



---

# TRABALHO - PARTE 1

## RELATÓRIO - QUESTÃO 1

---

ELIANE RAMOS DE SIQUEIRA RA:155233

GUILHERME PAZIAN RA:160323

HENRIQUE CAPATTO RA:146406

MURILO SALGADO RAZOLI RA:150987

Disciplina: **ME731 - Análise Multivariada**  
Professor: **Caio Lucidius Naberezny Azevedo**

Campinas - SP  
18 de Novembro de 2017

## 1. Introdução

O conjunto de dados utilizado neste relatório são relativos a moscas chamadas, em inglês, de “bitting fly”. Neste conjunto de dados foram consideradas no total 70 moscas, sendo 35 da espécie *Leptoconops carteri* e 35 da *Leptoconops torrens*. Tais espécies serão tratadas a partir de agora por Carteri e Torrens, respectivamente.

Para cada uma das espécies, foram medidas sete variáveis referentes ao aspecto morfológico das moscas e para cada unidade amostral (mosca), foram medidas oito variáveis, sendo elas: espécie (0 – Torrens e 1 – Carteri), comprimento da asa (CA), largura da asa (LA), comprimento do 3º palpo (CP3), largura do 3º palpo (LP3), comprimento do 4º palpo (CP4), comprimento do 12º segmento da antena (SA12) e comprimento do 13º segmento da antena (SA13).

Dado que as duas espécies são bastante semelhantes morfolologicamente (Johnson e Wichern (2007)), o objetivo desta análise é realizar a comparação das médias de todas as variáveis consideradas entre as espécies, com o intuito de verificar se e quais variáveis diferem entre os grupos.

Como abordagem inicial, será utilizada a Análise de Variância Multivariada (MANOVA) (veja mais em Johnson e Wichern (2007)), para verificar a possível existência de diferenças entre as médias. Posteriormente, caso a hipótese de igualdade entre as médias for rejeitada serão realizados testes do tipo  $CBU = M$  (veja mais em Azevedo (2017)), com o intuito de descobrir onde essas diferenças se encontram.

Todas as análises serão realizadas com o suporte dos softwares R versão 3.4.2 e R Studio versão 1.1.383

## 2. Análise descritiva

A tabela 1 mostra algumas medidas resumo para as todas variáveis citadas anteriormente, separadas por espécie. É possível notar que as médias amostrais para as variáveis comprimento da asa, comprimento do 3º palpo e do 4º palpo são relativamente diferentes entre as espécies, enquanto que para as demais as médias amostrais são relativamente iguais. Pode se notar também, que os desvios padrões para as variáveis largura da asa e comprimento do 4º palpo aparentemente são diferentes entre as espécies, enquanto que para as demais variáveis os desvios padrão são consideravelmente próximos.

A figura 1 apresenta uma matriz de diagramas de dispersão entre as variáveis, separadas por espécie. De um modo geral, podemos notar que a espécie Torrens apresenta valores inferiores à espécie Carteri, como por exemplo, no gráfico de Largura por Comprimento do 3º Palpo, em que os pontos versem se concentram abaixo dos azuis.

Cada diagrama apresenta seu respectivo valor do coeficiente de correlação linear, então pode-se notar que aparentemente as variáveis que mais se destacaram em relação à uma possível associação linear foram as variáveis referentes ao comprimento do 12º palpo e ao comprimento do 13º palpo (0,81) e as referentes ao comprimento e largura da asa (0,6). Além disso, observam-se também possíveis associações entre as demais variáveis, porém com grau menor de correlação, como por exemplo entre as variáveis referentes ao comprimento da asa e comprimento do 3º palpo (0,45) e entre comprimento da asa e comprimento do 4º palpo (0,48).

Na figura 2, temos os boxplots para todas as variáveis separadas por espécie. Algumas distribuições apresentam ligeiras assimetrias, observando de um modo geral. Por exemplo na variável largura da asa para a espécie Torrens, essa assimetria é um pouco mais evidente. Podemos afirmar que as distribuições quando comparadas entre as espécies são diferentes observando de um modo geral. É possível afirmar que as medianas são ligeiramente maiores para todas as variáveis na espécie Carteri, exceto para a variável referente ao comprimento do 13º segmento da antena. Por fim, nota-se a presença de alguns pontos discrepantes, sendo que em maioria das vezes, estes pontos estão mais presentes na espécie Carteri.

