Universidade de Brasília

DEPARTAMENTO DE ESTATÍSTICA

23 junho 2023

Lista 7 - Normal Multivariada

Prof. Dr. George von Borries

Análise Multivariada 1

Aluno: Bruno Gondim Toledo | Matrícula: 15/0167636

Questão 55

```r
mu <- c(3, 2)
sigma <- matrix(c(1, -1.5, -1.5, 4),2)
set.seed(150167636)
mvrnorm(20, mu, sigma)
```

```
##           [,1]       [,2]
##  [1,] 2.713068  2.2421326
##  [2,] 3.168527  1.8305926
##  [3,] 3.281616  3.6377464
##  [4,] 1.225270  3.8406363
##  [5,] 2.788789  1.6138664
##  [6,] 2.971249  0.3776516
##  [7,] 3.139249 -2.2801139
##  [8,] 2.729603  0.9393494
##  [9,] 2.447450  2.9861925
## [10,] 3.184863 -0.9861318
## [11,] 1.798995  3.6115228
## [12,] 1.712520  4.5234829
## [13,] 3.756708 -0.3153063
## [14,] 2.796773  1.9171775
## [15,] 3.142316  1.5452393
## [16,] 3.364824  1.9736312
## [17,] 1.625479  4.5668874
## [18,] 3.363095  1.8630202
## [19,] 3.756838  1.1431897
## [20,] 3.831592 -0.5266128
```

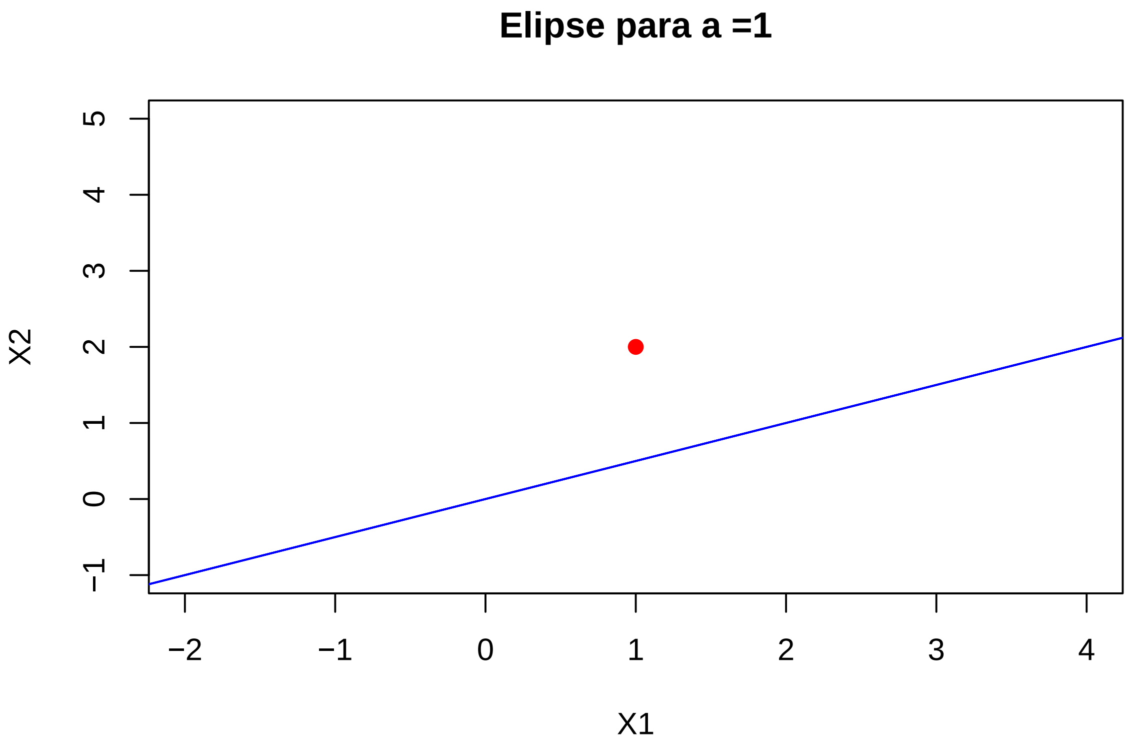
Questão 56 (caderno)

