

UNICAMP - Universidade Estadual de Campinas  
IMECC - Instituto de Matemática, Estatística e Computação Científica  
Departamento de Estatística  
ME731 - Métodos em Análise Multivariada

## **Trabalho**

Carlos Caminho Oco RA 168624  
marfloy Chaves eduardo RA 173733  
Caio Henrique de Sousa Lima RA 214144  
Piupiu Mendes Panela RA 219494  
Nathan Blusa Amaarela Soltinho RA 222854

Campinas  
2020

# 1 Questão 1:

Lista 3 questão 4 MANOVA e CBU

## 1.1 Introdução

O banco de dados é composto pelo perfil de tamanho e formato de 48 tartarugas, sendo 24 machos e 24 fêmeas. As informações obtidas foram largura, comprimento e altura. O estudo tinha como objetivo observar e comparar o sexo ao qual o animal pertencia e sua relação com essas medidas físicas. Para realizar a análise desse banco, foi utilizado o software estatístico R e aplicado algumas metodologias como a Análise de Variância Multivariada (MANOVA).

## 1.2 Análise descritiva

Tabela 1: Medidas resumo (F = Fem, M = Masc)

	Média	Var.	DP	CV(%)	Min.	Med.	Max.	CA	Cur.
length (F)	136.00	451.39	21.25	15.62	98.00	136.50	177.00	-0.23	2.26
width (F)	102.58	171.73	13.10	12.77	81.00	102.00	132.00	0.31	2.55
height (F)	51.96	66.65	8.16	15.71	38.00	51.00	67.00	-0.03	2.19
length (M)	113.38	138.77	11.78	10.39	93.00	115.00	135.00	-0.08	2.10
width (M)	88.29	50.04	7.07	8.01	74.00	89.00	106.00	0.20	3.15
height (M)	40.71	11.26	3.36	8.24	35.00	40.00	47.00	0.18	2.20

Inicialmente, visualizando as medidas resumo para cada sexo, podemos observar que aparentemente o sexo feminino tem, em média, medidas maiores que o sexo masculino.

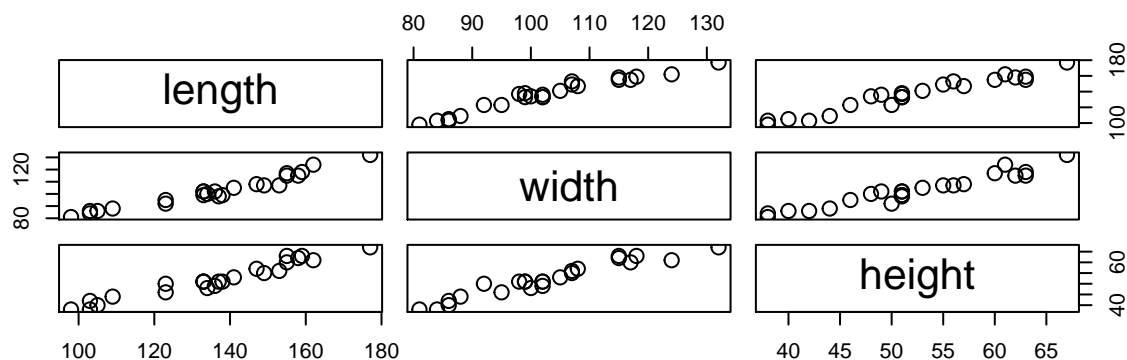


Figura 1: Matriz de gráfico de dispersões sexo feminino.

