8.1-Noções gerais.

(Introdução, amostragem, parâmetro, estimador, estimativa, distribuição de amostragem e Teorema do limite central.(TLC)).

Introdução.

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: A amostra deverá ser representativa da população, caso contrário, não poderemos tirar conclusões fiáveis. Quando uma amostra não é representativa, dizemos que é **enviesada**.

Métodos de amostragem probabilística- Qualquer elemento da população tem alguma probabilidade de fazer parte da amostra.

1- **Amostragem aleatória simples de n elementos**.

2- **Amostragem aleatória de n elementos com reposição**.

3- **Amostragem aleatória sistemática-** criamos uma regra para extrair os números.

4- **Amostragem aleatória estratificada**- Escolhemos a amostra respeitando alguns estratos da população que acreditamos influenciar as respostas ao inquérito.

5- Amostragem aleatória por conglomerados. Quando analisamos grupos de indivíduos que correspondam ao modo como se agrupam naturalmente na população em que estão inseridos.

Parâmetro e estatística. Estimativa pontual

O **parâmetro** é referente à população. A estatítica ou **estimador** ou **estimativa** é referente à amostra.

Símbolos a utilizar:

Dimensão da população **N** / Tamanho da amostra: $\boldsymbol{n}$

Valor médio populacional: $\boldsymbol{\mu}$ / média amostral: $\overline{\boldsymbol{X}}$ ou $\overline{\boldsymbol{x}}$

Proporção populacional: $\boldsymbol{p}$ / Proporção amostral: $\widehat{\boldsymbol{p}}$

Desvio padrão populacional: $\boldsymbol{\sigma}$ / Desvio padrão amostral: ***S***

Nota: por vezes, usamos um acento circunflexo (**^**) para indicar que se refere à amostra.

Um ***Parâmetro,*** $\boldsymbol{\theta}$ carateriza a população. Uma estatística ou estimador $\widehat{\boldsymbol{\theta}}$ carateriza a amostra.

Nota: **Parâmetro, estimador, estimativa.**

No caso da média, temos, como vimos μ **Parâmetro** (populacional)

$\bar{X}$ **<u>Estimador</u>**(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.

.

Estimação do valor médio.

$$\boldsymbol{E(\overline{X}) = \mu}$$

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.

O <u>desvio padrão da distribuição de amostragem da média</u>, isto é, o desvio padrão das médias amostrais pode ser dado por

$$\boldsymbol{\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}}$$ também se designa **<u>erro padrão</u>.**

Teorema do Limite Central.

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 **$n \geq 30$**

$$\boldsymbol{\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)}$$

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.