Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas
IESB

INFERÊNCIA ESTATÍSTICA
CONCEITOS INICIAIS
(Unidade 1)

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)

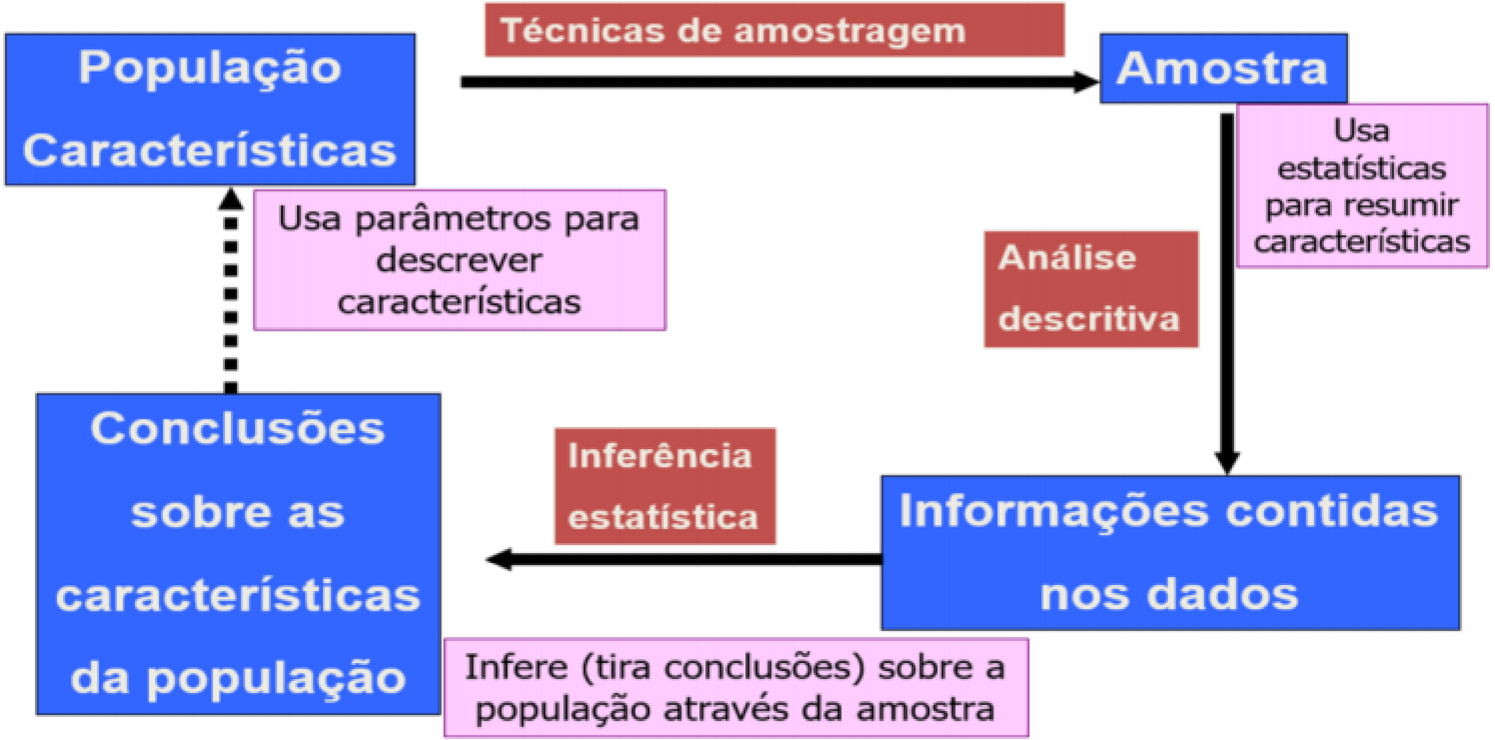
A Estatística é uma ciência que tem como objetivo a tomada de decisão em situações de incerteza. Esta ciência, como já vimos, divide-se basicamente em duas partes: a primeira parte é conhecida como **Estatística Descritiva** que trata da organização e descrição de dados e, a segunda, é a **Estatística Inferencial** que se preocupa em fazer afirmações e/ou testar hipóteses sobre características populacionais a partir de dados amostrais.

Além das divisões apresentadas anteriomente, podemos a grosso modo dividir a Estatística em uma terceira parte ao qual chamamos de **Probabilidade**. A teoria das probabilidades nos permite modelar fenômenos aleatórios, ou seja, aqueles em que está presente a ***incerteza***. É uma ferramenta fundamental para a Inferência Estatística.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)



Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Definição de População e Amostra

- População: é um conjunto formado por todos os elementos que possuem pelo menos uma característica em comum associada ao interesse da pesquisa.
- Amostra: é apenas uma parte da população, ou seja, é um subconjunto da população.