A figura 3 apresenta o gráfico de quantis-quantis com envelopes, para a distância de Mahalanobis (Azevedo(2017)), para cada espécie. Pode-se observar que a suposição de normalidade multivariada dos dados parece não ser razoável, uma vez que há grandes fugas para quantis maiores da distribuição F na parte superior dos gráficos.

Tabela 1: Medidas Resumo das variáveis por espécie

	Espécie	n	Média	Variância	Desvio Padrão	CV(%)	Mínimo	Mediana	Máximo
Comprimento da Asa	Carteri	35	96,457	40,726	6,382	6,616	85	95	109
	Torrens	35	99,343	31,291	5,594	5,631	82	99	112
Largura da Asa	Carteri	35	42,914	7,492	2,737	6,378	38	44	49
	Torrens	35	43,743	25,785	5,078	11,608	19	45	50
Comprimento 3º Palpo	Carteri	35	35,371	4,829	2,197	6,212	31	36	39
	Torrens	35	39,314	8,045	2,836	7,215	33	39	44
Largura 3º Palpo	Carteri	35	14,514	3,375	1,837	12,657	11	14	18
	Torrens	35	14,657	2,703	1,644	11,216	11	15	19
Comprimento 4º Palpo	Carteri	35	25,629	6,24	2,498	9,747	21	26	31
	Torrens	35	30,00	21,294	4,615	15,382	20	31	38
Comprimento 12º Seg antena	Carteri	35	9,571	0,84	0,917	9,577	8	9	13
	Torrens	35	9,657	1,585	1,259	13,036	6	10	12
Comprimento 13º Seg antena	Carteri	35	9,714	0,798	0,893	9,198	8	10	13
	Torrens	35	9,371	1,182	1,087	11,599	7	9	11