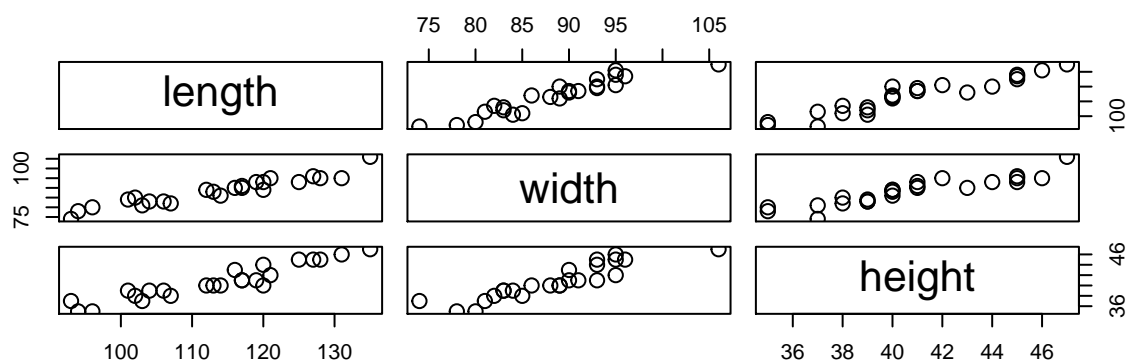


Figura 2: Matriz de gráfico de dispersões sexo masculino.

Para ambos os sexos, as medidas de altura, largura e comprimento são bastante correlacionadas.

Tabela 2: Matriz de covariâncias sexo feminino

	length	width	height
length	451.39	271.17	168.70
width	271.17	171.73	103.29
height	168.70	103.29	66.65

A matriz de covariância para ambos os sexos parecem ser diferentes, na qual as medidas do sexo masculino aparentam ter variância bem menor se comparado ao sexo feminino.

#### 4) Matriz de correlações

Tabela 3: Matriz de covariâncias sexo masculino

	length	width	height
length	138.77	79.15	37.38
width	79.15	50.04	21.65
height	37.38	21.65	11.26

Feminino

```
##           length      width      height
## length 1.0000000 0.9739706 0.9725816
## width   0.9739706 1.0000000 0.9654187
## height  0.9725816 0.9654187 1.0000000
```

Masculino

```
##           length      width      height
## length 1.0000000 0.9497846 0.9455580
## width   0.9497846 1.0000000 0.9122648
## height  0.9455580 0.9122648 1.0000000
```

Como dito anteriormente, as correlações parecem ser bem altas entre as medidas (já era esperado). #####

### 5) Boxplot

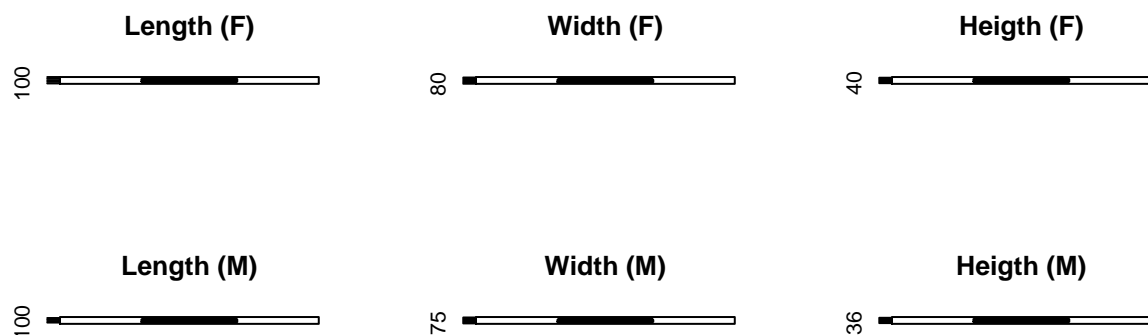


Figura 3: Boxplots das variáveis por sexo.

#histograms

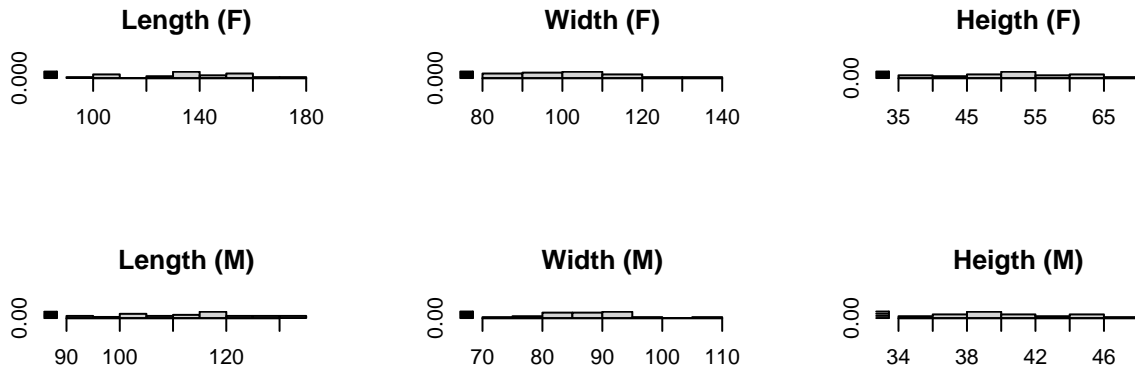


Figura 4: Histogramas das variáveis por sexo.

