Capítulo 5- Introdução à Inferência estatística.

(Versão: para o manual a partir de 2016/17)

1.1) Introdução.(222)(Vídeo 39)

Na inferência estatística, analisamos e interpretamos amostras com o objetivo de tirar conclusões acerca da população de onde se extraiu a amostra.

População- é o conjunto de todos os elementos em estudo.

Amostra- é um subconjunto finito da população.

Nota: usamos amostras porque, muitas vezes não é viável utilizarmos toda a população.

Exemplo 1(223)   ***Exemplo 2***   ***Exemplo 3.***

Amostragem Probabilística / não probabilística.

☞***Exercício(266): 1.***

1.1.1- Métodos de amostragem probabilística. (225).

1- amostragem aleatória simples de n elementos. ***Exemplo7*** (225).

2- amostragem aleatória de n elementos com reposição.

3- Amostragem aleatória sistemática. ***Exemplo 8***(226).

4- Amostragem aleatória estratificada. ***Exemplo 9***(226).

5- Amostragem aleatória por conglomerados.(227)

Exercícios(266):3, 4, 5.

1.2- Parâmetro e estimador (ou estatística). Estimativa.(228).

Exemplo 1(228).

População	**N** Dimensão da população	**μ** Valor médio populacional	**p** Proporção populacional	**σ** Desvio padrão populacional.
Amostra	**n** Tamanho da amostra.	$\bar{X}$ $\bar{x}$ Média amostral.	$\hat{p}$ Proporção amostral.	**S** Desvio padrão amostral.

No caso da média, temos:

μ **Parâmetro** (populacional)

$\bar{X}$ **Estimador**(fórmula ou processo para estimar). O estimador é uma variável aleatória, pois os seus valores variam de amostra para amostra.

$\bar{x}$ **Estimativa**- resultado concreto de uma amostra particular.

Exemplo 2(230).

☞***Exercícios(266):6, 7, 8, 9.***

1.3- Distribuição de amostragem de uma estatística (ou estimador). (231)

1.4- Estimação de um valor médio. (232).

Exemplo 1(232).

$$\boldsymbol{E(\bar{X}) = \mu} \qquad \boldsymbol{\frac{\sigma}{\sqrt{n}} = \sigma_{\bar{X}}} \quad \textbf{(235).}$$

Dizemos que $\bar{X}$ é um **estimador centrado** ou cêntrico ou não enviesado, pois o seu valor médio é igual ao parâmetro que pretendemos estimar. $E(\bar{X}) = \mu$.

O desvio padrão da distribuição de amostragem da média:

$$\boldsymbol{\frac{\sigma}{\sqrt{n}} = \sigma_{\bar{X}}}, \qquad \text{também se designa } \textbf{erro padrão.}$$

Exemplo 2 (237).

☞***Exercício(267):10.***

Nota: O aumento da dimensão da amostra, a partir da qual é calculado o estimador, faz diminuir a variabilidade das estimativas. (236).

1.5- Importância do teorema do Limite Central.

Teorema do Limite Central.(241)

Seja X uma população com valor médio μ e desvio padrão σ, da qual se recolhem amostras de dimensão n.

Então, se $n \geq 30$, a distribuição de amostragem da média $\bar{X}$ pode ser aproximada a uma distribuição normal com valor médio μ e desvio padrão $\frac{\sigma}{\sqrt{n}}$,

isto é, para $\mathbf{n \geq 30}$ $\quad \boldsymbol{\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)}$

Nota: este teorema diz-nos que as médias amostrais, apesar de variarem de amostra para amostra, tendem a concentrar-se em torno do valor médio da população à medida que a dimensão das amostras aumenta, uma vez que o desvio-padrão $\frac{\sigma}{\sqrt{n}}$ diminui.

Exemplo 1(241).

☞***Exercício(267):13***.

Exemplo 2.1 (241).

☞***Exercícios(267): 14, 15.1, 15.2).***

☞***Atividade 2.1(244).***

1.6- Intervalos de confiança para o valor médio.(245) (Vídeo 40).

$$\left]\boldsymbol{\bar{X} - Z\frac{\sigma}{\sqrt{n}}; \ \bar{X} + Z\frac{\sigma}{\sqrt{n}}}\right[$$

$\boldsymbol{\bar{X}}$ é a <u>média amostral</u> $\boldsymbol{\sigma}$ desvio padrão

n é o número de elementos da amostra

Z é o valor relacionado com a confiança. Os valores usuais para Z são os que constam da tabela seguinte:

Nível de confiança	90%	95%	99%
Valor de Z	1,645	1,960	2,576

A diferença entre os dois extremos do intervalo: $\left(\bar{X} + Z\frac{\sigma}{\sqrt{n}}\right) - \left(\bar{X} - Z\frac{\sigma}{\sqrt{n}}\right)$ é a **amplitude do intervalo**.

