Métodos Estatísticos
Capítulo 4

Resumos e exemplos

Exercícios

Capítulo 4

Amostra aleatória

Uma amostra aleatória de v.a. X é um conjunto finito de n v.a. $x_1, x_2 \dots x_n$, tais que

→ $x_1, x_2 \dots x_n$ são independentes;
→ $x_1, x_2 \dots x_n$ têm a mesma distribuição de X;

Podemos dizer que $x_1, x_2 \dots x_n$ são independentes e identicamente distribuídas com X
iid

Teorema do Limite Central (TLC)

- Se $x_1, x_2 \dots x_n$ são variáveis iid de uma v.a. X;
- Se $n > 30$;
- Se tem uma média μ e desvio-padrão σ;

Então:
$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} \overset{\cdot}{\sim} N(0,1)$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i; \quad E(x) = \mu; \quad V(\bar{X}) = \frac{V(X)}{n} = \frac{\sigma^2}{n}$$
Regras gerais

- Se não estiver centrada e reduzida e seguir uma distribuição normal:

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \Rightarrow \underbrace{\sqrt{n}\,\frac{\bar{X} - \mu}{\sigma} \sim N(0,1)}_{\text{Usar para determinar quantidade média}}$$

→ **Decompor uma TLC:**

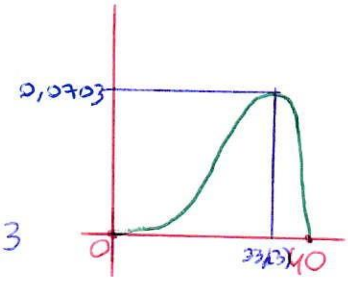
$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} \overset{\cdot}{\sim} N(0,1) = \frac{\times n\left(\frac{1}{n}\sum_{i=1}^{n} x_i - \mu\right)}{\times n\left(\frac{\sigma}{\sqrt{n}}\right)} \overset{\cdot}{\sim} N(0,1) = \frac{\sum_{i=1}^{n} x_i - n\mu}{\sqrt{n}\cdot\sigma} \overset{\cdot}{\sim} N(0,1)$$

multiplicar por $\boxed{n}$ para remover

$$= \underbrace{\sum_{i=1}^{n} x_i \overset{\cdot}{\sim} N(n\mu, \sqrt{n}\cdot\sigma)}_{\text{Usar para determinar quantidade total}}$$

1) X = "Quantidade de chuva que cai por dia, em l/m^2";

$$f(x) = \begin{cases} \frac{21}{8192 \cdot 10^7} \cdot (40x^5 - x^6), & \text{se } 0 \leq x \leq 40 \\ 0 & \text{, se c.c.} \end{cases}$$



$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\, dx = \int_0^{40} x \cdot f(x)\, dx = 30$$

$$V(X) = E(X^2) - E(X)^2 = \int_0^{40} x^2 \cdot f(x)\, dx - \int_0^{40} x \cdot f(x) dx^2 = 33,33$$

Seja $X_1, X_2 \ldots X_{100}$ uma amostra aleatória de X

X_i = "Quantidade de chuva que cai no dia i, em l/m^2";

a) $X_1, X_2 \ldots X_{100}$ têm que ser variáveis aleatórias iid com X

b) Considere $\overline{X}_{100} = \frac{1}{100} \sum_{i=1}^{100} X_i = \frac{X_1 + X_2 + \ldots + X_{100}}{100}$

i) $E(\overline{X}_{100}) = E\left(\frac{1}{100} \sum_{i=1}^{100} X_i\right) = \underbrace{\frac{1}{100} \sum_{i=1}^{100}}_{100 \cdot} \underbrace{E(X_i)}_{E(X)=30} = \frac{1}{100} \cdot 100 \cdot 30 = 30$

$V(\overline{X}_{100}) = V\left(\frac{1}{100} \sum_{i=1}^{100} X_i\right) = \left(\frac{1}{100}\right)^2 \underbrace{\sum_{i=1}^{100}}_{100 \cdot} \underbrace{V(X_i)}_{V(X)=33,33} = \frac{1}{100^2} \cdot 100 \cdot 33,33 = 0,3333$

ii) $\frac{\overline{X} - E(\overline{X}_{100})}{\sqrt{V(\overline{X}_{100})}} \dot{\sim} N(0,1) = \frac{\overline{X}_{100} - 30}{\sqrt{0,3333}} = \frac{\overline{X}_{100} - 30}{0,577} \dot{\sim} N(0,1)$

ou $\overline{X}_{100} \dot{\sim} N\left(30, \underbrace{\boxed{\frac{\sqrt{33,33}}{\sqrt{100}}}}_{=0,577}\right) \Rightarrow \frac{\overline{X}_{100} - 30}{0,577} \dot{\sim} N(0,1)$

iii) Se $\frac{\overline{X}_{100} - 30}{0,577} \dot{\sim} N(0,1)$ então $\overline{X} \sim N(30; 0,577)$