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

INFERÊNCIA ESTATÍSTICA (CONCEITOS INICIAIS)
DEFINIÇÃO DE PARÂMETRO E ESTATÍSTICA

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas
IESB

- Parâmetro: é uma característica numérica da população.
- Estatística: é uma característica numérica da amostra que será usada para revelar informações sobre a população.

População: *os eleitores da cidade de Campina Grande* ***Amostra:*** *650 eleitores escolhidos de preferência aleatoriamente (ao acaso)* ***Característica de interesse:*** *percentual de eleitores que planejam votar num candidato A nas próximas eleições.*

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

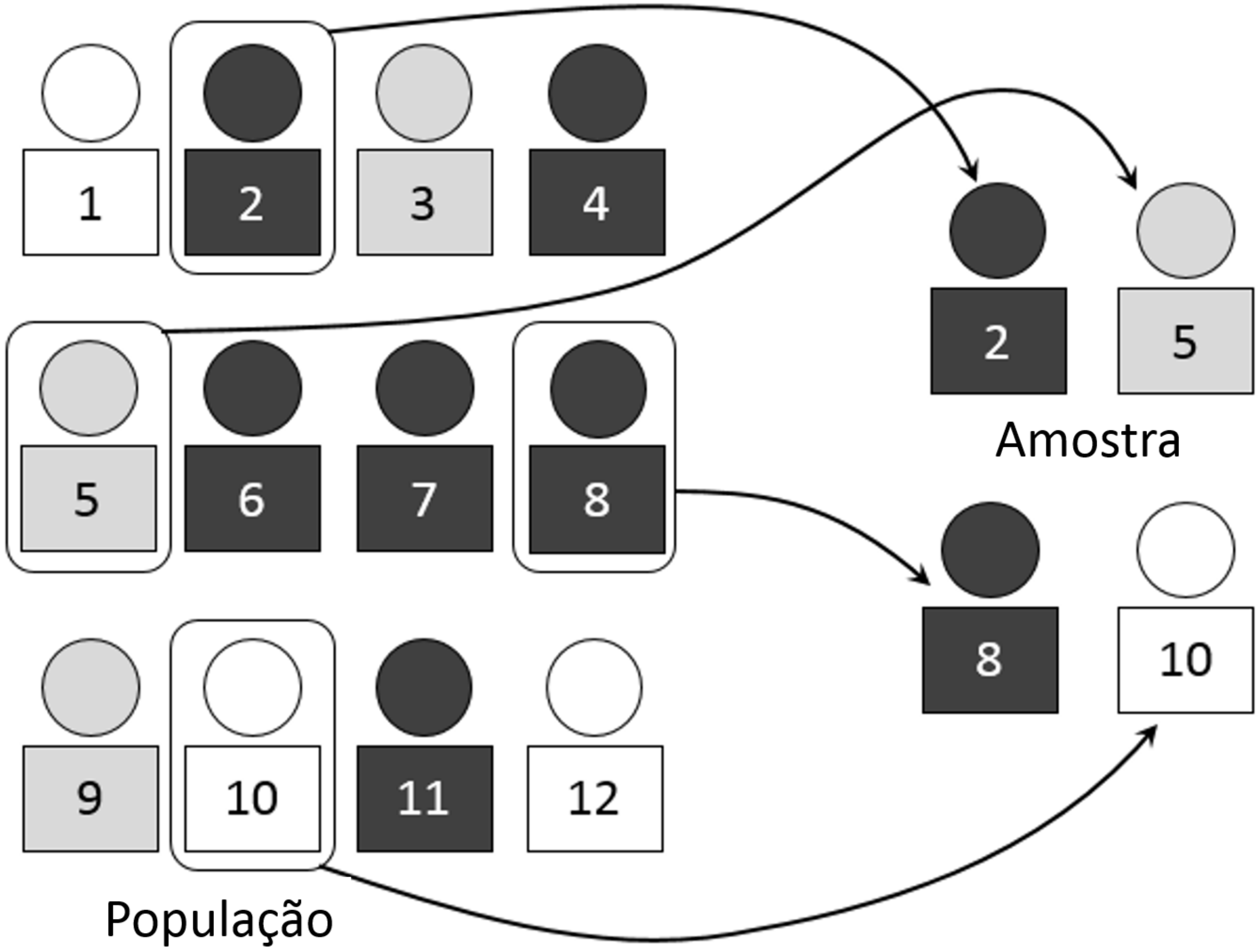
Inferência Estatística (Conceitos Iniciais)
Amostra Aleatória Simples

- Amostragem aleatória simples:

- ✓ tipo de amostragem probabilística mais utilizada;
- ✓ todos os elementos da população têm a mesma probabilidade de pertencerem à amostra;
- ✓ população homogênea;
- ✓ selecionar uma amostra “n” a partir de uma população “N”;
- ✓ sorteio / tabela de números aleatórios;
- ✓ com / sem reposição.

