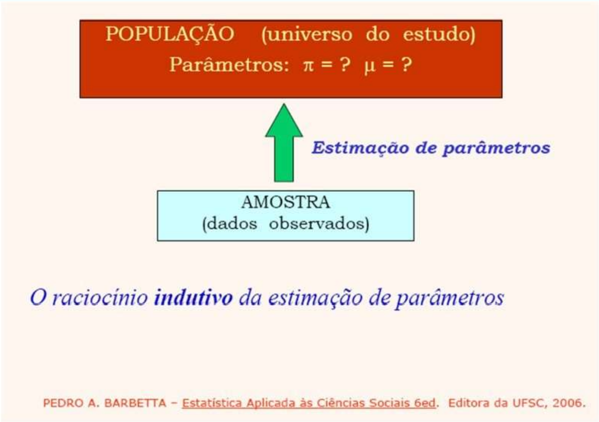
relembrando

ESTIMAÇÃO DE UMA MÉDIA

características dos elementos da população, a partir de operações com os dados de uma amostra;

- Raciocínio indutivo: resultados da *parte* (amostra) são generalizados para o *todo* (população) → este procedimento é denominado **estimação de parâmetros**.

ESTIMAÇÃO DE PARÂMETROS



ESTIMAÇÃO DE PARÂMETROS

- Conceitos básicos:

  - **Parâmetro**: alguma característica descritiva dos elementos da população, como por exemplo, a média e a variância de alguma variável, a proporção de algum atributo;

  - **Estatística**: alguma operação com os dados da amostra a ser selecionada. Por exemplo, uma média ou uma proporção a serem calculadas com estes dados;

ESTIMAÇÃO DE PARÂMETROS

- Conceitos básicos:

  - A estatística, quando usada com o objetivo de avaliar, ou *estimar*, o valor de algum parâmetro, também é chamada de **estimador**;

  - **Erro amostral**: corresponde à diferença entre a estatística (a ser calculada a partir de uma amostra) e o parâmetro (característica dos elementos de uma população).

- Com essas suposições, o desvio padrão da distribuição amostral de **P**, σ_P, também conhecido como erro padrão de **P**, pode ser estimado pelos dados da amostra, usando a expressão:

$$S_p = \sqrt{\frac{P\cdot(1-P)}{\underbrace{n}_{\text{tamanho da amostra}}}}$$

onde P é a proporção do atributo, na amostra; e n é o tamanho da amostra.

ESTIMAÇÃO DE UMA PROPORÇÃO

- Se o tamanho da população, N, for conhecido e não muito grande (N < 20n), então estima-se o erro padrão da proporção por:

$$S_p = \sqrt{\frac{P\cdot(1-P)}{n}}\cdot\sqrt{\frac{N-n}{N-1}}$$

ESTIMAÇÃO DE UMA PROPORÇÃO

- **Intervalo de confiança (IC) para π**

  - De modo geral:

IC(π;nível de confiança) = P ± E ou (P-E, P+E)

é dito um **intervalo de confiança** para o parâmetro π, com nível de confiança de 90%, 95% ou 99% (valores mais usuais) e margem de erro = E.

2° Cálculo → Margem de erro (E)

ESTIMAÇÃO DE UMA PROPORÇÃO

- **Intervalo de confiança (IC) para π**

  - Cálculo da margem de erro:

$$E = \underbrace{Z}_{\text{tabela}} . \underbrace{Sp}_{\text{erro padrão}}$$

*Valores de Z (relacionados com o nível de confiança elegido) vem da tabela da distribuição Normal e correspondem:

Tabelas de distribuição normal padrão

$$\boxed{\begin{array}{l}\text{90\% de confiança: Z = 1,65}\\ \text{95\% de confiança: Z = 1,96}\\ \text{99\% de confiança: Z = 2,58}\end{array}}$$

- Como o valor de X vai depender da amostra selecionada, podemos falar em erro padrão e em distribuição amostral de X . O erro padrão de $\overline{X}$ pode ser estimado com os dados da amostra por (Passo 1):

1° PASSO

$$\underbrace{S_{\overline{X}}}_{\text{Erro padrão de }\overline{x}} = \frac{\overset{\text{desvio padrão}}{S}}{\sqrt{n}}$$

ESTIMAÇÃO DE UMA MÉDIA

Obs.: Se o tamanho da população, N, for conhecido e não muito grande (N < 20n), então estima-se o erro padrão da média por:

$$S_{\overline{X}} = \frac{S}{\sqrt{n}}\cdot\sqrt{\frac{N-n}{N-1}}$$

$$\boxed{\text{if } (n \geq 30)}$$

∴

ESTIMAÇÃO DE UMA MÉDIA

2° MARGEM DE ERRO

- Margem de erro na estimação de uma média (Passo 2):

