

BIOESTATÍSTICA

AULA 5

Anderson Castro Soares de Oliveira
Jose Nilton da Cruz

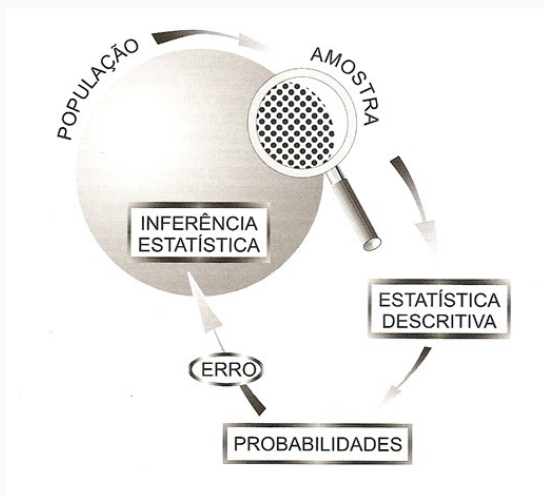
Departamento de Estatística/ICET/UFMT

Teoria de Estimação

INFERÊNCIA ESTATÍSTICA

- Vários tipos de estudos tem o objetivo de obter conclusões (fazer inferências) a respeito de parâmetros de uma população.
- Inferência estatística consiste em tomar decisões acerca de uma população com base nos resultados obtidos na amostra aleatória proveniente desta população.
- A inferência estatística está apoiada em probabilidades

INFERÊNCIA ESTATÍSTICA



TEORIA DA ESTIMAÇÃO

- **Parâmetro:** É uma constante (um número) que caracteriza uma população. Exemplo: média populacional μ , variância populacional σ^2 , etc. Em geral, os parâmetros são desconhecidos.
- **Estimador:** É uma expressão algébrica utilizada para obter um valor aproximado de um parâmetro. Exemplo:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

- **Estimativa:** É o valor numérico de um estimador. É determinada usando os dados amostrais.

- Estimação: conjunto de procedimentos que permitem obter dos dados uma aproximação (bem como uma medida da qualidade da aproximação) para uma quantidade de interesse cujo valor é desconhecido, denominado parâmetro e denotado genericamente por θ

- De uma dada população são extraídas n amostras:
 - Os elementos destas n amostras podem variar, não sendo necessariamente os mesmo de uma amostra para a outra
 - Os parâmetros estimados devem ser diferentes entre as n amostras e aos parâmetros populacionais.
 - Estas diferenças seguem ou não um comportamento regular e previsível?
 - A distribuição de probabilidade de uma estatística é chamada de distribuição amostral.

- Estimação: conjunto de procedimentos que permitem obter dos dados uma aproximação (bem como uma medida da qualidade da aproximação) para uma quantidade de interesse cujo valor é desconhecido, denominado parâmetro e denotado genericamente por θ

Distribuição Amostral

- A partir da distribuição amostral é possível fazer inferências para populações
- As principais distribuições amostrais são:
 - Distribuição Normal - associada a médias e proporções
 - Distribuição t-student - associada a média
 - Distribuição Qui-quadrado - associada a variância
 - Distribuição F-Snedecor - associada a variância

Intervalo Confiança

INTERVALOS DE CONFIANÇA

- Uma maneira de se calcular uma estimativa de um parâmetro desconhecido, é construir um intervalo de confiança $[a, b]$ da seguinte forma:
 - Verifica a distribuição amostral associada ao parâmetro de interesse
 - Estabelece a probabilidade de $1 - \alpha$ (nível de confiança) de que o intervalo contenha o verdadeiro parâmetro
 - Calcula-se os valores $[a, b]$

INTERVALOS DE CONFIANÇA

- O intervalo de confiança pode ser representado da seguinte forma

$$P(a \leq \mu \leq b) = 1 - \alpha$$

- Dessa maneira α será o nível de significância, isto é, o erro que se estará cometendo ao afirmar que o parâmetro está entre o limite inferior e o superior calculado.