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Amostra Aleatória Simples



Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

INFERÊNCIA ESTATÍSTICA (CONCEITOS INICIAIS)
AMOSTRA ALEATÓRIA SIMPLES

Exemplo: um pesquisador deseja comparar os teores médios de proteína de três cultivares de cevada. Para executar o experimento ele dispõe de uma área de terra homogênea (mesma fertilidade, mesma umidade, etc.) de tamanho 288 m². Portanto, as três cultivares serão comparadas em igualdade de condições. Um princípio básico da experimentação é o uso de repetições, ou seja, são necessários pelo menos dois valores para cada cultivar. Assim, a área total vai ser dividida em 12 canteiros de tamanhos 6m x 4m, totalizando 24m²/canteiro.

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

(tabela de números aleatórios) → {03,05,01,08,07,02,12,10,11,04,09,06}.

Canteiro 1 **Cultivar A**	Canteiro 2 **Cultivar B**	Canteiro 3 **Cultivar A**
Canteiro 4 **Cultivar C**	Canteiro 5 **Cultivar A**	Canteiro 6 **Cultivar C**
Canteiro 7 **Cultivar B**	Canteiro 8 **Cultivar A**	Canteiro 9 **Cultivar C**
Canteiro 10 **Cultivar B**	Canteiro 11 **Cultivar C**	Canteiro 12 **Cultivar B**

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Distribuição amostral da média: considere uma população identificada pela variável X, cujos parâmetros média populacional $\mu = E(X)$ e variância populacional $\sigma^2 = Var(X)$ são supostos conhecidos.

Retire todas as AAS de tamanho n dessa população e, para cada uma calcular $\bar{X}$.

Em seguida vamos considerar a distribuição.

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Exemplo: seja a população {1, 3, 5, 5, 7}. Retire todas as amostras de tamanho 2 com reposição. A distribuição conjunta da variável bidimensional (X_1, X_2) é:

X_2 \ X_1	1	3	5	7	Total
1	1/25	1/25	2/25	1/25	1/5
3	1/25	1/25	2/25	1/25	1/5
5	2/25	2/25	4/25	2/25	2/5
7	1/25	1/25	2/25	1/25	1/5
Total	1/5	1/5	2/5	1/5	1

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Qual a distribuição da estatística $\bar{X}$?

$\bar{X} = 3$ equivale à A={(1,5); (3,3) ; (5,1)}.

P(A) = 1/5.

De forma análoga,

$\bar{x}$	1	2	3	4	5	6	7	Total
$P(\bar{X} = \bar{x})$	1/25	2/25	5/25	6/25	6/25	4/25	1/25	1,00

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Exemplo: sabe-se que $\mu=4{,}2$ e $\sigma^2=4{,}16$. A distribuição amostral da média foi dada anteriormente, da qual obtemos:

$$E(\bar{X}) = \sum_i \bar{x}_i \, p_i$$

$$= 1x\frac{1}{25} + 2x\frac{2}{25} + 3x\frac{5}{25} + 4x\frac{6}{25} + 5x\frac{6}{25} + 6x\frac{4}{25} + 7x\frac{1}{25} = 4{,}2$$

De modo análogo,

$$Var(\bar{X}) = 2{,}08$$

O que vocês podem observar?

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Teorema: seja X uma v.a. com média μ e variância σ^2, e seja $(X_1, \ldots, X_n)$ uma AAS de X. Então:

$$E(\bar{X}) = \mu e Var(\bar{X}) = \frac{\sigma^2}{n}$$

Determinamos a média e a variância da distribuição amostral da média.

$$E(\bar{X}) = \frac{1}{n}\left(E(X_1)+\ldots+E(X_n)\right) = \frac{1}{n} x n\mu = \mu$$

$$Var(\bar{X}) = \frac{1}{n^2}\left(Var(X_1)+\ldots+Var(X_n)\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Qual a forma da distribuição dessa estatística?

