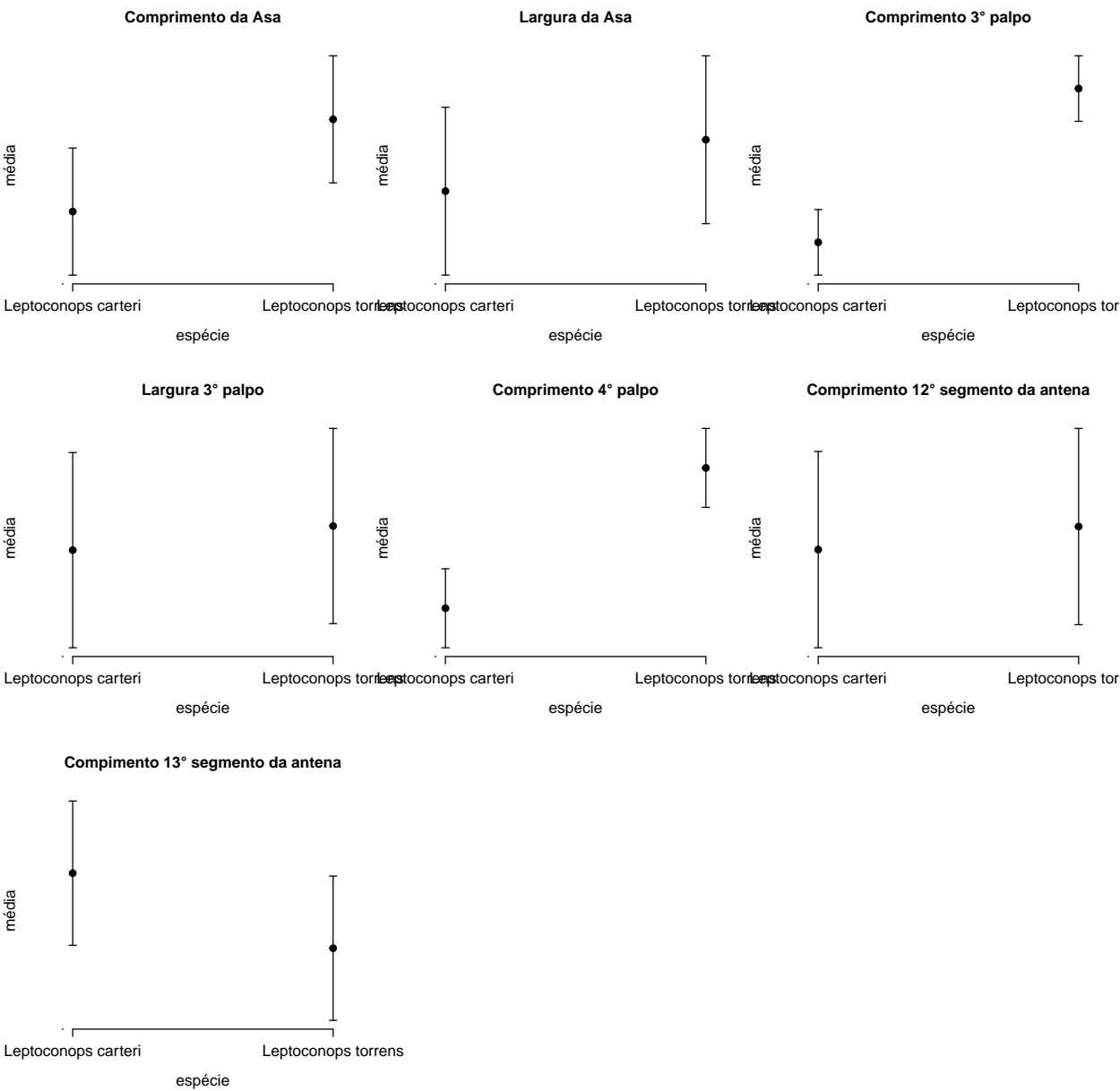


Modelo MANOVA

```
## tidyverse      plyr      plotrix      car      xtable
##      TRUE      TRUE      TRUE      TRUE      TRUE
```