$$P(28,5 < \overline{X} < 31,5) = P\left(\frac{28,5 - 30}{0,577} < \frac{\overline{X}_{100} - 30}{0,577} < \frac{31,5 - 30}{0,577}\right) = P(-2,5997 < Z < 2,5997)$$

$$\simeq P(Z < 2,6) - P(-2,6 < Z) = P(Z < 2,6) - 1 + P(Z < 2,6)$$

$$\simeq 0,9953 - 1 + 0,9953 = 0,9906$$

Z	0,00
2,6	0,9953

R: Em aproximadamente 99% dos 100 dias observados, a quantidade média de chuva, em l/m^2, situou-se entre $28,5\ l/m^2$ e $31,5\ l/m^2$

2) $X =$ "Energia, em J, de uma partícula"

$X \sim \mathcal{E}(2) \longrightarrow E(X) = \frac{1}{2}$ e $V(X) = \frac{1}{2^2} = \frac{1}{4}$ } (Ver Tabela distribuições 1.1)

Lei da energia do sistema $\beta = \sum_{i=1}^{1600} X_i$

a) Como $X_1, X_2 \dots X_{1600}$ são iid com média $\frac{1}{2}$ e desvio-padrão $\sqrt{\frac{1}{4}} = \frac{1}{2}$, então podemos aplicar TLC:

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} \dot{\sim} N(0,1) \left\{ \begin{array}{l} E(\bar{X}) = E(X) = \frac{1}{2} \\ V(\bar{X}) = \frac{V(X)}{1600} = \frac{\frac{1}{4}}{1600} = \frac{1}{6400} \end{array} \right.$$

$$= \frac{\frac{1}{1600}\sum_{i=1}^{1600} X_i - \frac{1}{2}}{\sqrt{\frac{1}{6400}}} \dot{\sim} N(0,1) =_{\times 1600} \frac{\frac{1}{1600}\sum_{i=1}^{1600} X_i - \frac{1}{2}}{\frac{1}{80}} \dot{\sim} N(0,1)$$

$\times 1600$

$$= \frac{\sum_{i=1}^{1600} X_i - 800}{20} \dot{\sim} N(0,1) = \sum_{i=1}^{1600} X_i \dot{\sim} N(800, 20)$$

b) $P(780 < \sum_{i=1}^{1600} X_i < 840) = P\left(\frac{780-800}{20} < \frac{\sum_{i=1}^{1600} X_i - 800}{20} < \frac{840-800}{20}\right)$

$= P(-1 < Z < 2) = P(Z<2) - P(-1<Z) = P(Z<2) - 1 + P(Z<1)$

$= 0{,}9772 - 1 + 0{,}8413 = 0{,}8185$

3) $X =$ "Erro da medição do raio de um círculo, em mm";

$X \sim N(0, 5)$; $X_1, X_2 \dots X_{10}$ uma amostra aleatória de X, são independentes entre si e têm a mesma distribuição de X;

$$T_1 = \frac{1}{10}\sum_{i=1}^{10} X_i = \frac{X_1 + X_2 + \dots + X_{10}}{10} \quad \text{média amostral } \bar{X}_{10}$$

$$T_2 = \frac{5X_1 + 5X_{10}}{10} = \frac{5}{10}X_1 + \frac{5}{10}X_{10} = \frac{1}{2}X_1 + \frac{1}{2}X_{10} \quad \text{média de } X_1 \text{ e } X_{10}$$

a) $T_1 = \frac{1}{10}(X_1 + X_2 \dots + X_{10})$, com $X_1, X_2 \dots X_{10}$ a serem iid, pela estabilidade da Normal, T_1 também terá distribuição Normal.

$T_1 \sim N(\mu, \sigma)$

$$E(T_1) = E\left(\frac{X_1 + X_2 + \dots X_{10}}{10}\right) = \frac{1}{10}E(X_1 + \dots + X_{10}) = \frac{1}{10}E(\underbrace{E(X_1)}_{=E(X)} + \dots + \underbrace{E(X_{10})}_{=E(X)})$$

$$= \frac{1}{10}E(E(X) + \dots + E(X)) = \frac{1}{10} \cdot 10 \cdot E(X) = E(X) = 0$$

$$V(T_1) = V\left(\frac{X_1 + \dots + X_{10}}{10}\right) = \frac{1}{10^2}V(X_1 + \dots + X_{10}) \xrightarrow{\text{pq são independentes}} = \frac{1}{10^2}(V(X_1) + \dots + V(X_{10}))$$

$$= \frac{1}{10^2}(V(X) + \dots + V(X)) = \frac{1}{10^2} \cdot 10 \cdot V(X) = \frac{V(X)}{10}$$

$$\sigma = V(T_1) = \sqrt{\frac{V(X)}{10}} = \frac{\sqrt{V(X)}}{\sqrt{10}} = \frac{\sigma_{T_1}}{\sqrt{10}} = \frac{5}{\sqrt{10}}$$

$$T_1 \sim N\left(0, \frac{5}{\sqrt{10}}\right)$$

b) $P(T_1 > 3) = P\left(\frac{T_1 - 0}{5/\sqrt{10}} > \frac{3 - 0}{5/\sqrt{10}}\right) = P(Z > 1{,}9) = 1 - F_Z(1{,}9) = 1 - 0{,}9713$

