

Probabilidade e Estatística

Margem de Erro

Determinação do Tamanho da Amostra



Margem de Erro

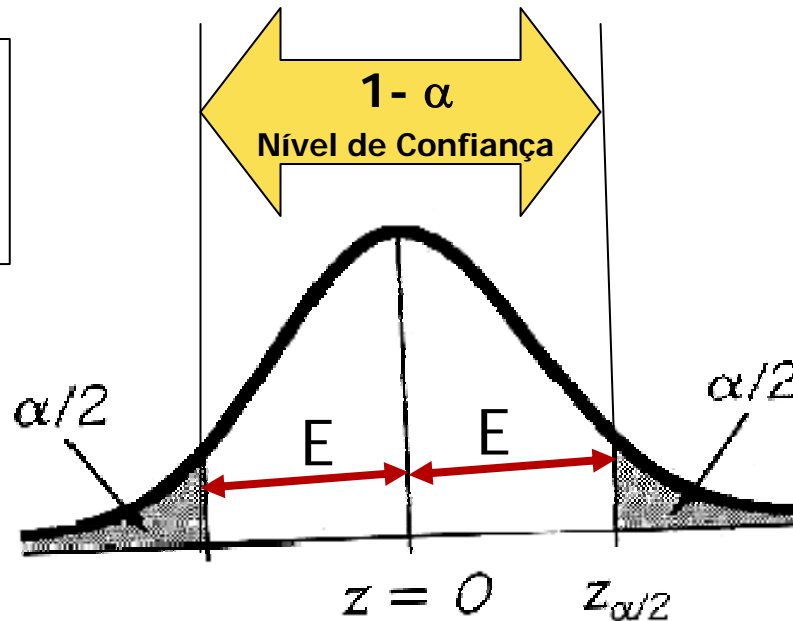
- Quando utilizamos dados amostrais para estimar uma média populacional μ , a margem de erro (E) é a diferença máxima provável (com probabilidade $1-\alpha$) entre a média amostral observada \bar{x} e a verdadeira média da população (μ)

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Margem de Erro

- Ou seja:
 - Há uma probabilidade de $1-\alpha$ de uma média amostral conter um erro não superior a E , e uma probabilidade de α de uma média amostral conter um erro superior a E .

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$



Problema!!!

- Como geralmente não conhecemos o real valor de σ , podemos aplicar as seguintes considerações:
 - $n > 30 \rightarrow$ pode-se adotar para σ o desvio-padrão amostral 's';
 - $n \leq 30 \rightarrow$ a população deve ter distribuição normal e devemos ter σ para aplicar a fórmula

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Para que serve isto?

- Com o conhecimento de E , podemos determinar o intervalo de confiança como:

$$\bar{x} - E < \mu < \bar{x} + E \quad \text{ou}$$

$$\mu = \bar{x} \pm E$$

$$(\bar{x} - E; \bar{x} + E)$$

Onde:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Entenda

$$z = \frac{x - \mu}{\sigma}$$

Como estamos tratando da distribuição das médias amostrais

$$x = \bar{x}$$
$$\sigma = \frac{\sigma}{\sqrt{n}}$$

Portanto:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Assim, para

$$\left\{ \begin{array}{l} -z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \Rightarrow -z \cdot \frac{\sigma}{\sqrt{n}} = \bar{x} - \mu \Rightarrow \mu = \bar{x} + z \cdot \frac{\sigma}{\sqrt{n}} \\ z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \Rightarrow z \cdot \frac{\sigma}{\sqrt{n}} = \bar{x} - \mu \Rightarrow \mu = \bar{x} - z \cdot \frac{\sigma}{\sqrt{n}} \end{array} \right.$$

Exemplo

- Numa pesquisa, foram coletadas 106 amostras de temperatura, obtendo-se uma média (\bar{x}) de 98,20 °F e desvio padrão $s=0,62$ °F. Para um nível de confiança de 95%, determine:
 - (a) A margem de erro da estimativa
 - (b) O Intervalo de confiança para μ

Exemplo

$$NC=95\% \Rightarrow \alpha=0,05 \Rightarrow Z_{\alpha/2} = 1,96$$

$$(a) E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1,96 \cdot \frac{0,62}{\sqrt{106}} = 0,12$$

(b) Como $\bar{x} = 98,20$ e $E = 0,12$;

$$\bar{x} - E < \mu < \bar{x} + E$$

$$98,2 - 0,12 < \mu < 98,2 + 0,12$$

$$98,08 < \mu < 98,32$$

$$\text{OU} \quad \mu = 98,20 \pm 0,12 \\ (98,08;98,32)$$

Se colhermos muitas amostras de tamanho 106 e construirmos um intervalo de confiança com NC=95% para cada um, 95% deles conteriam o valor da média populacional μ .

A temperatura média do ser humano é 98,6 °F

Determinação do Tamanho da Amostra

- Uma das perguntas mais importantes numa análise estatística é determinar qual o melhor tamanho de amostras que devemos ter.
 - Amostras muito grandes são dispendiosas e demandam mais tempo de manipulação e estudo
 - Amostras pequenas são menos precisas e pouco confiáveis

Determinação do Tamanho da Amostra

- Pode-se estimar o melhor tamanho da amostra pela fórmula:

$$n = \left[\frac{z_{\alpha/2} \cdot \sigma}{E} \right]^2$$

Obs.: Aproxime sempre para cima!!

- Observa-se que o tamanho da amostra depende do grau de confiança desejado, da margem de erro pretendida e do σ .

**CONTINUAMOS PRECISANDO DE σ ,
QUE AINDA É DESCONHECIDO**

Determinação do Tamanho da Amostra

- A fórmula exige que se substitua por algum valor o desvio-padrão populacional σ , mas se este for desconhecido, devemos poder utilizar um valor preliminar obtido por processos como:
 - $\sigma \approx \text{amplitude}/4$
 - Realizar um estudo piloto iniciando o processo de amostragem. Com base na primeira coleção de pelo menos 31 valores amostrais selecionados aleatoriamente, calcular o desvio padrão amostral 's' e utilizá-lo em lugar de σ . Este valor pode e deve ser refinado com a obtenção de mais dados amostrais.

Exemplo

- Queremos estimar a renda média no primeiro ano de um profissional. Quantas coletas devemos realizar se queremos 95% de confiança em que a média esteja a menos que R\$1.000,00 da renda média verdadeira da população. Suponha σ conhecido e igual a R\$3.000,00.

Exemplo

$$n = \left[\frac{z_{\alpha/2} \cdot \sigma}{E} \right]^2 = \left[\frac{1,96 \cdot 3000}{1000} \right]^2 = 34,54 \Rightarrow 35 \text{ amostras}$$

Aceitemos agora um $E = 2000$;

$$n = \left[\frac{z_{\alpha/2} \cdot \sigma}{E} \right]^2 = \left[\frac{1,96 \cdot 3000}{2000} \right]^2 = 8,64 \Rightarrow 9 \text{ amostras}$$

Ou seja, dobrando o erro admissível, podemos reduzir em aproximadamente $\frac{1}{4}$ o número de amostras.