Questão 57

```r
mu <- c(1, 2)
a <- 0
sigma1 <- matrix(c(2, a, a, 2),2)
a <- (-1/2)
sigma2 <- matrix(c(2, a, a, 2),2)
a <- (1/2)
sigma3 <- matrix(c(2, a, a, 2),2)
a <- 1
sigma4 <- matrix(c(2, a, a, 2),2)

ellipse_params <- ellipse::ellipse(mu, sigma4, level = 0.95)

plot(0, 0, type = "n", xlim = c(mu[1] - 3, mu[1] + 3), ylim = c(mu[2] - 3, mu[2] + 3),
       xlab = "X1", ylab = "X2", main = paste0("Elipse para a =", a))
polygon(ellipse_params, border = "blue", col = NA)
points(mu[1], mu[2], pch = 19, col = "red")  # Marcar a média com um ponto vermelho
```



Questão 58 (caderno)

Questão 59

Questão 60

Ex. 4.26 | Johnson & Wichern

a)

De x_1 e x_2, obtemos: o o vetor de médias $\mu =$

```
##        [,1]
## [1,]  5.200
## [2,] 12.481
```

A matriz $\mathbf{S} =$

```
##           [,1]      [,2]
## [1,]  10.62222 -17.71022
## [2,] -17.71022  30.85437
```

E a inversa $\mathbf{S^{-1}} =$

```
##          [,1]      [,2]
## [1,] 2.189813 1.2569395
## [2,] 1.256939 0.7538861
```

Com isso, podemos calcular as distâncias estatísticas quadradas

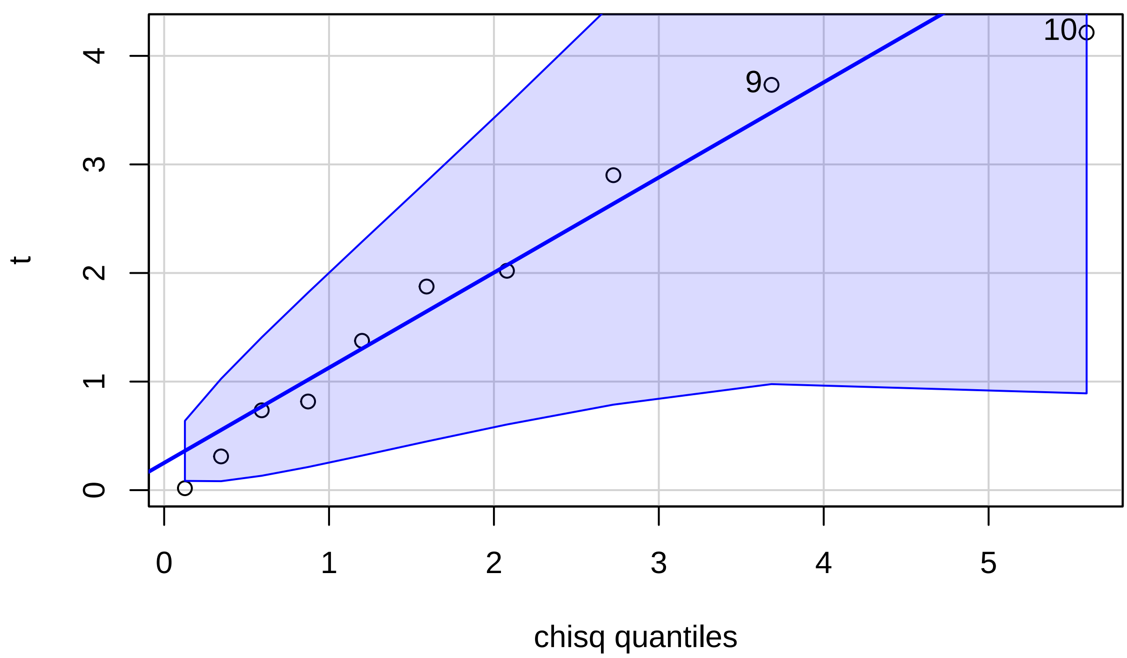
$d_j^2 = (\mathbf{x_j} - \mathbf{\bar{x}})^{\mathbf{T}}\mathbf{S}^{-1}(\mathbf{x_j} - \mathbf{\bar{x}}) =$ [1.8753045, 2.0203262, 2.9009088, 0.7352659, 0.3105192, 0.0176162, 3.7329012, 0.8165401, 1.3753379, 4.2152799]