## 2 qqplots

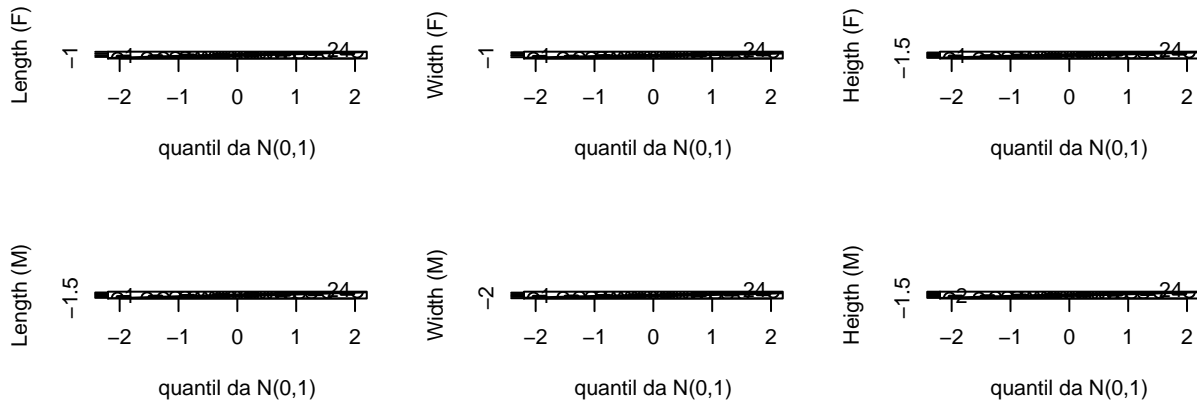


Figura 5: QQplots das variáveis por sexo.

Olhando para boxplot, histogramas e qqplots, as variáveis não parecem seguir uma distribuição normal.

### 2.1 Análise inferencial

## Estatística do Teste: 24.03738

```

## nível descritivo: 0.0005140546
## Matrizes de Covariâncias por grupo:
##      grupo    length    width    height
## length    1 451.39130 271.17391 168.69565
## width     1 271.17391 171.73188 103.28623
## height    1 168.69565 103.28623  66.65036
## length    2 138.76630  79.14674  37.37500
## width     2  79.14674  50.04167  21.65399
## height    2  37.37500  21.65399  11.25906

##      Df  Wilks approx F num Df den Df    Pr(>F)
## sex      1 0.40796    21.284      3    44 1.141e-08 ***
## Residuals 46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##      Df  Pillai approx F num Df den Df    Pr(>F)
## sex      1 0.59204    21.284      3    44 1.141e-08 ***
## Residuals 46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##      Df Hotelling-Lawley approx F num Df den Df    Pr(>F)
## sex      1      1.4512    21.284      3    44 1.141e-08 ***
## Residuals 46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##      Df    Roy approx F num Df den Df    Pr(>F)
## sex      1 1.4512    21.284      3    44 1.141e-08 ***
## Residuals 46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Tabela 4: Teste de Box-Cox para homocedasticidade

	Estatística do teste	p-valor
Box-Cox	24.04	0.0005

Tabela 5: Resultados da MANOVA

	Valor	Aproximação pela distribuição F	p-valor
Wilks	0.41	21.28	< 0.0001
Pillai	0.59	21.28	< 0.0001
Hotelling-Lawley	1.45	21.28	< 0.0001
Roy	1.45	21.28	< 0.0001

Para observar inferencialmente a diferença entre as matrizes de covariância de cada grupo, foi realizado um teste de Box-cox, e tivemos como resultado que as matrizes realmente são diferentes. Após isso, foi realizado uma MANOVA com o objetivo de visualizar se existe diferença entre os vetores de médias para ambos os sexos, e também foi rejeitada a hipótese de que elas são iguais. Dessa forma, queremos ver em qual medidas essas diferenças são mais acentuadas, portanto aplicamos um teste CBU.

