Cap. 5 Combinações lineares de variáveis aleatórias e teorema do limite central

5.1 Combinações lineares de v.a.'s

- Dadas k variáveis aleatórias $X_1, X_2, \dots, X_k$, e k constantes $c_1, c_2, \dots, c_k \in \mathbb{R}$, diz-se que a v.a. Y definida por

$$Y = c_1 X_1 + c_2 X_2 + \dots + c_k X_k = \sum_{i=1}^{k} c_i X_i$$

é uma combinação linear de $X_1, \dots, X_k$.

- $$\boxed{E(Y) = E\left(\sum_{i=1}^{k} c_i X_i\right) = \sum_{i=1}^{k} c_i E(X_i)}$$

- $Y = \sum_{i=1}^{k} c_i X_i$, $W = \sum_{j=1}^{m} d_j Z_j$

$$\Rightarrow \boxed{\operatorname{cov}(Y, W)} = E(YW) - E(Y)E(W) =$$

$$= E\left(\left(\sum_{i=1}^{k} c_i X_i\right)\left(\sum_{j=1}^{m} d_j Z_j\right)\right) - E\left(\sum_{i=1}^{k} c_i X_i\right) E\left(\sum_{j=1}^{m} d_j Z_j\right)$$

$$= E\left(\sum_{i=1}^{k} \sum_{j=1}^{m} c_i d_j X_i Z_j\right) - \left(\sum_{i=1}^{k} c_i E(X_i)\right)\left(\sum_{j=1}^{m} d_j E(Z_j)\right)$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{m} c_i d_j \left(E(X_i Z_j) - E(X_i) E(Z_j)\right)$$

$$\boxed{= \sum_{i=1}^{k} \sum_{j=1}^{m} c_i d_j \operatorname{cov}(X_i, Z_j)}$$

- $V(Y) = cov(Y,Y) = \sum_{i=1}^{k} \sum_{j=1}^{k} c_i c_j \, cov(X_i, X_j)$

$$\Rightarrow \boxed{V(Y) = \sum_{i=1}^{k} (c_i)^2 V(X_i) + 2 \sum_{i=1}^{k} \sum_{j>i} c_i c_j \, cov(X_i, X_j)}$$

Casos particulares: $V(X_1 + X_2) = V(X_1) + V(X_2) + 2\,cov(X_1, X_2)$

$V(X_1 - X_2) = V(X_1) + V(X_2) - 2\,cov(X_1, X_2)$

- $\boxed{cov(X_i, X_j) = 0, \ \forall i \neq j \ \Rightarrow V(Y) = \sum_{i=1}^{k} (c_i)^2 V(X_i)}$

Isto acontece em particular quandos as v.a.'s X_i forem independentes duas a duas.

- <u>Casos especiais discretos</u>:

① $X_i \sim Bin(n_i, p)$, independentes, $i = 1, \ldots, k$

$\Rightarrow \ Y = \sum_{i=1}^{k} X_i \sim Bin\left(\sum_{i=1}^{k} n_i, p\right)$.

[$X_i \sim Ber(p) \equiv Bin(1, p)$, indep. $\Rightarrow X_1 + \cdots + X_n \sim Bin(n, p)$]

② $X_i \sim Poisson(\lambda_i)$, independentes, $i = 1, \ldots, k$

$\Rightarrow Y = \sum_{i=1}^{k} X_i \sim Poisson\left(\sum_{i=1}^{k} \lambda_i\right)$.

- Exemplos

(1) $X \sim Bin(5, 0.7)$, $Y \sim Bin(7, 0.7)$, independentes

$P(X+Y=3) = ?$

$Z = X + Y \sim Bin(12, 0.7)$ e $P(Z=3) = F_Z(3) - F_Z(2)$

Para usar tabela, consideramos

$W = 12 - Z = 12 - (X+Y) \sim Bin(12, 0.3)$

$\Rightarrow P(Z=3) = P(W=9) = F_W(9) - F_W(8)$

$= 0.9998 - 0.9983 = 0.0015$

(2) X = nº de pessoas que chegam por mês a uma urgência hospitalar c/ determinada doença rara

$X \sim Poisson(\lambda = 1)$, independente de mês p/ mês.

Y = nº de pessoas que chegam por ano à mesma urgência com essa doença rara.

