

Aula Passada:

- Distribuições Amostrais

- Distribuição Amostral de $\bar{X}$ quando
 X tem distribuição normal $X \sim N(\mu, \sigma^2)$

am. aleatória tamanho n

$$\Rightarrow \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

- Distribuição Amostral de $\bar{X}$ quando

X tem distribuição qualquer $E(X) = \mu$

$$\text{Var}(X) = \sigma^2$$

am. aleatória de tamanho n

$$\Rightarrow \bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right) \text{ para } n \text{ grande}$$

- Exemplos de aplicações dos dois resultados

Aplicação do TLC:

Distribuição da Proporção Amostral

p : proporção de elementos portadores de uma certa característica na população

Deseja fazer o vestibular

Possui carro

E^- favorável a determinado projeto

E^+ portador de determinada doença

p - parâmetro populacional desconhecido

Tomada uma amostra aleatória de tamanho n dessa população sejam

$$Y_i = \begin{cases} 1 & \text{se o } \overset{\text{i-ésimo}}{\text{indivíduo}} \text{ apresenta a característica} \\ 0 & \text{caso contrário} \end{cases}$$

$$i = 1, 2, \dots, n \quad e$$

$$\hat{p} = \frac{\text{n.º de elementos na amostra com a característica}}{n}$$

$\rightarrow$ proporção amostral, estimador de p

$\hat{p}$ pode ser escrito como

$$\hat{p} = \frac{Y_1 + Y_2 + \dots + Y_n}{n} = \bar{Y}$$

Amostra aleatória $\Rightarrow$

$Y_1, Y_2, \dots, Y_n$ são v. a. independentes com distribuição de Bernoulli com

$$E(Y_i) = P(Y_i = 1) = p = P(\text{apresentar a característica})$$

$$\text{Var}(Y_i) = p(1-p)$$

$$E(\hat{p}) = E\left(\frac{Y_1 + Y_2 + \dots + Y_n}{n}\right) = \frac{1}{n} E(Y_1 + Y_2 + \dots + Y_n)$$

$$= \frac{np}{n} = p \quad \Rightarrow \hat{p} \text{ é um estimador não viciado para } p$$

$$\text{Var}(\hat{p}) = \frac{1}{n^2} \text{Var}(Y_1 + Y_2 + \dots + Y_n) = \rightarrow \text{indep. } Y_1, Y_2, \dots, Y_n$$

$$= \frac{1}{n^2} [\text{Var}(Y_1) + \text{Var}(Y_2) + \dots + \text{Var}(Y_n)] = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

$$\Rightarrow \lim_{n \rightarrow \infty} \text{Var}(\hat{p}) = 0 \Rightarrow \hat{p} \text{ é consistente para } p \quad (\text{não viciado})$$

Além disso, pelo TLC, $\hat{p} (= \bar{Y})$ é tal que

$$\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{n \rightarrow \infty} N(0, 1) \quad \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{p(1-p)}}{\sqrt{n}}$$

Exemplo 7.17 pag 242 Magalhães e Lima

Proporção de peças fora de especificação em uma fábrica = $p = 0,4$

(Não é um problema de estimação pois p é conhecido)

Amostra de tamanho 30, qual é a probabilidade da amostra fornecer proporção amostral de peças defeituosas inferior a 0,5?

Supondo amostra casual simples com reposição seja

X - n.º de peças defeituosas no lote

$$X \sim B(30; 0,4) \quad \text{e} \quad \hat{p} = \frac{X}{n} \quad n=30$$

$$\left[\begin{array}{l} \text{Obs: } X = Y_1 + Y_2 + \dots + Y_n \quad Y_i = \begin{cases} 1 & \text{se a } i\text{-ésima peça} \\ & \text{é defeituosa} \\ 0 & \text{caso contrário} \end{cases} \\ \text{Soma de } n \text{ v.a. independentes Bernoulli}(p) \\ \text{tem distribuição } B(n, p) \end{array} \right]$$

$$P(\hat{p} < 0,50) = P\left(\frac{X}{n} < 0,5\right) = P(X < 15) =$$

$$= \sum_{i=0}^{14} P(X=i) = \sum_{i=0}^{14} \binom{30}{i} (0,4)^i (0,6)^{30-i} = 0,8250$$

←
Tabela da Binomial ou R

Cálculo aproximado pelo TLC

$$\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{n \rightarrow \infty} N(0, 1)$$

$$E(\hat{p}) = p = 0,4$$

$$\begin{aligned} \text{Var}(\hat{p}) &= \frac{p(1-p)}{n} \\ &= \frac{0,4 \cdot 0,6}{30} = 0,008 \end{aligned}$$