Uma suposi  o adotada para se realizar um teste de an  lise de variancia multivariada   a de igualdade das matrizes de

```
##      Estat stica Valor Aprox..distr..F p.valor
## 1      Wilks 0,391      13,824 <0,001
## 2      Pillai 0,609      13,824 <0,001
## 3 Hotelling-Lawley 1,561      13,824 <0,001
## 4      Roy 1,561      13,824 <0,001
```



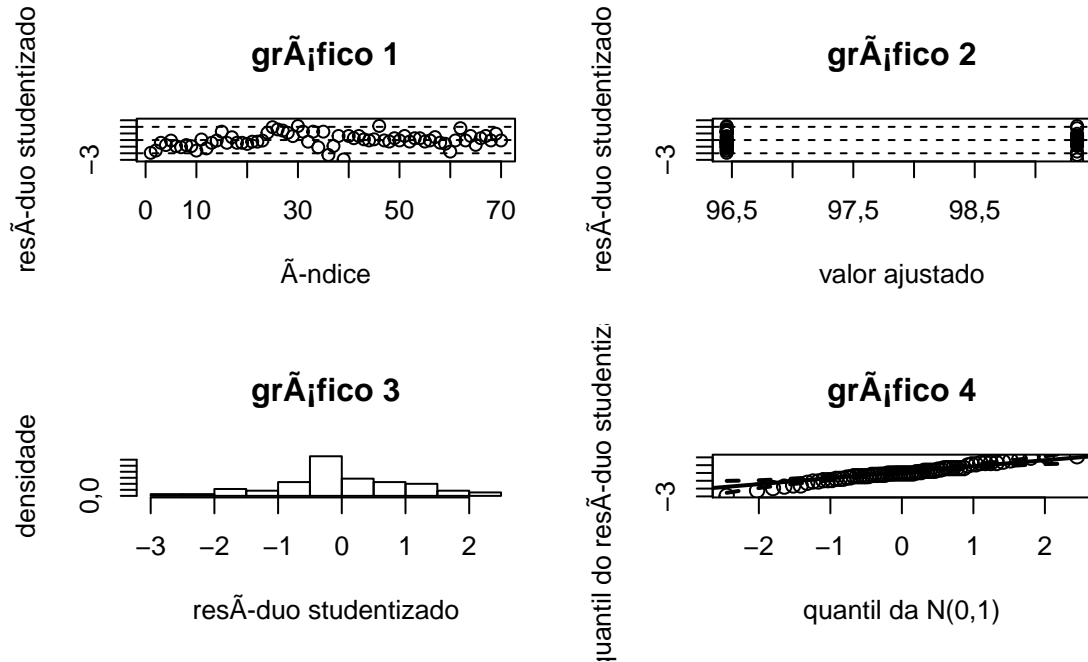


Figure 1: Gráficos para os resíduos referentes à variável Comprimento da Asa

Note pelos gráficos da figura (X-1) que os Intervalos de Confiança para as médias preditas para as espécies de moscas se interceptam num intervalo grande para as variáveis Largura da Asa, Largura 3º palpo e Comprimento 12º segmento da antena, portanto é razoável conjecturar que as espécies de moscas tem médias iguais para estas variáveis. Por meio da metodologia CBU=M (veja Azevedo,2017), testou-se simultaneamente a igualdade das médias destas variáveis entre as espécies de moscas, ao qual resultou num p-valor = 0,864, ou seja, não temos evidências estatísticas suficientes para rejeitar a hipótese de igualdade simultânea das médias entre as espécies para as variáveis Largura da Asa, Largura 3º palpo e Comprimento 12º segmento da antena. A fim de identificar melhor onde residem as diferenças entre as espécies de moscas, aplicamos esta mesma metodologia acrescentando as demais variáveis na hipótese de igualdade (uma de cada vez), os resultados deste teste constam na tabela (n XX “tabela_CBU”).

```
## Hipótese
## [1,] "$\alpha_{12} = \alpha_{14} = \alpha_{16} = 0$"
## [2,] "$\alpha_{12} = \alpha_{14} = \alpha_{16} = \alpha_{11} = 0$"
## [3,] "$\alpha_{12} = \alpha_{14} = \alpha_{16} = \alpha_{13} = 0$"
## [4,] "$\alpha_{12} = \alpha_{14} = \alpha_{16} = \alpha_{15} = 0$"
## [5,] "$\alpha_{12} = \alpha_{14} = \alpha_{16} = \alpha_{17} = 0$"
## p-valor do teste CBU=M
## [1,] "0,864"
## [2,] "0,342"
## [3,] "< 0,001"
## [4,] "< 0,001"
## [5,] "0,029"
```