$P(10 < Y < 20) = ?$

$Y = \sum_{i=1}^{12} X_i$ e $Y \sim Poisson(12)$

$\Rightarrow P(10 < Y < 20) = F_Y(19) - F_Y(10) =$

$= 0.9787 - 0.3472 = 0.6315$

- Caso especial contínuo

$X_i \sim N(\mu_i, \sigma_i^2)$, independentes, $i = 1, \dots, k$

$$\Rightarrow Y = \sum_{i=1}^{k} c_i X_i \sim N\left(\sum_{i=1}^{k} c_i \mu_i, \sum_{i=1}^{k} c_i^2 \sigma_i^2\right)$$

Exemplo: um elevador de acesso a uma mina tem capacidade p/ 3800 Kg. Peso de cada mineiro $\sim N(75, 8^2)$. Qual é a prob. de 50 mineiros excederem a capacidade do elevador?

$X_i \sim N(75, 8^2)$, independentes e identicam/ distribuídas

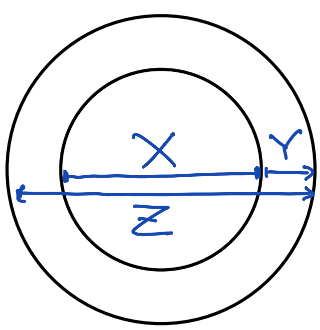
$$\Rightarrow Y = \sum_{i=1}^{50} X_i \sim N(50 \times 75, 50 \times 8^2) = N(3750, 3200)$$

e $P(Y > 3800) = 1 - P(Y \leq 3800)$

$$= 1 - P\left(\frac{Y - 3750}{\sqrt{3200}} < \frac{3800 - 3750}{\sqrt{3200}}\right)$$

$$= 1 - \Phi\left(\frac{50}{10 \times 4 \times \sqrt{2}}\right) \simeq 1 - \Phi(0.88) \simeq 1 - 0.8106 = 0.1894.$$

- Exercício 4.3



[cm]

$X \sim N(3, (0.02)^2)$

$Y \sim N(0.3, (0.005)^2)$

$X \perp\!\!\!\perp Y$

$$\Rightarrow Z = X + 2Y \sim N\left(3.6,\ \underbrace{(0.02)^2 + 4(0.005)^2}_{= 0.0005}\right)$$

$$\Rightarrow \sigma_Z = \sqrt{5 \times 10^{-4}} = 10^{-2} \times \sqrt{5} = 2{,}236/100 .$$

$$P(Z > 3.62) = 1 - P(Z \leq 3.62)$$

$$= 1 - P\left(\frac{Z - 3.6}{0.0236} \leq \frac{0.02}{0.0236}\right)$$

$$\simeq 1 - \Phi(0.89) = 1 - 0.8133 = \underline{0.1867}$$

5.2 Teorema do Limite Central

Teorema: Seja $(X_1, \dots, X_n, \dots)$ uma sucessão de v.a.'s independentes e identicamente distribuídas (i.i.d.), com valor esperado μ e variância σ^2. Considere-se

$$S_n = \sum_{i=1}^{n} X_i \quad \text{e} \quad \bar{X}_n = \frac{S_n}{n}, \quad \forall n \in \mathbb{N}$$

Então

$$\lim_{n \to +\infty} P\left(\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \leq z\right) = \Phi(z)$$

e

$$\lim_{n \to +\infty} P\left(\frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{V(\bar{X}_n)}} \leq z\right) = \Phi(z).$$

Notas:

- $$E(S_n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) \overset{\text{id. dist.}}{=} n\mu$$

$$V(S_n) = V\left(\sum_{i=1}^{n} X_i\right) \underset{\text{ind.}}{=} \sum_{i=1}^{n} V(X_i) \underset{\text{id. dist.}}{=} n\sigma^2$$

ind. + id. dist.

$$\underset{\text{TLC}}{\Longrightarrow} \quad \boxed{\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \overset{a}{\sim} N(0,1)}$$

- $$E(\bar{X}_n) = E\left(\sum_{i=1}^{n} X_i / n\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) \overset{\text{id. dist.}}{=} \mu$$

$$V(\bar{X}_n) = V\left(\sum_{i=1}^{n} X_i / n\right) \underset{\text{ind.}}{=} \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) \underset{\text{id. dist.}}{=} \frac{\sigma^2}{n}$$

ind. + id. dist.

$$\underset{\text{TLC}}{\Longrightarrow} \quad \boxed{\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \overset{a}{\sim} N(0,1)}$$

- <u>Observações</u>:

(1) TLC é para ser usado quando as v.a.'s i.i.d. X_i não têm distribuição normal.

(2) As v.a.'s X_i podem ser discretas ou contínuas.

(3) A demonstração do TLC está fora do âmbito desta cadeira.

- <u>Próxima aula</u>: mais observações, exemplos e exercícios.