  → Se a amostra for grande (n > 30): $E = z\cdot S_{\overline{X}}$

  - onde z vem da distribuição normal

*Valores de Z (relacionados com o nível de confiança elegido) vem da tabela da distribuição Normal e correspondem:

90% de confiança: Z = 1,65
95% de confiança: Z = 1,96
99% de confiança: Z = 2,58

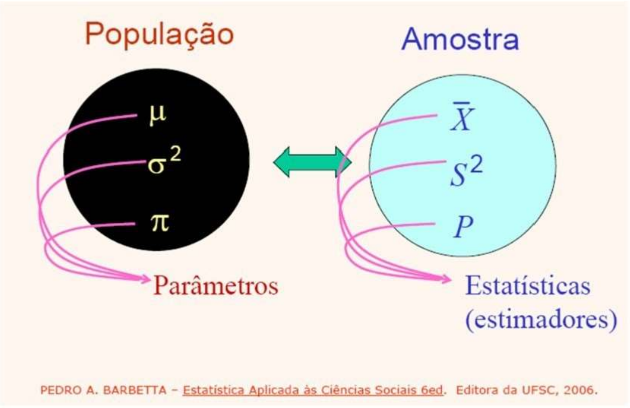
ESTIMAÇÃO DE UMA MÉDIA

- **Intervalo de confiança (IC) para μ**

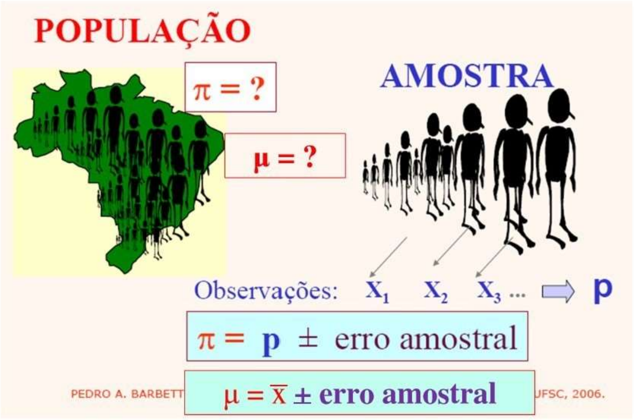
  - De modo geral (Passo 3):

IC(μ;nível de confiança) = $\overline{x}$ ± E ou ($\overline{x}$-E, $\overline{x}$+E)

ESTIMAÇÃO DE PARÂMETROS



ESTIMAÇÃO DE PARÂMETROS



ESTIMAÇÃO DE PARÂMETROS

- Alguns parâmetros e as respectivas estatísticas que geralmente são usadas para estimá-los:

PARÂMETROS (características da **população**)	ESTATÍSTICAS (características da **amostra**)
π = proporção de algum atributo, dentre os elementos da população.	P = proporção de elementos com o atributo, dentre os que serão observados na amostra.
μ = média de alguma variável quantitativa, nos elementos da população.	$\overline{X}$ = média da variável, a ser calculada com os elementos da amostra.
σ = desvio padrão de uma variável, dentre os elementos da população.	S = desvio padrão da variável, a ser calculado com os elementos da amostra.

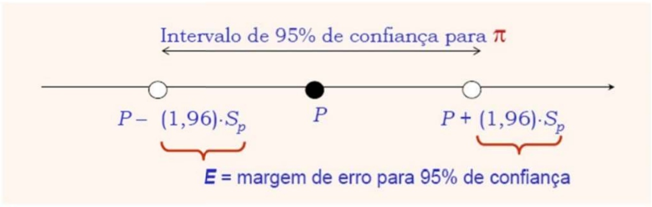
ESTIMAÇÃO DE PARÂMETROS

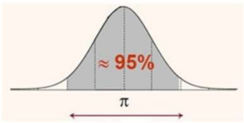
- Ao observar uma particular amostra, podemos calcular o valor da estatística que estamos usando como estimador:

- o valor encontrado é chamado de **estimativa**;

- Exemplo: Se em uma amostra n = 400 moradores encontrarmos 240 favoráveis a um projeto implementado pela prefeitura do

ESTIMAÇÃO DE UMA PROPORÇÃO





ESTIMAÇÃO DE UMA PROPORÇÃO

- Em linhas gerais...

  - Passo 1: cálculo do erro padrão Sp

  - Passo 2: cálculo da margem de erro E, dado um nível de confiança

  - Passo 3: obtenção do intervalo de confiança para o parâmetro π

  IC(π;nível de confiança)

ESTIMAÇÃO DE UMA PROPORÇÃO

- Voltando ao exemplo das aulas anteriores...