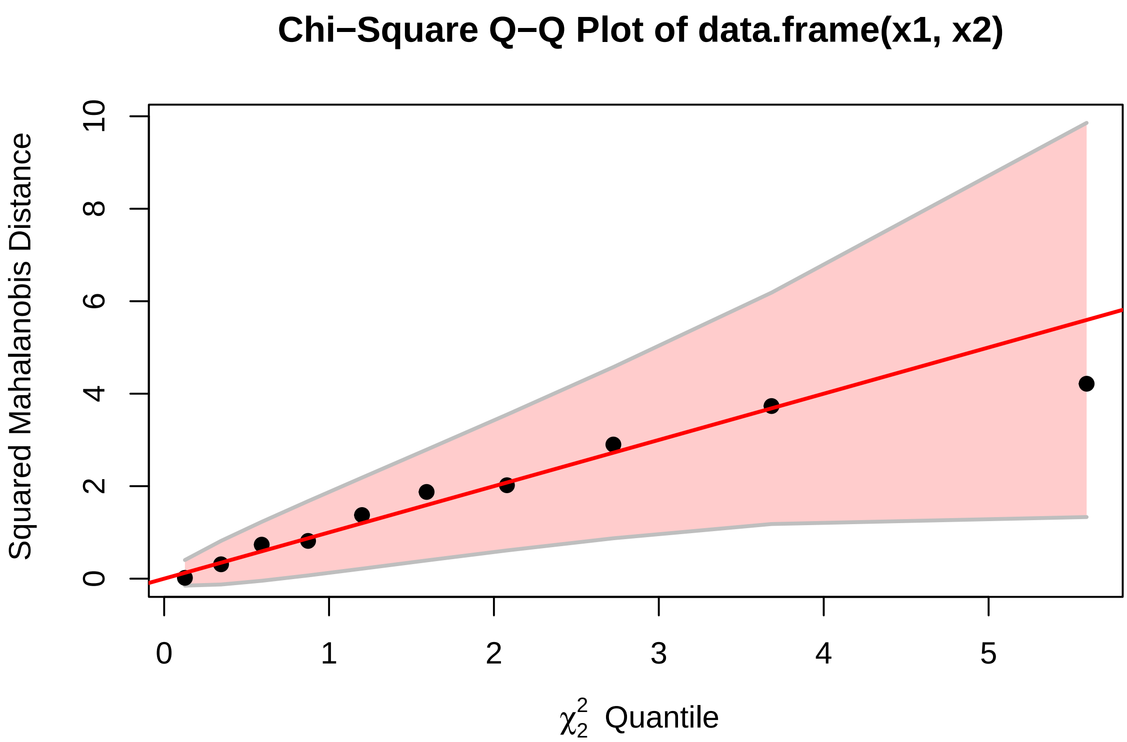
b)

Neste caso, iremos comparar os valores d_j^2 com o quantil $\chi_2^2(0,5) = 1.3862944$ e avaliar a proporção de observações na margem de aceitação, que para este caso é 50%

c)

Duas representações gráficas análogas:





d)

Pelo resultado da proporção de distâncias não rejeitadas pelo quantil qui-quadrado, pelo baixo número de dados e pelos gráficos acima, creio não haver evidências suficientes para rejeitar a normalidade bivariada destes dados