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

$$\hat{p} \approx N(0,4; 0,008)$$

$$Z \sim N(0, 1)$$

$$P(\hat{p} < 0,5) = P\left(\frac{\hat{p} - 0,4}{\sqrt{0,008}} < \frac{0,5 - 0,4}{\sqrt{0,008}}\right) \approx$$

$$\approx P(Z < 1,12) = 0,8686$$

Dimensionamento da amostra para estimar a proporção

Qual é o tamanho de amostra para que a estimativa pontual $\hat{p}$ não se desvie da proporção populacional p por mais que um erro ϵ , com probabilidade γ ?

$$n? \quad P(|\hat{p} - p| \leq \varepsilon) = \gamma$$

$$= P(-\varepsilon \leq \hat{p} - p \leq \varepsilon)$$

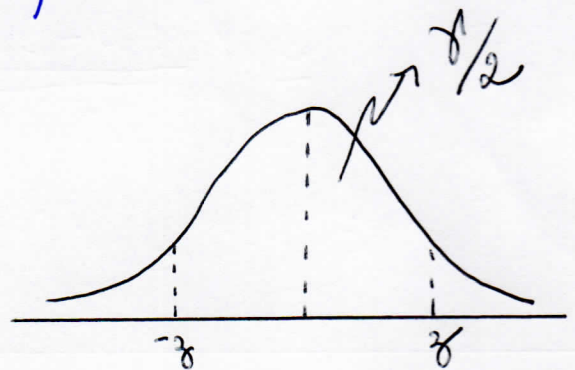
$$\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

$$= P(p - \varepsilon \leq \hat{p} \leq p + \varepsilon) \simeq$$

$$\simeq P\left(\frac{p - \varepsilon - p}{\sqrt{\frac{p(1-p)}{n}}} \leq Z \leq \frac{p + \varepsilon - p}{\sqrt{\frac{p(1-p)}{n}}}\right) \quad Z \sim N(0,1)$$

$$= P\left(\frac{-\varepsilon\sqrt{n}}{\sqrt{p(1-p)}} \leq Z \leq \frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) = \gamma$$

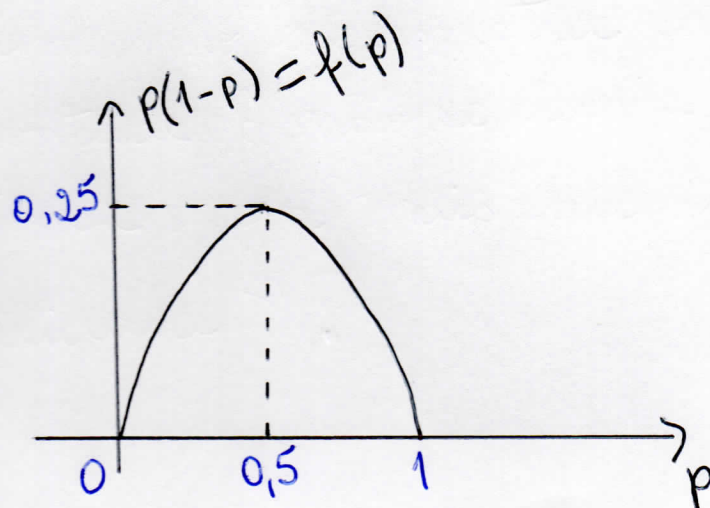
$$\frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}} = z$$



$$n = \underbrace{p(1-p)}_{\text{desconhecido}} \left(\frac{z}{\varepsilon}\right)^2$$

$$z \mid P(-z \leq Z \leq z) = \gamma$$

desconhecido



Substitui-se $p(1-p)$ pelo seu máximo
valor = 0,25

Utiliza-se

$$n = \left(\frac{z}{\varepsilon} \right)^2 \times 0,25$$

Exemplo

Uma cidade do interior de São Paulo deseja estimar a proporção de moradores contaminados com dengue.

a) Determine o tamanho da amostra para que o erro cometido na estimação seja no máximo 0,02 com probabilidade 0,99.

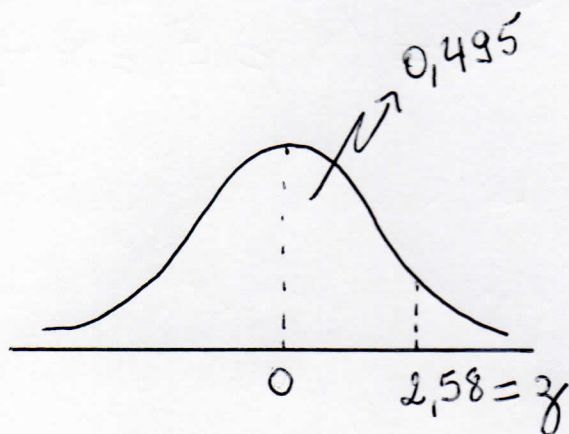
$$P(|\hat{p} - p| \leq 0,02) = 0,99$$

$$\gamma = 0,99 \quad \varepsilon = 0,02$$

$$n = \left(\frac{2,58}{0,02} \right)^2 \times 0,25 = \frac{129^2}{4}$$

→ 129

$$n = 4161$$



b) Se tivermos a informação adicional que a proporção de contaminados é no máximo 0,2, qual deve ser o tamanho da amostra?