Note pela tabela (n XX “tabela_CBU”) que os dois primeiros testes indicam a não rejeição da hipótese apresentada, portanto temos evidências estatisticamente significantes de que as espécies tem médias conjuntamente iguais para as variáveis Largura da Asa, Largura 3º palpo, Comprimento 12º segmento da antena e Comprimento da Asa, ou seja, as diferenças parecem residir nas variáveis Comprimento 3º palpo, Comprimento 4º palpo e Comprimento 13º segmento da antena.

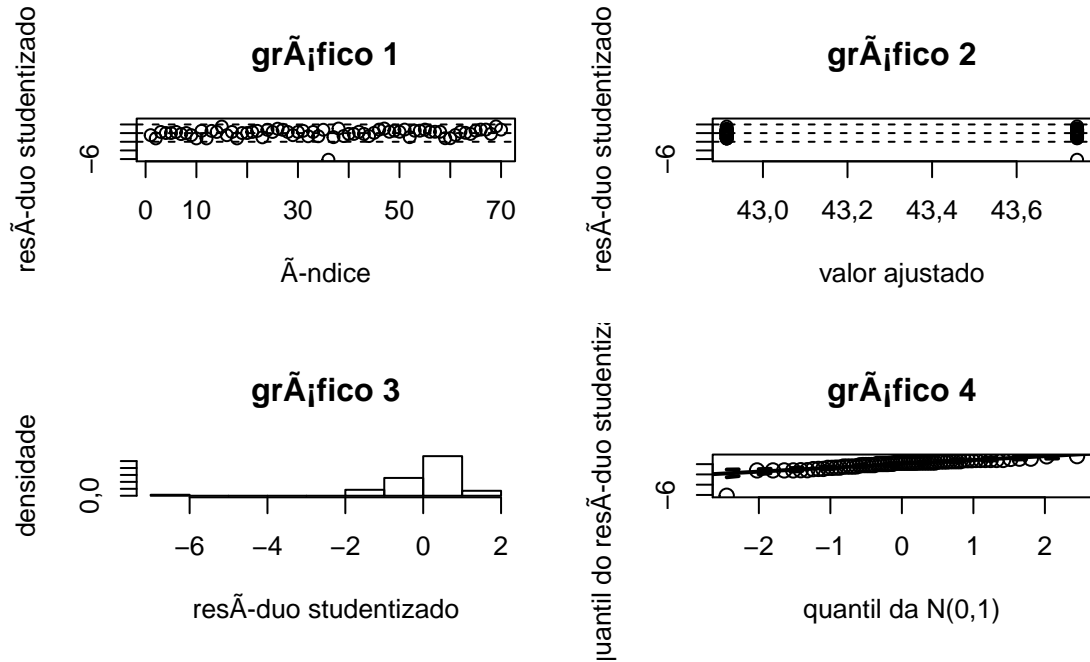


Figure 2: Gráficos para os resíduos referentes à variável Largura da Asa

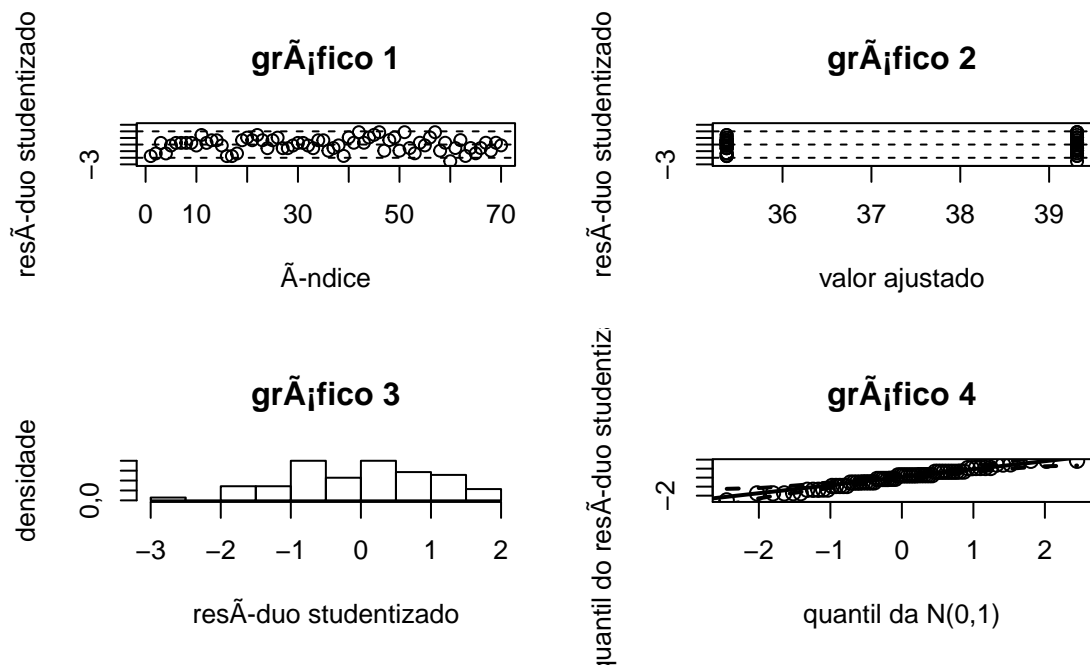


Figure 3: Gráficos para os resíduos referentes à variável Comprimento 3º palpo