Intervalo de Confiança para proporção p

INTERVALO DE CONFIANÇA PARA PROPORÇÃO P

- Consideremos uma população cujos elementos podem ser classificados em dois tipos: Sucesso e Insucesso.
- Pretende-se estimar a proporção p de sucessos na população.
- Dada uma amostra de tamanho n , uma estimativa pontual de p da proporção de sucessos é dada por

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

INTERVALO DE CONFIANÇA PARA PROPORÇÃO P

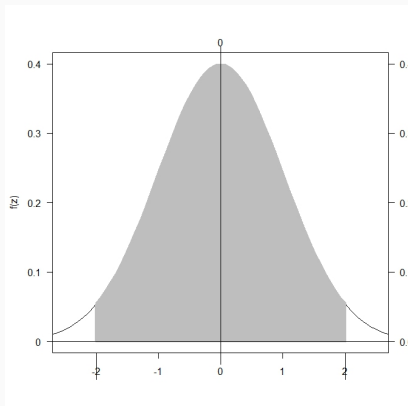
- Fixando uma probabilidade de confiança $(1 - \alpha)$, o intervalo de confiança para uma proporção pode ser obtido da seguinte forma:

$$P(\hat{p} - e \leq p \leq \hat{p} + e) = 1 - \alpha$$

- $e = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$ é a margem de erro da proporção
- $z_{\frac{\alpha}{2}}$ é o valor da curva normal padrão acima do qual encontramos uma área de $\frac{\alpha}{2}$.
- Assim, a distribuição de probabilidade associada a proporção seria a distribuição normal

INTERVALO DE CONFIANÇA PARA PROPORÇÃO P

- Podemos representar graficamente o intervalo de confiança da seguinte forma.



INTERVALO DE CONFIANÇA PARA PROPORÇÃO P

- É possível determinar o tamanho de uma amostra para que esta satisfaça um certo intervalo de confiança
- Fixando uma margem de erro, podemos determinar o tamanho da amostra

$$e = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} \Rightarrow n = \left(z_{\frac{\alpha}{2}}\right)^2 \frac{\hat{p}\hat{q}}{e^2}$$

Exemplo

- Numa certa região foi feita uma amostragem aleatória em propriedades para verificar se os pecuaristas fazem vacinação do rebanho contra a raiva bovina. Num total foram amostradas 220 propriedades, em que 132 delas os rebanhos foram vacinados.
- Assim, a proporção estimada de produtores que vacinam o rebanho é

$$\hat{p} = \frac{x}{n} = \frac{132}{220} = 0,60$$

INTERVALO DE CONFIANÇA PARA PROPORÇÃO P

Exemplo

- Para obter o intervalo de confiança temos que
 - Como $\hat{p} = 0,60$, temos que $\hat{q} = 1 - \hat{p} = 1 - 0,60 = 0,40$
 - Como queremos o intervalo de confiança a 95%, temos que:

$$1 - \alpha = 0,95 \Rightarrow \alpha = 1 - 0,95 = 0,05 \Rightarrow \frac{\alpha}{2} = \frac{0,05}{2} = 0,025$$

- Assim, temos que o valor tabelado de $z_{\frac{\alpha}{2}} = z_{0,025} = 1,96$

INTERVALO DE CONFIANÇA PARA PROPORÇÃO P

Exemplo

- A margem de erro

$$\begin{aligned}e &= z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} \\&= 1,96 \sqrt{\frac{0,60 \times 0,40}{220}} \\&= 0,06\end{aligned}$$

- Assim, o intervalo de confiança

$$P(\hat{p} - e \leq p \leq \hat{p} + e) = 1 - \alpha$$

$$P(0,60 - 0,06 \leq p \leq 0,60 + 0,06) = 0,95$$

$$P(0,54 \leq p \leq 0,66) = 0,95$$

INTERVALO DE CONFIANÇA PARA PROPORÇÃO P

- Quantas propriedades deverão ser amostradas no futuro para se estimar p com um erro de 3% com 95% de confiança?
- Temos

$$n = \left(z_{\frac{\alpha}{2}}\right)^2 \frac{\hat{p}\hat{q}}{e^2} = (1,96)^2 \frac{0,60 \times 0,40}{(0,03)^2} = 1024,43$$