  - Suponha que na amostra de n = 400 pessoas, encontramos 60% de favoráveis. Temos, então, P = 0,60 (ou 60%), com erro padrão (Passo 1):

1º

$$S_p = \sqrt{\frac{P\cdot(1-P)}{n}} = \sqrt{\frac{(0{,}60)\cdot(0{,}40)}{400}} = 0{,}0245$$

ESTIMAÇÃO DE UMA PROPORÇÃO

= 1,96

- Usando nível de confiança de 95%, temos a margem de erro (Passo 2):

2º

$E = 2 \cdot S_p$

$$E = (1{,}96)\cdot S_p = (1{,}96)\cdot(0{,}0245) = 0{,}048 \text{ (ou 4,8\%)}$$

3º

- Passo 3, o IC é representado por:

60,0% ± 4,8%

é dito um **intervalo de confiança** para o parâmetro μ, com nível de confiança de 90%, 95% ou 99% (valores mais usuais) e margem de erro = E.

ESTIMAÇÃO DE UMA MÉDIA

- **Exemplo**: seja μ = ganho médio de peso durante o primeiro ano letivo, na população de crianças da rede municipal de ensino, devido a uma merenda especial.

- Deseja-se estimar esse parâmetro μ

ESTIMAÇÃO DE UMA MÉDIA

- Solução:

Numa amostra aleatória simples de n = 100 crianças do primeiro ano letivo, em que se estava servindo a merenda especial, foram obtidos os seguintes resultados:

Ganho médio de peso: $\overline{X}$ = 6,0 kg;

Desvio padrão: S = 2,0 kg.

Procedendo a estimativa do erro padrão de $\overline{X}$:

$$S_{\overline{X}} = \frac{S}{\sqrt{n}} = \frac{2{,}0}{\sqrt{100}} = 0{,}2\ kg \qquad \text{Passo 1}$$

O limite superior para o erro amostral (nível de confiança de 95%):

$n > 30 \therefore E = z \cdot S_{\overline{X}}$

Margem de erro: $E = (1{,}96)\cdot(0{,}2) = 0{,}392$ kg Passo 2

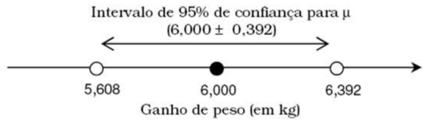
donde resulta o seguinte intervalo de 95% de confiança para μ:

6,000 ± 0,392 kg. Passo 3

(6-0,392;6+0,392) = (5,608;6,392)

ESTIMAÇÃO DE UMA MÉDIA

- Ou seja, a partir do acompanhamento da amostra das cem crianças, chegamos à conclusão de que o intervalo de 5,608 a 6,392 kg contém, com 95% de confiança, o ganho médio de peso, μ, de todas as crianças do primeiro ano da rede municipal de ensino, que venham a ser submetidas à merenda especial. Esquematicamente:



município, então temos a seguinte estimativa para o parâmetro π : $P = \frac{240}{400} = 0{,}60$ (ou, 60%)

ESTIMAÇÃO DE PARÂMETROS

- Importante:

  - não devemos esperar que o valor P coincida com o parâmetro π, devido ao que chamamos de erro amostral (o mesmo vale para a média);

  - Dizemos que uma estimativa é tão mais *precisa* quanto menor for o seu erro amostral.

ESTIMAÇÃO DE PARÂMETROS

- Um dos principais objetivos na teoria da estimação é estimar um *limite superior provável* para o erro amostral → este valor será a base para avaliarmos a precisão da nossa estimativa.
- Nesta disciplina estaremos preocupados em avaliar a precisão de estimativas de parâmetros do tipo π (proporção de algum atributo) e do tipo μ (média de alguma variável quantitativa).

o intervalo de limite inferior 60,0% – 4,8% = 55,2% e de limite superior 60,0% + 4,8% = 64,8%.

IC(π;95%) = (55,2%; 64,8%)

ou seja

ESTIMAÇÃO DE UMA PROPORÇÃO

- Podemos dizer, com nível de confiança de 95%, que o intervalo 60,0% ± 4,8% contém o parâmetro π (proporção de favoráveis em toda a população).

ESTIMAÇÃO DE UMA PROPORÇÃO

- Outros níveis de confiança:

Área	0,800	0,900	**0,950**	0,980	0,990	0.995	0,998
z	1,282	1,645	**1,960**	2,326	2,576	2,807	3,090

Margem de erro: $\boxed{E = z \cdot S_P}$

Intervalo de confiança para π: $\boxed{P \pm E}$