$\bar{X}$ e S^2 são independentes - Aplicação do Teorema 2

$$\bar{X} = \frac{\underset{\sim}{1}'\underset{\sim}{X}}{n} \qquad S^2 = \frac{1}{n-1}\underset{\sim}{X}'\left[I - \frac{\underset{\sim}{1}\underset{\sim}{1}'}{n}\right]\underset{\sim}{X}$$

$$A = \frac{1}{n-1}\left[I - \frac{\underset{\sim}{1}\underset{\sim}{1}'}{n}\right.$$

$$B = \frac{\underset{\sim}{1}'}{n} \quad 1\times n$$

$\underset{\sim}{X}'A\underset{\sim}{X}$ e $B\underset{\sim}{X}$ são independentes $\Leftrightarrow$ $\underset{1\times n}{B}\,\underset{n\times n}{\Sigma}\,\underset{n\times n}{A} = \underset{1\times n}{0}$

$$B\Sigma A = \frac{\underset{\sim}{1}'}{n}\,\sigma^2 I\,\frac{1}{n-1}\left[I - \frac{\underset{\sim}{1}\underset{\sim}{1}'}{n}\right] =$$

$$= \frac{\sigma^2}{n(n-1)}\left[\underset{\sim}{1}' - \frac{\overbrace{\underset{\sim}{1}'\underset{\sim}{1}}^{n}\underset{\sim}{1}'}{n}\right] =$$

$$= \frac{\sigma^2}{n(n-1)}\left[\underset{\sim}{1}' - \underset{\sim}{1}'\right] = \underset{\sim}{0}_{1\times n}$$

Como gerar um vetor aleatório $\underset{\sim}{X}$ tal que

$\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$

Gerar $Z_1, Z_2 \dots Z_p$ independentes, $Z_i \sim N(0,1)$

$\underset{\sim}{Z} = (Z_1, Z_2 \dots Z_p) \sim N_p(\underset{\sim}{0}, I_p)$ $\qquad i = 1, 2, \dots p$

Σ positiva definida $\Rightarrow \Sigma = LL'$.

Determino L.

$$\underset{\sim}{X} = \underset{\sim}{\mu} + L\underset{\sim}{Z}$$

$$\underset{\sim}{X} \sim N_p$$

$$E(\underset{\sim}{X}) = \underset{\sim}{\mu} + L\,E(\underset{\sim}{Z}) = \underset{\sim}{\mu} + L\underset{\sim}{0} = \underset{\sim}{\mu}$$

$$Var(\underset{\sim}{X}) = Var(\underset{\sim}{\mu} + L\underset{\sim}{Z}) = Var(L\underset{\sim}{Z}) =$$

$$= L\,Var(\underset{\sim}{Z})\,L' = L I L' = LL' = \Sigma.$$

$$\therefore \underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma).$$

COMO GERAR DADOS DE DISTRIBUIÇÕES NORMAIS P-VARIADAS

__Aplicativo:__ *RGui*

Para gerar dados de distribuições normais p-variadas através do aplicativo RGui é necessário, em primeiro lugar, carregar o pacote "MASS". Com esse pacote carregado, podemos utilizar a função "mvrnorm" que gera dados de distribuições Normais p-variadas. Essa função, que produz uma ou mais observações da distribuição normal multivariada especificada, é dada por:

$$mvrnorm(n,\ mu,\ Sigma,\ tol = 1e^{-6},\ empirical = \text{FALSE})$$

em que:

__*n*__ representa o número de observações que serão geradas. Se $n = 1$, um vetor do mesmo comprimento de *mu* é gerado. Se $n > 1$, uma matriz $n \times p$ é gerada, onde p é o comprimento de *mu*;

__*mu*__ representa o vetor de média das variáveis;

__*Sigma*__ representa a matriz simétrica positiva-definida, especificando a matriz de covariância das variáveis.

__*Tol*__ é a tolerância (relativa à maior variância) caso a matriz *Sigma* não seja positiva-definida.

__*Empirical*__ é uma função lógica (verdadeiro ou falso). Se *empirical* é verdadeiro, então *mu* e *Sigma* especificam, respectivamente, a média populacional e a matriz de covariância não empíricas. Caso contrário, *mu* e *Sigma* especificam, respectivamente, a média populacional e a matriz de covariância empíricas.

__Referências:__ B. D. Ripley (1987) Stochastic Simulation. Wiley. Page 98.

Exemplo: Os vetores aleatórios $X' = [X_1\ X_2\ X_3]$ e $Y' = [Y_1\ Y_2\ Y_3]$ são distribuídos independentemente com distribuição normal trivariada com parâmetros, respectivamente,

$$\mu_1 = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}, \quad \Sigma_1 = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}, \quad \Sigma_2 = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 4 \end{bmatrix}.$$

Considerando esses vetores, gerar 10 dados de suas distribuições e também da distribuição da nova variável $D = [X - Y]$.

Resolução:

```
mu_x <- c(2,2,2)
sigma_x<- matrix(c(3,2,1,2,4,1,1,1,2),3,3, byrow=T)          # Sem o comando byrow=T, os valores são colocados em coluna.

mu_y <- c(3,4,2)
sigma_y<- matrix(c(4,2,0,2,4,2,0,2,4),3,3, byrow=T)

XT<- mvrnorm(10, mu_x, sigma_x, empirical = TRUE)
XF <- mvrnorm(10, mu_x, sigma_x)          #Neste caso o default do R assume empirical = FALSE
```

```
Var_XT<- var(XT),          media_XT<- apply(XT, 2, mean)
#Sem o comando apply, mas somente com o comando mean(XT), obtemos a media do vetor mu
Var_XF<- var(XF),          media_XF<- apply(XF, 2, mean)
```

```
Var_XT
     [,1] [,2] [,3]
[1,]    3    2    1
[2,]    2    4    1
[3,]    1    1    2

media_XT
[1] 2 2 2

Var_XF
         [,1]      [,2]      [,3]
[1,] 2.098150 1.2584740 1.6846992
[2,] 1.258474 3.1000440 0.8563819
[3,] 1.684699 0.8563819 2.6581171

media_XF
[1] 2.637510 1.673638 2.099733
```

```
YT<- mvrnorm(10, mu_y, sigma_y, empirical = TRUE)
YF <- mvrnorm(10, mu_y, sigma_y)

Var_YT<- var(YT),          media_YT<- apply(YT, 2, mean)
Var_YF<- var(YF);          media_YF<- apply(YF, 2, mean)
```

```
Var_YT
             [,1] [,2]         [,3]
[1,] 4.000000e+00    2 1.268276e-15
[2,] 2.000000e+00    4 2.000000e+00
[3,] 1.268276e-15    2 4.000000e+00

media_YT
[1] 3 4 2

Var_YF
           [,1]      [,2]      [,3]
[1,]  3.4359223 0.9189359 -1.867392
[2,]  0.9189359 5.3945125  3.008983
[3,] -1.8673915 3.0089830  4.867940

media_YF
[1] 2.181026 2.853329 1.345639
```

```
DT<- XT - YT;          DF<- XF - YF
```

Referência para
Distribuição Normal
Multivariada.
Tong, Y. L. (1990)
The Multivariate Normal
Distribution.
Springer-Verlag
271 p.