Para amostras aleatórias simples (X_1,..., X_n) retiradas de uma população com média μ e variância σ^2 finita, a distribuição amostral da média aproxima-se para n grande, de uma distribuição normal com média e variância dada anteriormente (TLC).

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

(Teorema Limite Central) *Sejam* $(X_1, X_2, \ldots, X_n)$ *uma sequência de v.a's i.i.d. cada uma tendo média* μ *e variância* σ^2. *Então, para* n *grande, a distribuição de* $X_1 + X_2 + \ldots + X_n$ *é aproximadamente normal com média* $n\mu$ *e variância* $n\sigma^2$. *Segue do Teorema Central do Limite que* $\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \approx Normal(0,1)$. *Então, para* n *grande,*

$$\mathcal{P}\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq x\right) \approx \mathcal{P}(Z \leq z),\ \ Z \sim Normal(0,1).$$

Observações:

1. No teorema acima não fizemos nenhuma suposição sobre a natureza das distribuições das variáveis $X_1, X_2, \ldots, X_n$, ou seja, independente de como se comportam essas variáveis, sejam elas discretas ou contínuas, o teorema continua válido.

2. Se as variáveis $X_1, X_2, \ldots, X_n$ têm distribuição normal, então $\overline{X}$ terá também distribuição normal e não apenas uma aproximação.

A demonstração completa deste teorema exigiria recursos dos quais não dispomos, portanto não será dada, mas o importante é sabermos como esse resultado pode ser usado. O aluno interessado pode consultar Bussab (2008), por exemplo.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Exemplo: considere uma máquina que enche pacotes cujos pesos seguem uma distribuição N(500, 100). Colhendo uma amostra de n=100 pacotes e pesando-os, pelo que foi dito acima, a média terá uma distribuição normal com média 500 e variância 1. Logo, se a máquina estiver regulada, a probabilidade de encontrarmos a média de 100 pacotes diferindo de 500g de menos de 2 gramas será?

$$P(|\bar{X} - 500| < 2) = P(498 < \bar{X} < 502) = 95\%$$

Ou seja, dificilmente 100 pacotes terão uma média fora do intervalo (498,502).

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Obtenção da distribuição amostral da estatística T

O problema da inferência consiste em fazer uma afirmação sobre os parâmetros da população através da amostra.

Suponha que a afirmação deva ser feita sobre um parâmetro Θ da população (média, variância, proporção, dentre outros).

Será usada uma AAS de n elementos sorteados dessa população.

A decisão será baseada na estatística T que é uma função da amostra $(X_1, X_2, \ldots, X_n)$, ou seja, $T=f(X_1, X_2, \ldots, X_n)$.

Após a coleta da amostra será observado um valor qualquer de T (t_0) e, baseado nesse valor é que se faz a afirmação sobre Θ.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

OBTENÇÃO DA DISTRIBUIÇÃO AMOSTRAL DA ESTATÍSTICA T

A resposta sobre essa afirmação seria mais rica se soubéssemos o que aconteceria com a estatística T, quando retiramos todas as amostras de uma população conhecida segundo o plano amostral adotado.

Qual a distribuição de T quando $(X_1, X_2, \ldots, X_n)$ assume todos os valores possíveis? Essa distribuição é chamada de *distribuição amostral da estatística T* e desempenha importante papel na inferência.