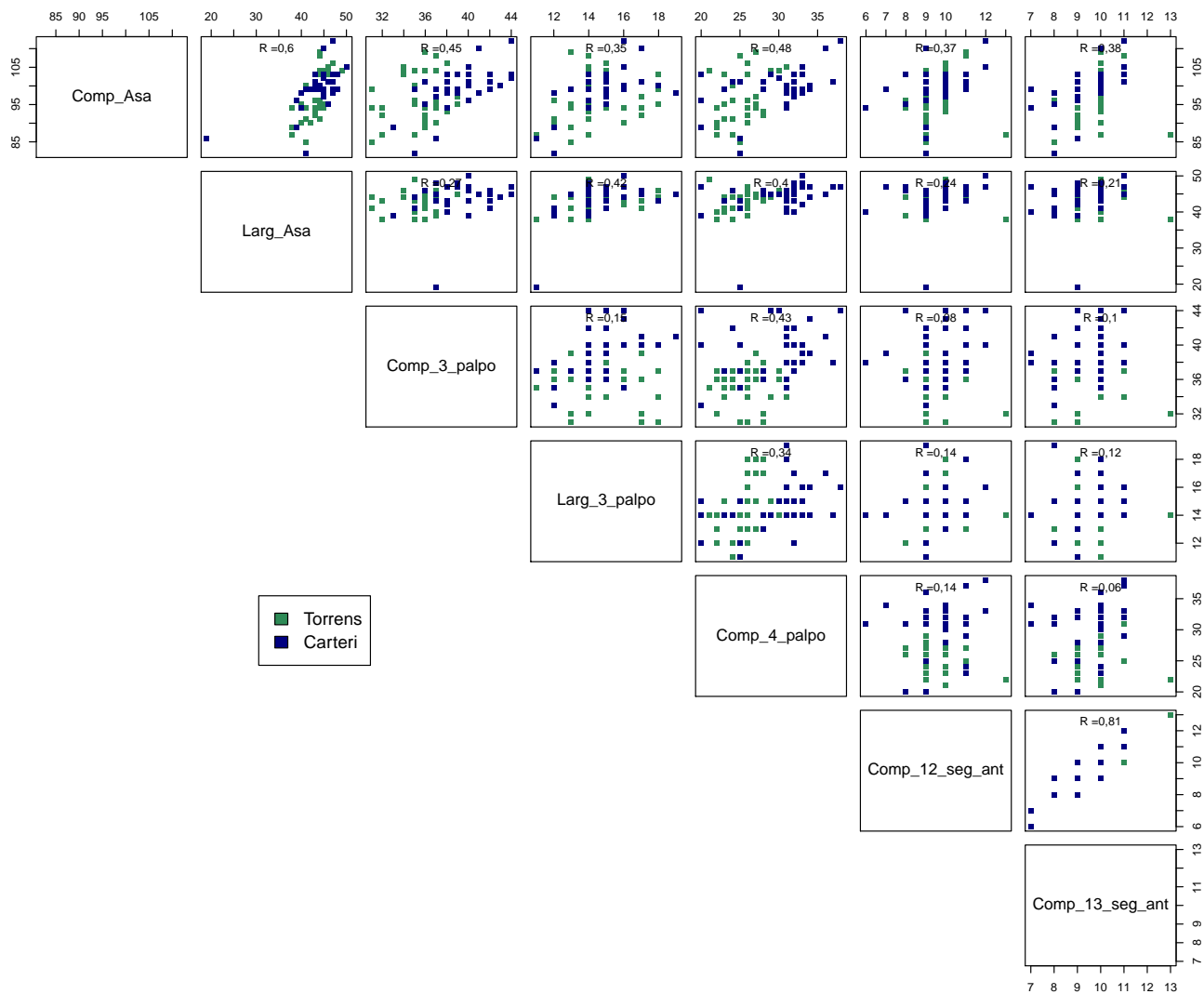


Figura 1: Matriz de diagramas de dispersão entre as variáveis

### 3. Análise Inferencial

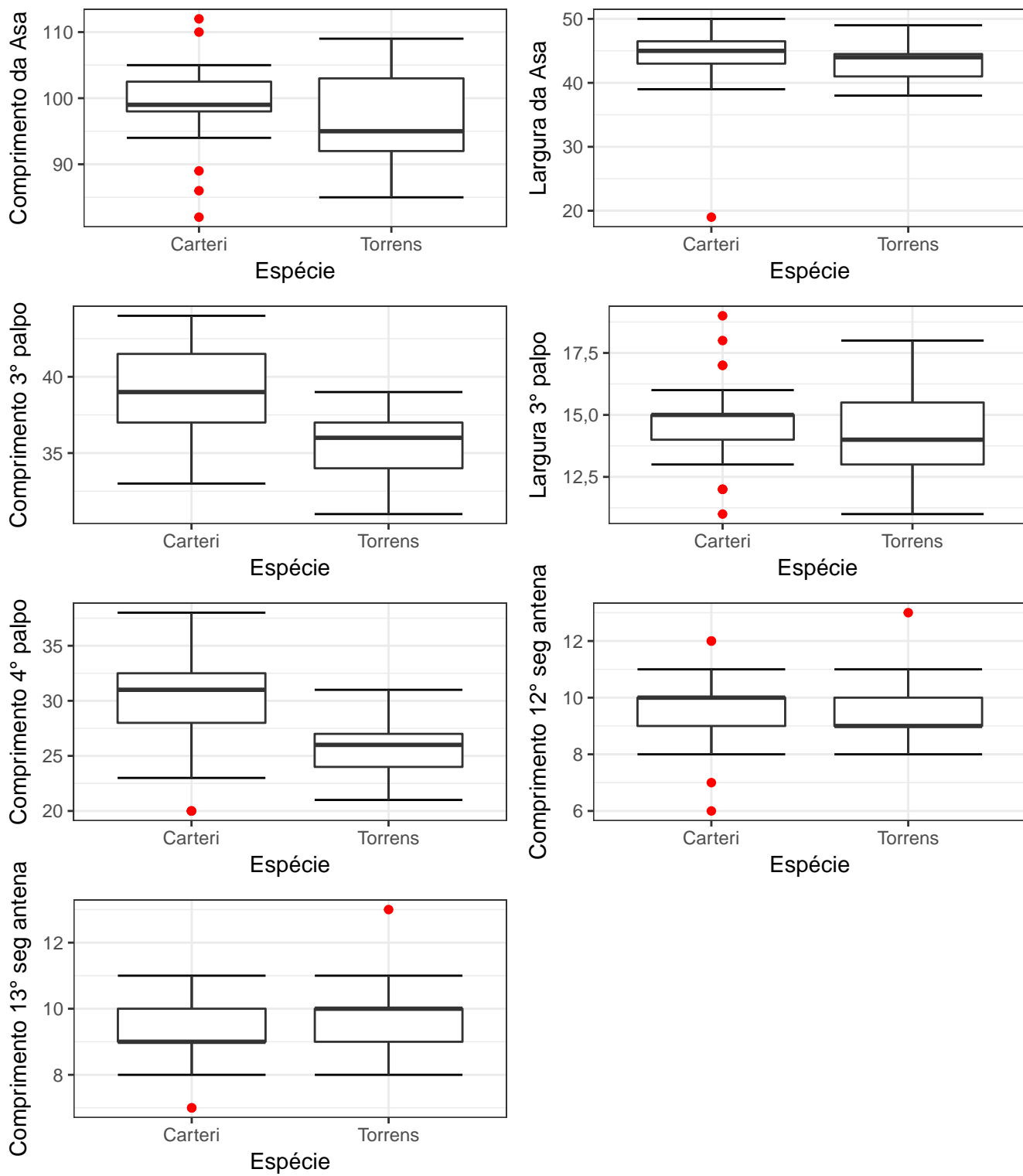


Figura 2: Box-plot das variáveis por espécie