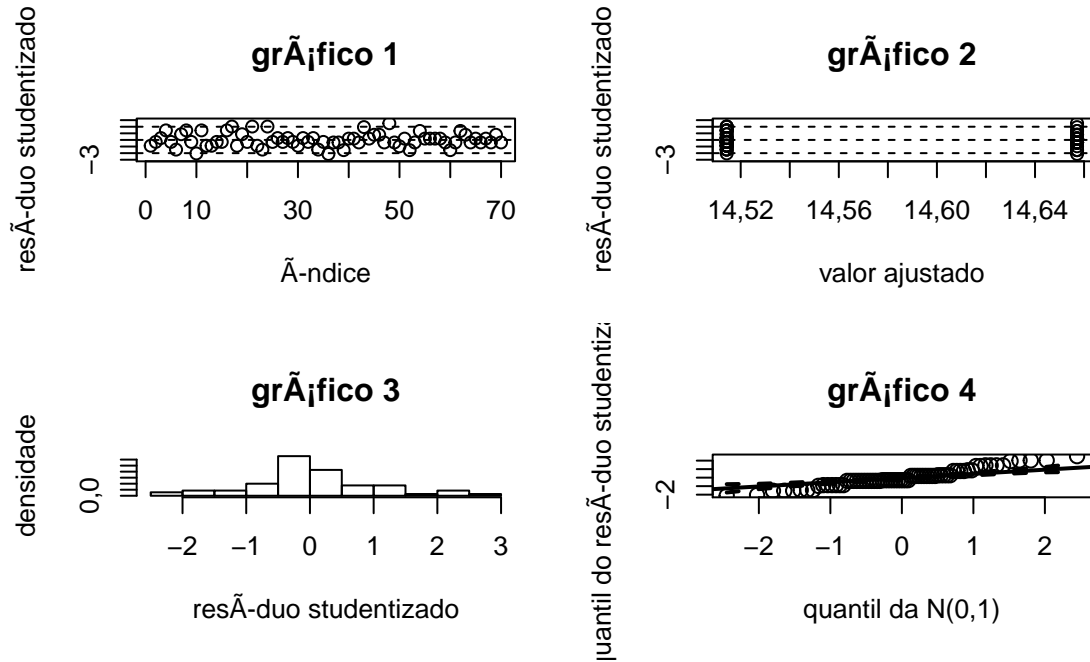


Figure 4: Gráficos para os resíduos referentes à variável Largura 3º palpo

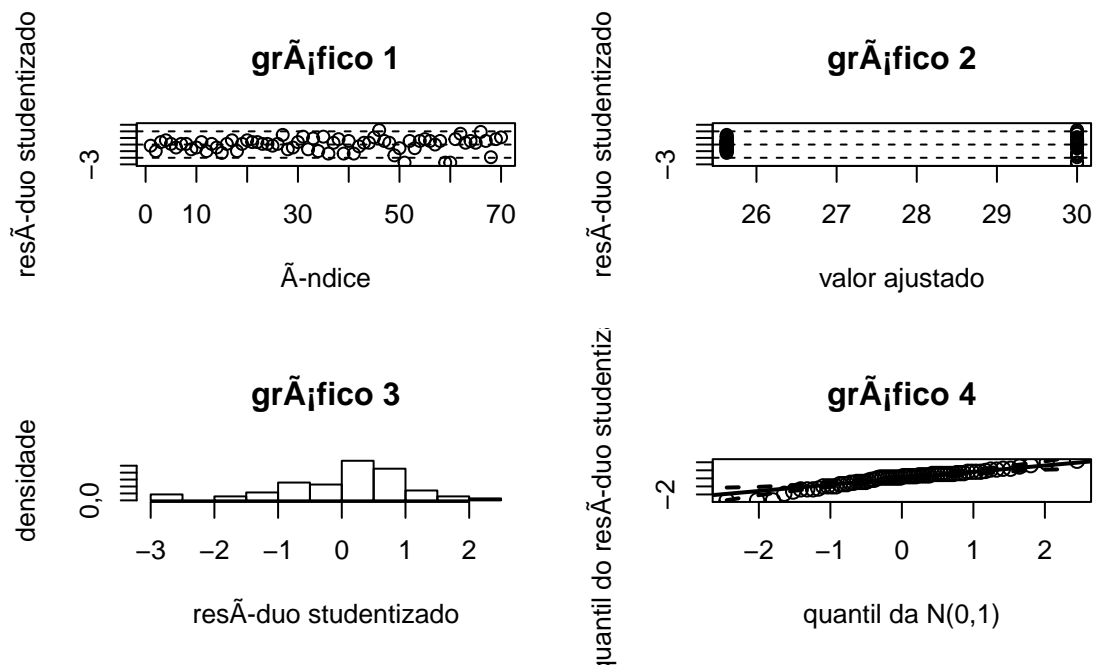


Figure 5: Gráficos para os resíduos referentes à variável Comprimento 4º palpo

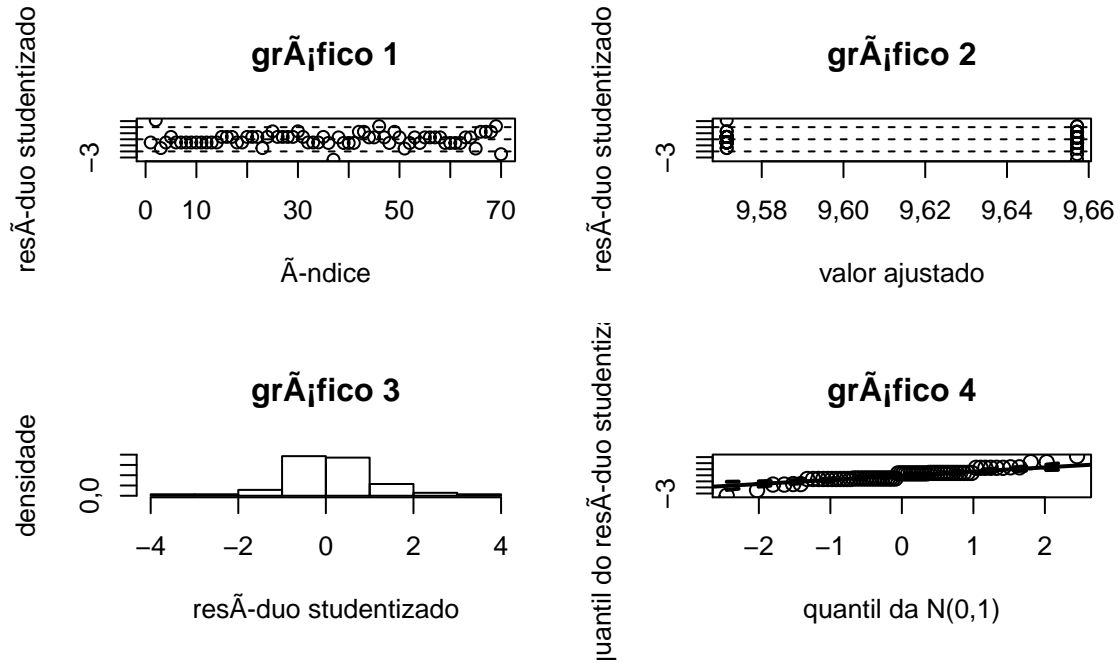


Figure 6: Gráficos para os resíduos referentes à variável Comprimento 12° segmento da antena