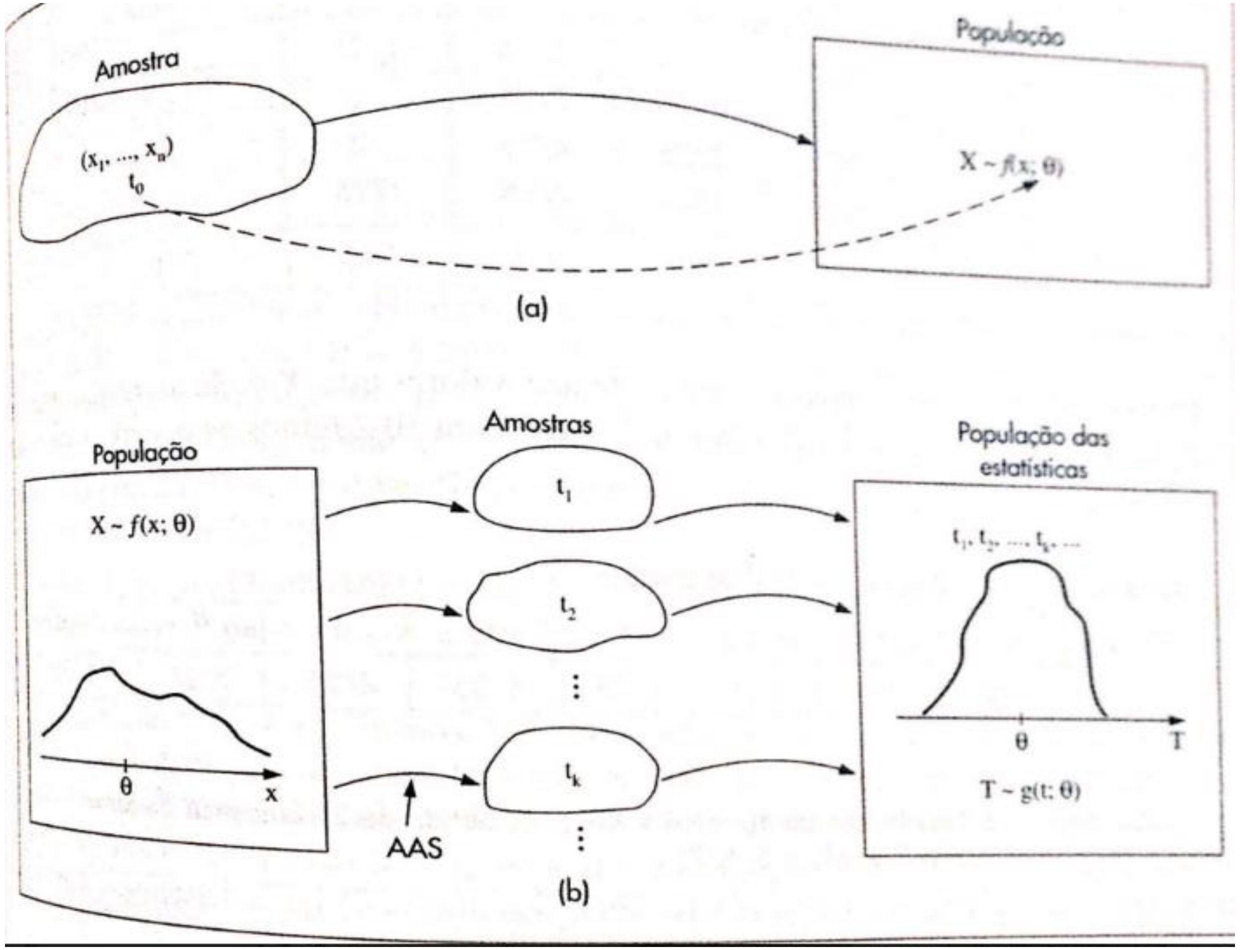
O principal objetivo é identificar um modelo que explique a distribuição amostral de T e, essa irá depender da distribuição de X e do plano amostral adotado.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Obtenção da distribuição amostral da estatística T

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas



Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Obtenção da distribuição amostral da estatística T

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas

(Estatística) *Uma estatística é uma característica da amostra, ou seja, dada uma amostra aleatória $X_1, X_2, ..., X_n$ de uma população X, definiremos uma estatística T como qualquer função de $X_1, X_2, ..., X_n$, ou seja $T = f(X_1, X_2, ..., X_n)$.*

Assim, dada uma amostra aleatória $X_1, X_2, ..., X_n$, um exemplo de estatística seria a média amostral

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n).$$

Sendo $X_1, X_2, ..., X_n$ uma amostra aleatória, uma pergunta interessante seria: O que acontece com a estatística T quando retiramos todas as amostras possíveis de uma população conhecida segundo o plano amostral adotado, ou seja, qual a distribuição da estatística T quando $X_1, X_2, ..., X_n$ assume todos os valores possíveis. Essa distribuição é chamada de **distribuição amostral da estatística** T e desempenha papel fundamental na teoria da inferência estatística.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Distribuição amostral da proporção: considere uma população em que a proporção de elementos com certa característica é p. Seja X uma v.a. da seguinte forma:

X=1, se o indivíduo for portador da característica

X=0, se o indivíduo não for portador da característica

Logo,

$$E(X)=p$$

$$Var(X)=p(1-p)$$

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Retirada uma AAS dessa população e indicando Y_n como o total de indivíduos portadores da característica na amostra,

$$Y_n \sim b(n,p)$$

Vamos definir por $\hat{p}$ a proporção de portadores da característica na amostra:

$$\hat{p} = \frac{Y_n}{n}$$

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Logo,

$$P(Y_n = k) = P(Y_n/n = k/n) = P(\hat{p} = k/n)$$

Ou seja, a distribuição amostral de $\hat{p}$ é obtida da distribuição de

Y_n = n x média amostral

Sendo E(X) = p e Var(X) = [p(1-p)].