Questão 61

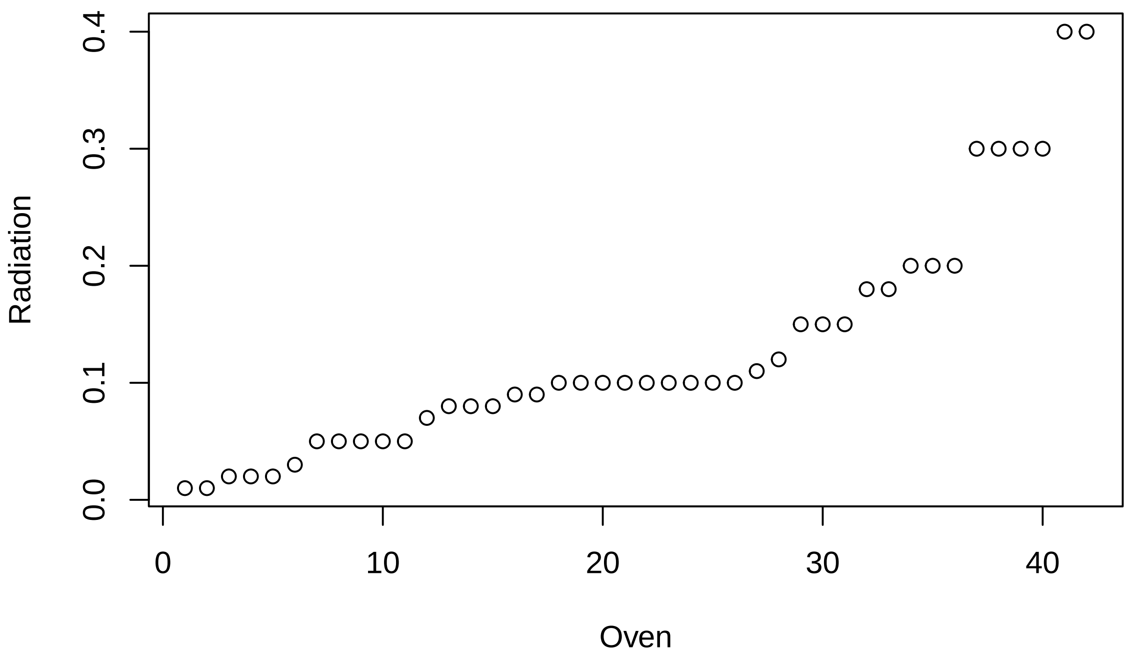
Ex. 4.27 | Johnson & Wichern

Algumas opções de teste de normalidade multivariada

Caso 1: Variáveis sem transformação

Test	Statistic	p-value	Result
Skewness	11.0013	0.0265	NO
Kurtosis	0.4296	0.6675	YES
MV Normality	NA	NA	NO

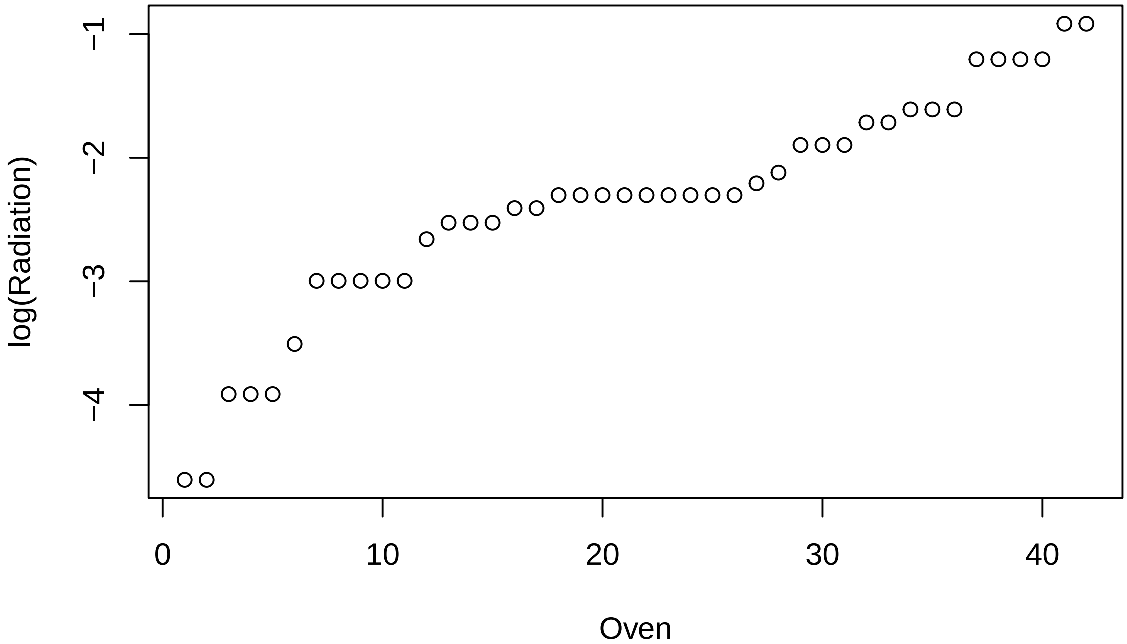
Test	HZ	p value	MVN
Henze-Zirkler	1.520541	0.0007484	NO



Caso 2: Variáveis com transformação $\lambda = 0$ (ln)

Test	Statistic	p-value	Result
Skewness	8.0963	0.0881	YES
Kurtosis	-0.5458	0.5852	YES
MV Normality	NA	NA	YES