```
## Response length :
##
## Call:
## lm(formula = length ~ sex)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -38.000 -10.625   0.812  11.969  41.000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   136.000      3.506  38.786 < 2e-16 ***
## sexM          -22.625      4.959  -4.563 3.75e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```

## Residual standard error: 17.18 on 46 degrees of freedom
## Multiple R-squared:  0.3116, Adjusted R-squared:  0.2966
## F-statistic: 20.82 on 1 and 46 DF,  p-value: 3.752e-05
##
##
## Response width :
##
## Call:
## lm(formula = width ~ sex)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.5833  -5.5417  -0.4375   4.8854  29.4167
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   102.583      2.149   47.725  < 2e-16 ***
## sexM          -14.292      3.040   -4.701 2.38e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.53 on 46 degrees of freedom
## Multiple R-squared:  0.3246, Adjusted R-squared:  0.3099
## F-statistic:  22.1 on 1 and 46 DF,  p-value: 2.376e-05
##
##
## Response height :
##
## Call:
## lm(formula = height ~ sex)
##
## Residuals:

```



```

##      Min      1Q   Median      3Q      Max
## -13.9583 -2.7708 -0.7083   4.1042  15.0417
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    51.958      1.274  40.783 < 2e-16 ***
## sexM          -11.250      1.802  -6.244 1.25e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.241 on 46 degrees of freedom
## Multiple R-squared:  0.4587, Adjusted R-squared:  0.447
## F-statistic: 38.99 on 1 and 46 DF,  p-value: 1.246e-07

##          vbeta   epbeta      estt      pvalor
## [1,] 136.00000 3.506416 38.786047 0.000000e+00
## [2,] 102.58333 2.149484 47.724628 0.000000e+00
## [3,]  51.95833 1.274014 40.783157 0.000000e+00
## [4,] -22.62500 4.958820 -4.562577 3.752223e-05
## [5,] -14.29167 3.039830 -4.701470 2.376030e-05
## [6,] -11.25000 1.801729 -6.244004 1.246117e-07

```

Tabela 6: Estimativas dos parâmetros do modelo

	Estimativa	EP	Estatística t	p-valor
$\mu_1$	136.00	3.51	38.79	< 0.0001
$\mu_2$	102.58	2.15	47.72	< 0.0001
$\mu_3$	51.96	1.27	40.78	< 0.0001
$\alpha_{21}$	-22.63	4.96	-4.56	< 0.0001
$\alpha_{22}$	-14.29	3.04	-4.70	< 0.0001
$\alpha_{23}$	-11.25	1.80	-6.24	< 0.0001

##Variável Length

## Estatística Qui-quadrado = 20.82

```

## pvalor = 0
## Matriz C :
##      [,1] [,2]
## [1,]    0    1
## Matriz U :
##      [,1]
## [1,]    1
## [2,]    0
## [3,]    0
## Matriz M :
## [1] 0

```

##Variável Width

```

## Estatistica Qui-quadrado = 22.1
## pvalor = 0
## Matriz C :
##      [,1] [,2]
## [1,]    0    1
## Matriz U :
##      [,1]
## [1,]    0
## [2,]    1
## [3,]    0
## Matriz M :
## [1] 0

```

##Variável Height

```

## Estatistica Qui-quadrado = 38.99
## pvalor = 0
## Matriz C :
##      [,1] [,2]
## [1,]    0    1

```

```
## Matriz U :  
##      [,1]  
## [1,]    0  
## [2,]    0  
## [3,]    1  
## Matriz M :  
## [1] 0
```

Observando o resultado dos teste CBU, podemos observar que as diferenças estão em todas as medidas (largura, comprimento e altura).

#MRLNM (estimativas parâmetros)

Aqui eu não sei o que falar

OBS: DEPOIS TEM QUE VER A LISTA FEITA PRA VER SE FALTA ALGO