Pelo TLC:

$$\bar{X} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

$$Y_n \sim N(np, np(1-p))$$

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Logo,

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Exemplo: suponha que p=30% dos estudantes de uma escola sejam mulheres. Coletamos uma AAS de n=10 estudantes e calculamos a proporção de mulheres na amostra. Qual a probabilidade de que o estimador desvie do parâmetro em menos de 0,01?

$$P(|\hat{p}-p|<0{,}01)=P(-0{,}01<\hat{p}-p<0{,}01)$$

$$\hat{p}-p\sim N\left(0,\frac{p(1-p)}{n}\right)$$

Substituindo p = 0,3.

$$P\left(\frac{-0{,}01}{\sqrt{0{,}021}}<\hat{p}-p<\frac{0{,}01}{\sqrt{0{,}021}}\right)=0{,}056$$

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

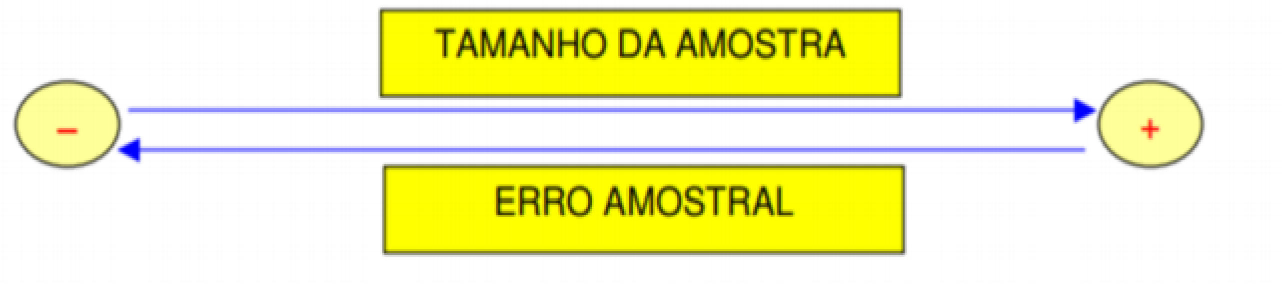
Inferência Estatística (Conceitos Iniciais) Distribuição amostral da proporção e determinação do tamanho da amostra

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas

Do mesmo modo que estudamos a distribuição amostral de $\overline{X}$ e de $\widehat{p}$, podemos encontrar a distribuição amostral de qualquer estatística $T = f(X_1, X_2, ..., X_n)$ mas quanto mais complexa for essa relação f mais difícil será a derivação matemática das propriedades dessa estatística. Este é o caso da variância amostral S^2 em que é possível demonstrar que, a menos de uma constante, tem distribuição Qui-Quadrado com $n-1$ graus de liberdade (χ^2_{n-1}), **desde que** a amostra seja proveniente de uma população com distribuição normal. Note, portanto, que no caso da distribuição amostral da variância S^2; mais suposições como a normalidade das variáveis $X_1, X_2, ..., X_n$ são exigidas enquanto que pelo TLC não há esta exigência.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br



Erro Amostral é a diferença entre um resultado amostral e o verdadeiro resultado populacional; tais erros resultam de flutuações amostrais aleatórias.

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais) Distribuição amostral da proporção e determinação do tamanho da amostra

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas
IESB

Suponha que estejamos estimando a média μ populacional e para isso usaremos a média amostral, $\bar{X}$, baseada numa amostra de tamanho n. Suponha ainda que se queira determinar o valor de n de modo que

$$P(|\bar{X} - \mu| \leq \epsilon) = \gamma,$$

com $0 < \gamma < 1$ e $\epsilon > 0$ sendo o erro amostral máximo admissível, ambos valores fixados.

Como $\bar{X} \approx N(\mu, \sigma^2/n)$, então $\bar{X} - \mu \approx N(0, \sigma^2/n)$ e, portanto,

$$P\left(|\bar{X} - \mu| \leq \epsilon\right) = P\left(-\epsilon \leq \bar{X} - \mu \leq \epsilon\right) = P\left(\frac{-\sqrt{n}\epsilon}{\sigma} \leq Z \leq \frac{\sqrt{n}\epsilon}{\sigma}\right) \cong \gamma$$

$$\Rightarrow 2P\left(0 \leq Z \leq \frac{\sqrt{n}\epsilon}{\sigma}\right) \cong \gamma,$$