Utilizando o R

- No pacote `RcmdrPlugin.EZR` também é possível obter os intervalos de confiança
- Para obter o IC da proporção utiliza-se:
 - `Statistical Analysis` $\Rightarrow$ `Discrete variables` $\Rightarrow$ `Confidence interval for a proportion`
 - Em seguida deve-se informar o tamanho da amostra n , o número de sucessos X e o coeficiente de confiança

Intervalo de Confiança para média

μ

INTERVALO DE CONFIANÇA PARA MÉDIA μ

- Utilizando a estimativa pontual $\bar{X}$ e a margem de erro, pode-se construir um intervalo de confiança para a média populacional μ

$$P(\bar{X} - e \leq \mu \leq \bar{X} + e) = 1 - \alpha$$

- A margem de erro e pode ser obtida de duas formas

- Para grandes amostras, se $n > 30$

$$e = z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

- $z_{\frac{\alpha}{2}}$ é o valor da curva normal padrão acima do qual encontramos uma área de $\frac{\alpha}{2}$.

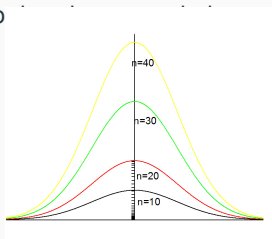
- Para pequenas amostras, se $n \leq 30$

$$e = t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

- $t_{\frac{\alpha}{2}}$ é o valor encontrado na tabela da distribuição t.

INTERVALO DE CONFIANÇA PARA MÉDIA μ

- A distribuição t-Student é bastante semelhante com a distribuição Normal
- Algumas características da distribuição t de student:
 - É simétrica em relação a zero;
 - Todas curvas tem máximo em $t = 0$;
 - Existe uma curva para cada tamanho de amostra (n) e o valor $\nu = n - 1$ (número de graus de liberdade) é usado para obtenção
 - A medida que n aumenta, a curva se aproxima da normal padrão z ;
- Notação: $X \sim t_\nu$



INTERVALO DE CONFIANÇA PARA MÉDIA μ

- Para determinar o tamanho da amostra utilizamos:

$$e = z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \Rightarrow n = \left(z_{\frac{\alpha}{2}} \frac{S}{e} \right)^2$$

INTERVALO DE CONFIANÇA PARA MÉDIA μ

Exemplo

- Deseja-se estimar o diâmetro pupilar de coelhos adultos normais, a partir de uma amostra de 12 animais, em foi encontrado média foi $5,2mm$ e desvio padrão é $1,2mm$. Construir o intervalo de confiança, considerando grau de confiança de 95%.
- Como $n < 30$, primeiramente temos que obter o valor tabelado de t , como queremos o intervalo de confiança a 95%, temos que:

$$1 - \alpha = 0,95 \Rightarrow \alpha = 1 - 0,95 = 0,05 \Rightarrow \frac{\alpha}{2} = \frac{0,05}{2} = 0,025$$

- Olhando o valor tabelado que deixa 0,025 de área acima com $\nu = 12 - 1 = 11$, temos $t_{\frac{\alpha}{2}} = 2,20$

INTERVALO DE CONFIANÇA PARA MÉDIA μ

Exemplo

- Temos que: $\bar{X} = 5,2$, $S = 1,2$

$$\begin{aligned} e &= t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \\ &= 2,2 \frac{1,2}{\sqrt{12}} \\ &= 0,76 \end{aligned}$$

- Assim, o intervalo de confiança

$$P(\bar{X} - e \leq \mu \leq \bar{X} + e) = 1 - \alpha$$

$$P(5,2 - 0,76 \leq \mu \leq 5,2 + 0,76) = 95\%$$

$$P(4,44 \leq \mu \leq 5,96) = 95\%$$

Utilizando o R

- No pacote `RcmdrPlugin.EZR` também é possível obter os intervalos de confiança
- Para obter o IC para média utiliza-se
 - `Statistical Analysis` $\Rightarrow$ `Continuous variables` $\Rightarrow$ `Confidence interval for a mean`
 - Em seguida deve-se informar a média, o desvio padrão, o tamanho da amostra n e o coeficiente de confiança