$Z\frac{\sigma}{\sqrt{n}}$ representa a **margem de erro** do intervalo.

Nota: Quando afirmamos que o intervalo tem 95% confiança, significa que, para 100 amostras diferentes, da mesma dimensão, selecionadas da mesma população, esperamos que 95 dos correspondentes intervalos obtidos, contenham o valor médio da população a estudar.

Exemplo 1 (246).

Atividade 1(247) ***Atividade 2(247)***

Exemplo 2(248). ***Atividade 3(248)*** ***Exemplo 3(249).***

Atividade 4(249) ***☞Exercícios(267):16, 17…até 26.***

Nota: Quando aumentamos o grau de confiança, a amplitude do intervalo aumenta, e consequentemente a informação fica menos precisa.

Nota: Quando aumentamos o número de elementos da amostra, a margem de erro $Z\frac{\sigma}{\sqrt{n}}$ diminui e a amplitude do intervalo também diminuir. Assim, melhoramos a precisão da nossa estimação.

1.7- Estimativa pontual de proporção.(250)(Vídeo41).

$$\boldsymbol{\hat{p} = \frac{f}{n}}$$

A **proporção** é o número de elementos favoráveis à nossa condição, a dividir pelo número total de elementos da amostra.

Exemplo 1(250).

Nota: como podemos constatar com o exemplo 1 da página 252 do livro, podemos transformar uma proporção numa média e assim adaptar o resultados anteriores ao caso das proporções.

☞Exercícios(269):27, 28.

Proporção/ Valor médio. Válido para amostras com dimensão maior ou igual a 30 elementos.

Distribuição de amostragem da proporção amostral

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

Exemplo 2(253).

☞Atividade 2(253).

☞Exercícios(269):30, 31.

1.8- Intervalos de confiança para a proporção (254).

$$\left]\hat{p} - Z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\,;\hat{p} + Z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right[$$

$\hat{p}$ Representa a proporção amostral.

$Z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ é a margem de erro.

Exemplo 1 (254)

Exemplo 2(255) Exemplo 3(255)

Nota: À semelhança das médias, também podemos verificar que nas proporções, quanto maior for a dimensão da amostra, menor é a amplitude do intervalo.

☞**Atividade 1(256).**

☞**Atividade 2(256).**

☞**Atividade 3(256).**

☞**Exercícios(269):32, 33, 34, 35.**

1.9- Interpretação do conceito de intervalo de confiança(257).(Vídeo 43)

Exemplo 1(257).

ε- Margem de erro.

Nota. Para uma boa estimação, devemos ter presente:

A Qualidade da amostra, isto é, garantir que esta é representativa da população. Se uma amostra for enviesada, as conclusões que tiramos serão pouco credíveis.

O tamanho da amostra- aumentando o tamanho da amostra, a margem de erro diminui e a precisão aumenta.

O grau de confiança- é importante ter um valor alto para o grau de confiança, no entanto, quando este é demasiado próximo de 100%, pode acontecer que o intervalo tenha uma amplitude muito grande, e seja pouco útil a informação fornecida.

☞Atividade 1(257).

☞Atividade 2(257).

☞Exercícios(270):36 e 37.

Qualidade da amostra (258).

Devemos garantir que a amostra é representativa da população. Se uma amostra for enviesada, as conclusões que tiramos serão pouco credíveis.

Exemplo 2(258).

Grau de confiança (258).

É importante ter um valor alto para o grau de confiança. No entanto, quando este é demasiado próximo de 100%, pode acontecer que o intervalo tenha uma amplitude muito grande, e a informação fornecida seja pouco útil.

Dimensão da amostra (259).

Já vimos que, aumentando o tamanho da amostra, a margem de erro diminui e a precisão aumenta

Exemplo 3(259).

Para estimarmos o tamanho da amostra, para o qual a margem de erro é inferior a um determinado valor, começamos por igualar a expressão da margem de erro ao valor pretendido, e depois resolvemos a equação até obtermos o valor de n.

No final da equação, obteremos uma expressão do tipo: $\boldsymbol{n = \left(\frac{Z.\sigma}{\varepsilon}\right)^2}$ no caso do intervalo da média.

Para a proporção, obtemos algo do tipo:

$$\boldsymbol{n = \left(\frac{Z}{\varepsilon}\right)^2 . \hat{p}. (1 - \hat{p})}$$

Nota: devemos usar sempre os cálculos a partir da margem de erro. Estas fórmulas serão apenas para verificar.

☞Atividade 3(259). ***Exemplo 4.(260)*** ***☞Atividade 4(260).***

☞Exercícios(270):38, 39, 40, 41. ***☞Exercíciosglobais(271):1 a 9.***