onde $Z = (\bar{X} - \mu)\sqrt{n}/\sigma$. Logo, podemos obter $z_{\gamma/2}$ da $N(0,1)$, tal que $P(0 \leq Z \leq z_{\gamma/2}) = \frac{\gamma}{2}$, de modo que

$$z_{\gamma/2} = \frac{\sqrt{n}\epsilon}{\sigma},$$

de onde obtemos finalmente

$$n = \frac{z_{\gamma/2}^2 \sigma^2}{\epsilon^2}.$$

Observação: Na prática, não se conhece o valor da variância populacional σ^2 e para resolver este problema pode-se utilizar uma pequena amostra piloto para estimar o valor da variância populacional ou então tomar como base alguma informação prévia sobre a mesma.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Distribuição amostral da proporção e determinação do tamanho da amostra

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas

Um economista deseja estimar a renda média para o primeiro ano de trabalho de um bacharel em direito. Quantos valores de renda devem ser tomados, se o economista deseja ter 95% de confiança em que a média amostral esteja a menos de R$500,00 da verdadeira média populacional? Suponha que saibamos, por um estudo prévio, que para tais rendas, $\sigma =$ R$6250,00.

Queremos determinar o tamanho de n da amostra, dado que $\alpha = 0,05(95\%)$ de confiança. Desejamos que a média amostral seja a menos de R$500,00 da média populacional, de fomra que $\epsilon = 500$. Supondo $\sigma = 6250$, e aplicando a equação (3), obtendo

$$n = \frac{z^2_{\gamma/2}\sigma^2}{\epsilon^2} = \left(\frac{z_{\gamma/2}\sigma}{\epsilon}\right)^2 = \left(\frac{1,96 \times 6250}{500}\right)^2 = 600,25 \approx 601 \quad \text{(arredondamos para cima).}$$

Devemos, portanto, obter uma amostra de ao menos 601 rendas de primeiro ano, selecionadas aleatoriamente, de bacharéis de faculdades que tenham feito um curso de direito. Com tal amostra teremos 95% de confiança que a média amostral $\overline{X}$ difira em menos de R$500,00 da verdadeira média populacional μ.

No R, basta executarmos

```
1-0.95=0.05/2=0.025. Daí, 0.95+0.025=0.975 (busca na tabela Normal ou no R qnorm(0.975))
```

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Estimação pontual: definição de espaço paramétrico

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas

- Estimação de parâmetros: um conjunto de técnicas baseadas em probabilidade, que a partir de dados amostrais nos permite tirar conclusões sobre a população de interesse.
- Estimador: é uma função dos elementos da amostra, que será usada no processo de estimação do parâmetro desejado. O estimador é, como vemos, uma estatística. Será, portanto, uma variável aleatória caracterizada por uma distribuição de probabilidade e seus respectivos parâmetros próprios.
- Estimativa: cada valor particular assumido por um estimador.
- Estimador pontual e intervalar.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Estimação pontual: definição de estimador pontual

θ : parâmetro (valor numérico constante e desconhecido da população)

$\hat{\theta}$: estimador pontual (estatística visando estimar o parâmetro)

Um estimador $\hat{\theta}$ é uma função dos valores amostrais que é usado para estimar o valor de um parâmetro desconhecido θ. O estimador é uma variável aleatória com uma distribuição de probabilidade. Quando uma amostra aleatória é selecionada de uma população e $\hat{\theta}$ é calculado a partir dos dados, o valor numérico obtido é chamado uma estimativa de da amostra considerada.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Estimação pontual: definição de estimador pontual

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas

Alguns estimadores pontuais:

Parâmetro da população (θ)	**Estimador ($\hat{\theta}$)**
Média (μ)	Média amostral ($\bar{x}$)
Proporçao (p)	Proporção amostral ($\bar{p}$)
Desvio-padrão (σ)	Desvio-padrão amostral (S)

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

INFERÊNCIA ESTATÍSTICA (CONCEITOS INICIAIS)
ESTIMADOR NÃO VIESADO E ESTIMADOR CONSISTENTE

(Vício de um Estimador) *O vício de um estimador é dado por*

$$B(T) = E(T) - \theta.$$

(Estimador não viciado) *Um estimador T é dito ser não viciado para o parâmetro θ se $B(T) = 0$. Ou seja, se $E(T) = \theta$.*

(Estimador consistente) *Uma sequência T_n de estimadores de θ é consistente se, à medida que o tamanho da amostra aumenta, seu valor esperado converge para o parâmetro de interesse e sua variância converge para zero. Ou seja, T_n é consistente se as duas propriedades abaixo são satisfeitas:*

(i) $\lim_{n\to\infty} E(T_n) = \theta$;

(ii) $\lim_{n\to\infty} Var(T_n) = 0$.