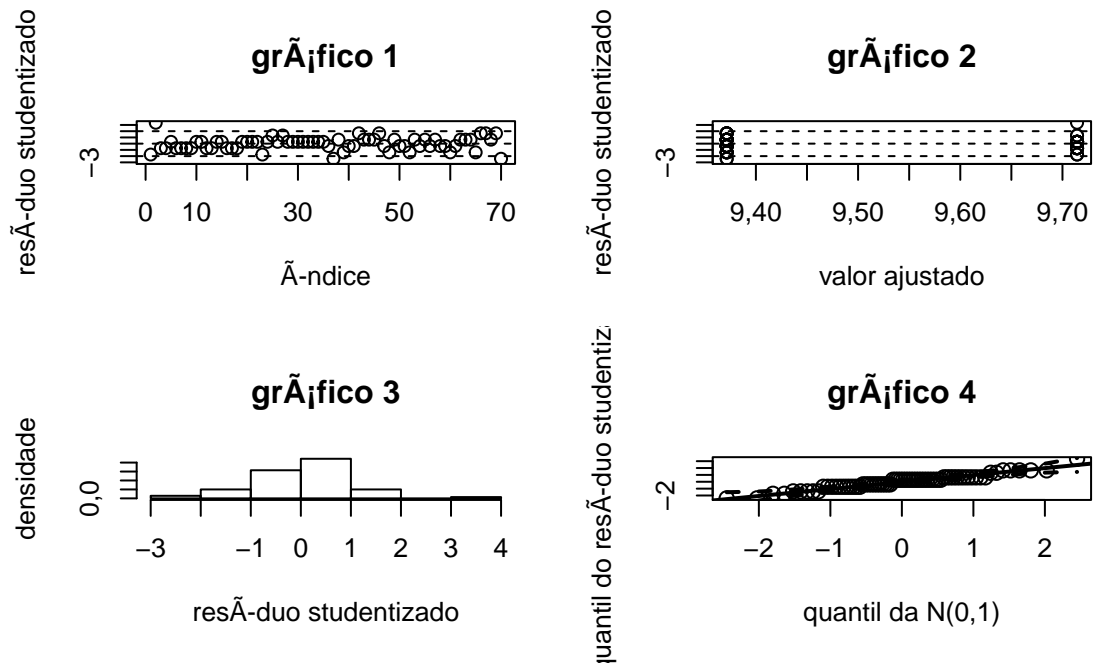
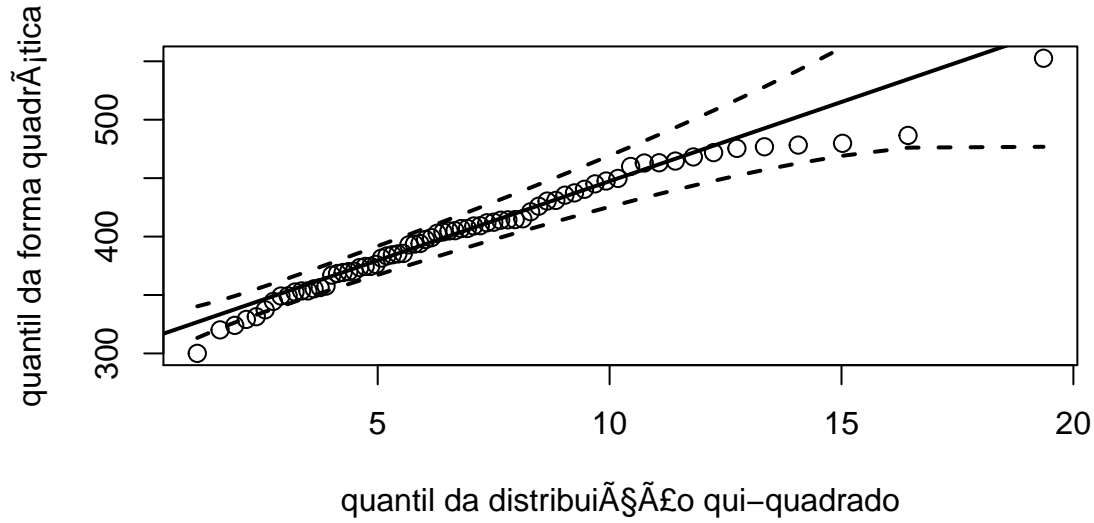