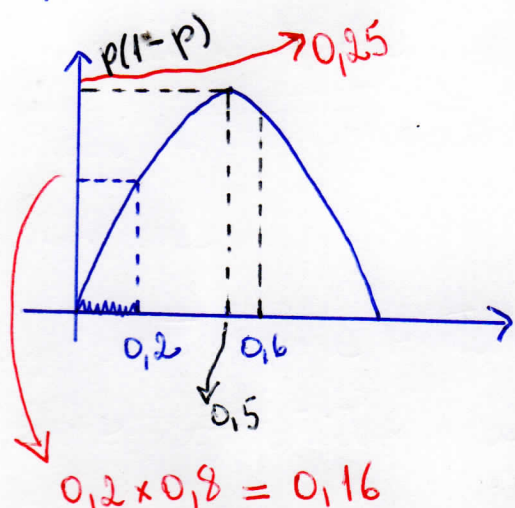
E se a informação é que essa proporção está entre 20% e 30%?

E se a proporção for menor que 60%?

$$p \leq 0,2$$

 $\Rightarrow$

$$p(1-p) \leq 0,16$$



$$n = p(1-p) \left(\frac{z}{\epsilon} \right)^2 = 0,16 \cdot 129^2 = 2663$$

$$\leq 0,16$$

Houve uma considerável redução no tamanho de amostra

$$0,2 \leq p \leq 0,3 \Rightarrow p(1-p) \leq 0,3 \cdot 0,7 \Rightarrow p(1-p) \leq 0,21$$

$$n = 0,21 \cdot 129^2 = 3495$$

$$p \leq 0,6 \Rightarrow p(1-p) \leq 0,25 \text{ (não há redução)} \quad n = 4161$$

Nas pesquisas eleitorais estima-se

p - proporção de eleitores que votarão no candidato A
por

$\hat{p}$ - proporção amostral de eleitores na amostra
que afirmam que votarão no candidato A

Utiliza-se margem de erro = 0,02 $\rightarrow \epsilon$

Coefficiente de confiança = 95% $\rightarrow \gamma$

$$\gamma = 0,95 \Rightarrow P(-z \leq Z \leq z) = 0,95 \Rightarrow z = 1,96$$

$$n = \left(\frac{1,96}{0,02} \right)^2 \cdot \frac{1}{4} = 2401$$

Obs: Nas pesquisas eleitorais não é usada amostra casual simples com reposição mas utiliza-se (erroneamente) a margem de erro da amostragem casual simples com reposição,

O erro da margem de erro - José Ferreira de
Carvalho

MAE -1512 Estatística para a Licenciatura II**2ª. Provinha - 2018**

- 1) Deseja-se estimar a nota média de Matemática dos alunos de uma grande escola. Por experiências anteriores pode-se admitir que o desvio padrão dessas notas é igual a 2, mas acredita-se que a média pode ter se modificado. Determine o tamanho de amostra necessário para que a estimativa da nota média não se afaste do verdadeiro valor por mais de 0,5 unidade com probabilidade 0,90.
- 2) Um procedimento de controle de qualidade foi planejado para garantir um máximo de 10% de itens defeituosos na produção. A cada 6 horas sorteia-se uma amostra de 40 peças e havendo mais de 15% de defeituosas, pára-se a produção para a verificação do processo. Qual é a probabilidade de uma parada desnecessária? Repita o problema para amostra de tamanho 60 e compare com o obtido para amostra de tamanho 40.

Exercício 2 Seção 7.3 pag 243 Magalhães e Lima

MAE 1512- Estatística para a Licenciatura II**2ª. Provinha - 2019**

- 1) O tempo de emissão de extratos do caixa eletrônico de um banco, em segundos, foi modelado por uma distribuição Exponencial com parâmetro $1/40$ (média 40 e variância 40^2). Para uma amostra aleatória de 50 clientes, determine a probabilidade de que o tempo médio de emissão seja inferior a 30 segundos.
Utilize o Teorema do Limite Central.
- 2) Em uma pesquisa de mercado, deseja-se estimar a proporção de pessoas que compram o sabonete B.
 - a) Qual é o tamanho de amostra necessário para que, com probabilidade 0,90, a estimativa não se desvie do verdadeiro valor do parâmetro por mais do que 0,05?
 - b) Se tivermos a informação adicional de que a probabilidade de aceitação do sabonete B é no mínimo 0,8, qual deve ser o tamanho da amostra ?
 - c) Decidiu-se colher uma amostra de 81 pessoas. Qual será o erro máximo que se comete com probabilidade 0,9?
 - d) Para amostra de tamanho 81, qual é a probabilidade de que o erro máximo seja 0,08?