Observação: Se o estimador T_n é não viciado para θ e deseja-se verificar sua consistência, basta observar a segunda condição da definição acima. Ou seja, um estimador T_n não viciado é consistente para θ se $\lim_{n\to\infty} Var(T_n) = 0$.

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Erro Quadrático Médio

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas

(Erro Quadrático Médio) *Chama-se erro quadrático médio (EQM) do estimador T ao valor*

$$EQM(T;\theta) = E(e^2) = E(T-\theta)^2.$$

Desenvolvendo , é possível chegar a seguinte fórmula:

$$EQM(T;\theta) = Var(T) + B^2(T),$$

onde $B(T) = E(T) - \theta$ é o viés de T.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Estimador ótimo e inadmissível

- Estimador ótimo: tem EQM menor ou igual ao EQM de qualquer outro estimador, para todos os valores de θ.

- Estimador inadmissível: dados os estimadores T e S do parâmetro θ, S será estimador inadmissível de θ, se EQM(T) < EQM(S) , para todos os valores de θ.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Estimador eficiente

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas

(Eficiência de um estimador) *Dados dois estimadores T_1 e T_2, não viciados para o parâmetro θ, dizemos que T_1 é mais eficiente que T_2 se*

$$Var(T_1) < Var(T_2)$$

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

ESTIMADORES POR MÁXIMA VEROSSIMILHANÇA

O princípio de máxima verossimilhança é um dos procedimentos usados para se obter estimadores. Ele trata o problema de estimação baseado nos resultados obtidos pela amostra e devemos determinar qual a distribuição, dentre todas aquelas definidas pelos possíveis valores de seus parâmetros, com maior possibilidade de ter gerado tal amostra. Consideremos uma população e uma variável aleatória X, relacionada a essa população, com função de probabilidade (se X é uma variável aleatória discreta) ou função densidade de probabilidade (se X é uma variável aleatória contínua) $f(x,\theta)$, sendo θ o parâmetro desconhecido. Seja $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}_{n\geq 1}$ sequência de modelos, com espaço paramétrico $\Theta \in \mathbb{R}^p$. Desta forma, retiramos uma amostra aleatória simples de X, de tamanho n, $X_1, \dots, X_n$, e sejam $x_1, \dots, x_n$ os valores efetivamente observados.

A função de verossimilhança L é definida por

$$L(\theta; x_1, \dots, x_n) = f(x_1;\theta) \times \dots \times f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta).$$

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Função escore

- A função escore é obtida a partir da função de verossimilhança:

$$U(\theta) = \frac{\partial \log L(\theta)}{\partial \theta}$$

- Para encontrar a estimativa de máxima verossimilhança basta igualar a função escore a 0 (zero) e isolar o parâmetro.

Prof. Natália Ribeiro S. Evangelista
natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Informação de Fisher

- A segunda derivada da função de log-verossimilhança tem um papel fundamental na construção de intervalos de confiança por aproximação quadrática da função de log-verossimilhança, uma vez que ela mede a curvatura da função no entorno do ponto de máximo.

$$I_{\mathrm{F}}(\theta) = \mathrm{E}\left[\left(\frac{\partial \log f(X|\theta)}{\partial \theta}\right)^2\right],$$

onde $f(X|\theta)=L(\theta)$

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

INFERÊNCIA ESTATÍSTICA (CONCEITOS INICIAIS)
DESIGUALDADE DA INFORMAÇÃO

Instituto de Educação Superior de Brasília
A Teoria e a Prática Juntas

- Teorema da Desigualdade da Informação: quando as condições de regularidade estão satisfeitas, a variância de qualquer estimador não viciado $\theta^{\wedge}$ do parâmetro θ satisfaz a desigualdade

$$\mathrm{Var}[\hat{\theta}] \geq \frac{1}{nI_{\mathrm{F}}(\theta)}.$$

A desigualdade da informação, inicialmente chamada de Cramér-Rao, não é um método de construção de estimadores, mas apenas possibilita verificar se determinado estimador é ou não eficiente. É então importante que sejam estabelecidos métodos para construção de estimadores que tenham boas propriedades.

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br

Inferência Estatística (Conceitos Iniciais)
Exercício

LISTA DE EXERCÍCIOS 1 – parte I

Data limite de entrega: vide cronograma em “Sistema de avaliação”

Prof. Natália Ribeiro S. Evangelista

natalia.evangelista@iesb.br