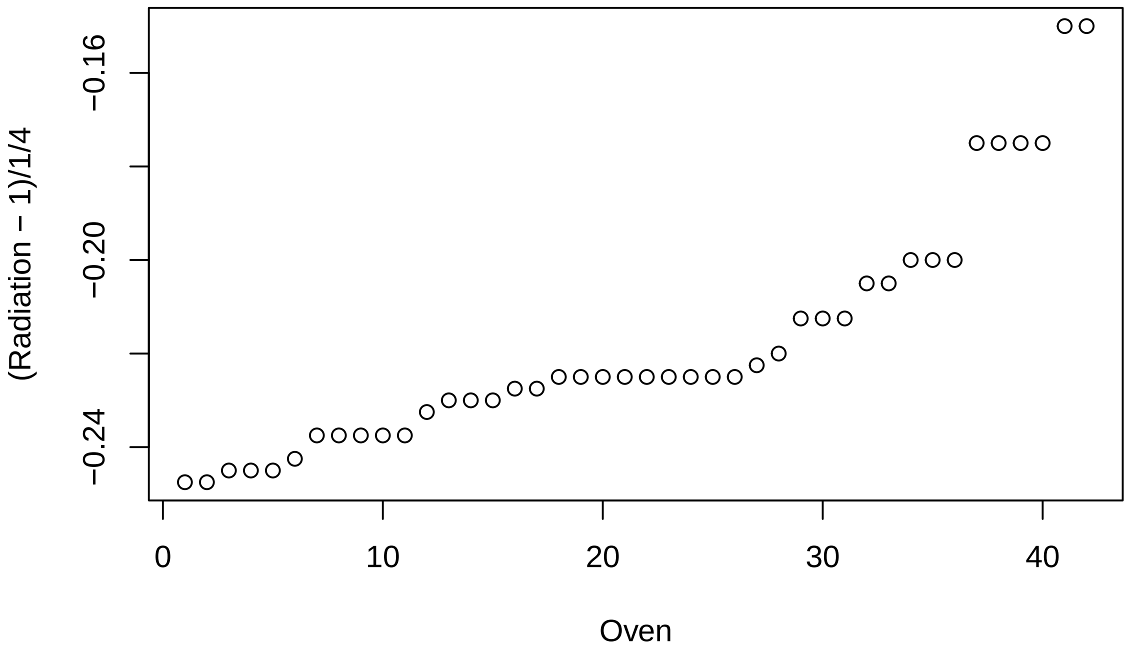
Test	HZ	p value	MVN
Henze-Zirkler	1.22835	0.0045145	NO



Caso 3: Variáveis com transformação $\lambda = 1/4$ $(\frac{x^{(\lambda)-1}}{\lambda})$

Test	Statistic	p-value	Result
Skewness	11.0013	0.0265	NO
Kurtosis	0.4296	0.6675	YES
MV Normality	NA	NA	NO

Test	HZ	p value	MVN
Henze-Zirkler	1.520541	0.0007484	NO



Portanto, apesar de ser bem difícil de inferir uma conclusão, a transformação $\lambda = 0$ aparenta ter trazido o melhor resultado de normalidade multivariada

Questão 62

Ex. 4.35 | Johnson & Wichern

Test	Statistic	p-value	Result
Skewness	127.17	0	NO
Kurtosis	10.4578	0	NO
MV Normality	NA	NA	NO

Test	HZ	p value	MVN
Henze-Zirkler	1.89379	4e-07	NO

Test	Variable	Statistic	p value	Normality
Anderson-Darling	Density	1.1852	0.0038	NO
Anderson-Darling	Strength_MachineDirection	0.3001	0.5661	YES
Anderson-Darling	Strength_CrossDirection	2.7420	<0.001	NO

	Beta-hat	kappa	p-val
Skewness	17.28145	118.089941	0
Kurtosis	30.62636	9.133963	0

```
##
##  Shapiro-Wilk normality test
##
## data:  Z
## W = 0.56907, p-value = 8.969e-10
```

Diversos testes de normalidade multivariada e marginal univariada foram testados, e à excessão de um teste de normalidade marginal da variável *Machine Direction*, todos os demais rejeitaram a hipótese nula de normalidade multivariada. Portanto, há evidências para descartar a hipótese nula de normalidade multivariada desses dados. Entretando, é possível que transformadas dessas variáveis não rejeitem a hipótese nula de normalidade multivariada.

Questão 63

Ex. 4.1 | Rencher & Christensen

Questão 64

Ex. 4.2 | Rencher & Christensen

Questão 65

Ex. 4.10 | Rencher & Christensen

Questão 66

Ex. 4.11 | Rencher & Christensen

Questão 67

Ex. 4.12 | Rencher & Christensen

Questão 68

Ex. 4.13 | Rencher & Christensen

Questão 69

Ex. 4.14 | Rencher & Christensen

Questão 70

Ex. 4.17 | Rencher & Christensen