$= 0{,}0287 \simeq 3\%$

Exe 4) X é uma média (v.a.r.)

$$E(X) = \frac{2}{3} \ ; \quad V(X) = \frac{8}{9}$$

$$X_1, X_2 \dots X_m \text{ iid} \to \bar{X} = \frac{1}{m} \sum_{i=1}^{m} X_i$$

a) Como $X_1, X_2 \dots X_m$ são uma amostra de X, e iid, pelo TLC:

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} \dot{\sim} N(0,1) \text{ em que } \begin{cases} E(\bar{X}) = E(X) = \frac{2}{3} \\ V(\bar{X}) = \frac{V(X)}{m} = \frac{8/9}{m} = \frac{8}{9m} \end{cases}$$

$$\Leftrightarrow \frac{\bar{X} - \frac{2}{3}}{\sqrt{\frac{8}{9m}}} \dot{\sim} N(0,1) \Leftrightarrow \bar{X} \dot{\sim} N\left(\frac{2}{3}, \sqrt{\frac{8}{9m}}\right) = N\left(\frac{2}{3}, \frac{2\sqrt{2}}{3\sqrt{m}}\right)$$

b) $? m : P(\bar{X}_m > 0{,}8) \leq 0{,}0119$

$$\Leftrightarrow P\left(\frac{\bar{X} - \frac{2}{3}}{\frac{2\sqrt{2}}{3\sqrt{m}}} > \frac{0{,}8 - \frac{2}{3}}{\frac{2\sqrt{2}}{3\sqrt{m}}}\right) \leq 0{,}0119$$

$$\Leftrightarrow P\left(Z \leq \frac{0{,}8 - \frac{2}{3}}{2\sqrt{2}/3\sqrt{m}}\right) > 1 - 0{,}0119$$

$$\Leftrightarrow P\left(Z \leq \frac{0{,}8 - \frac{2}{3}}{\frac{2\sqrt{2}}{3\sqrt{m}}}\right) > \underline{0{,}9881}$$

procurar na Tabela

Inversa Normal F_Z^{-1}

z	0,06
2,2	0,9881

$$\Leftrightarrow \frac{0{,}8 - \frac{2}{3}}{\frac{2\sqrt{2}}{3\sqrt{m}}} > 2{,}26$$

$$\Leftrightarrow 3\sqrt{m}\,\frac{0{,}8 - \frac{2}{3}}{2\sqrt{2}} > 2{,}26$$

$$\Leftrightarrow \sqrt{m} > \frac{2{,}26 \cdot 2\sqrt{2}}{3\left(0{,}8 - \frac{2}{3}\right)}$$

$$\Leftrightarrow m > 15{,}98^2$$

$$\Leftrightarrow m > 255{,}38$$

$$\Leftrightarrow m \geq 256$$

5) X = "Concentração diária de um poluente, em nanogramas por m^3"

$X \sim N(100, 9{,}71)$

a) $P(X>120) = 1 - P(X \leq 120)$

$= 1 - P\left(\frac{X-100}{9{,}71} \leq \frac{120-100}{9{,}71}\right)$

$= 1 - P(Z \leq 2{,}06)$

$F_Z(2{,}06)$

z	$0{,}06$
$2{,}0$	$0{,}9803$

$= 1 - 0{,}9803$

$= 0{,}0197$

$= 1{,}97\%$

b) Se X é a concentração diária, então necessitamos de $\bar{X}_{15}$:

$\bar{X}_{15} \sim N(E(\bar{X}_{15}), \sigma(\bar{X}_{15})) \Rightarrow Z = \frac{\bar{X}_{15} - E(\bar{X}_{15})}{\sigma(\bar{X}_{15})} \Rightarrow \ldots$

Conversão direta para centrada e reduzida $\ldots \Rightarrow \sqrt{15}\,\frac{\bar{X}_{15} - 100}{9{,}71} \sim N(0,1)$

$P(\bar{X}_{15} > 120) = 1 - P(\bar{X}_{15} < 120)$

$= 1 - P\left(Z \leq \sqrt{15}\,\frac{120-100}{9{,}71}\right)$

$= 1 - P(Z \leq 7{,}98)$

$\simeq 1 - 1$

$\simeq 0$

z	$0{,}09$
4	$0{,}99978$

Se $4{,}09$ tem $0{,}99978$, e $7{,}98 \gg 4{,}09$, podemos assumir que $F_Z(7{,}98)$ terá um valor extremamente próximo de 1 mas $\neq 1$