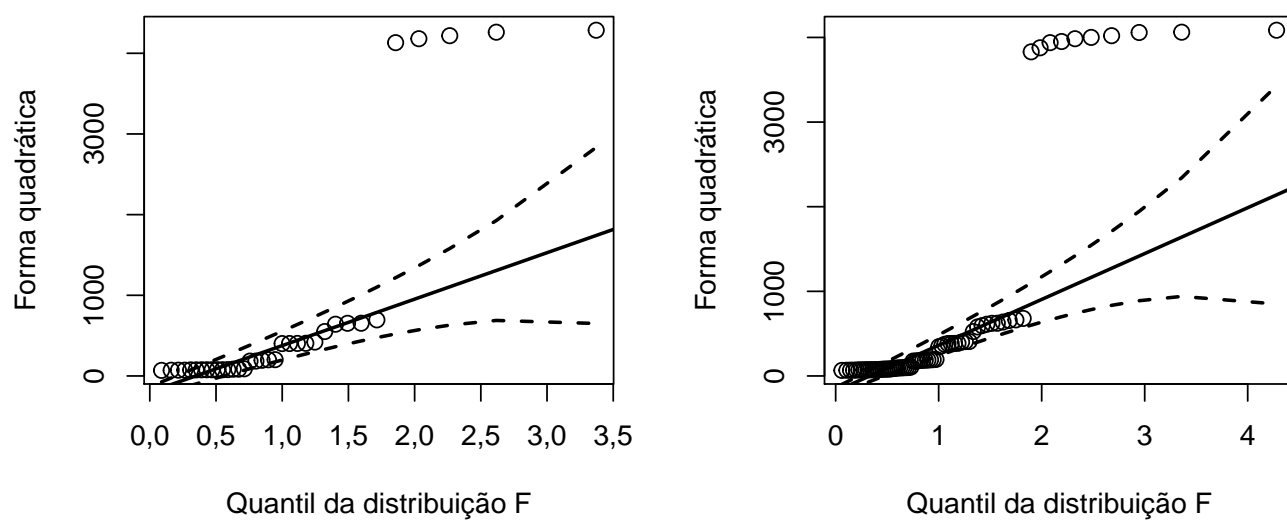


Figura 3: Gráfico de quantil-quantil com envelopes para a distância de Mahalanobis; A) Torrens, B) Carteri

#### 4. Conclusões

## 5. Bibliografia