Figure 7: Gráficos para os resíduos referentes à variável Comprimento 13° segmento da antena



Com o objetivo de identificar as diferenças entre as espécies de moscas, ajustou um modelo de regressão normal linear homocedástico multivariado ajustado via mínimos quadrados generalizados (veja Azevedo, 2017):

$$Y_{ijk} = \mu_k + \alpha_{ik} + \varepsilon_{ijk}, \quad \varepsilon_{ijk} \sim N_k(0, \Sigma),$$

$i = 1, 2$ (espécie, 1 - *Leptoconops torrens*, 2 - *Leptoconops carteri*), $j = 1, 2, \dots, 35$ (moscas) e

$k = 1, \dots, 7$ (variável, 1 - Comprimento da Asa, 2 - Largura da Asa, 3 - Comprimento 3º palpo, 4 - Largura 3º palpo, 5 - Comprimento 4º palpo, 6 - Comprimento 12º segmento da antena, 7 - Comprimento 13º segmento da antena),

em que $\alpha_{1k} = 0, k = 1, \dots, 7$.

Na tabela XX(tbl_resultados_MANOVA) estão apresentadas as quatro estatísticas referentes ao teste de análise de variância multivariada:

TABELA tbl_resultados_MANOVA

Note pela tabela XX(tbl_resultados_MANOVA) que todas as estatísticas apresentaram p-valor $< 0,05$ (Melhorar AQUÍ) portanto, todos os testes nos apresentam evidências estatisticamente significativas de que as espécies diferem em ao menos uma das variáveis presentes no banco de dados.

3.1 Análise dos resíduos

X <- A o número do primeiro gráfico de resíduos

A fim de avaliar a validade das suposições de normalidade multivariada dos dados considerando as espécies (consequentemente normalidade univariada) e homocedasticidade multivariada entre as espécies (consequentemente homocedasticidade univariada), podemos observar as figuras (X) a (X+7) que apresentam gráficos para os resíduos studentizados para cada uma das 7 variáveis, assim como a figura (X+8) que apresenta o gráfico de envelopes baseado na distância de Mahalanobis (veja Azevedo, 2017). A partir da observação destes gráficos, pode-se identificar muitos comportamentos e tendências não esperadas, as quais podemos destacar o comportamento apresentado no gráfico 4 das figuras (X), (X+3), (X+5) e (X+6) tendo muitos pontos fora dos limites das bandas de confiança, nas figuras (X+1), (X+2) parece existir uma pequena tendência nos valores dos resíduos e na figura (X+4) apresenta muitos pontos com quantis baixos fora das bandas de confiança. Adicionalmente, identificamos comportamento assimétrico

negativo no gráfico 3 das figuras (X), (X+1), (X+2) e (X+4) e assimétrico positivo apresentando na figura (X+6). Dadas as observações referentes aos gráficos 1 e 4 das figuras (X) a (X+6) temos um forte indício de que a suposição de normalidade não é razoável para nenhuma das variáveis presentes no banco de dados. Observando o gráfico 2 das figuras (X) a (X+7), identificamos evidências de presença de heterocedasticidade dos dados nas figuras (X), (X+2), (X+3) de maneira mais leve e nas figuras (X+4), (X+5) e (X+6) de maneira mais acentuada, já para a figura (X+1) não se nota, a menos de um valor extremo, a presença de indícios de heterocedasticidade. Não identificamos nenhum comportamento a ser destacado referente ao gráfico 1 das figuras (X) a (X+7). Na figura (X+8) observamos alguns valores fora das bandas de confiança para valores menores de quantis da forma quadrática, além disso, valores maiores de quantis da forma quadrática tendem a se apresentar abaixo da linha de referência baseada no quantil da distribuição qui-quadrado, deste modo temos indícios de que a suposição de normalidade multivariada dos dados não parece ser uma suposição razoável neste caso. Contudo, dadas as observações destacadas, temos que a única variável do banco de dados a qual não seria irrisuável supor normalidade e homocedasticidade dos dados seria a variável “Largura da Asa”, e todas as restantes apresentam ao menos um indício evidente da fuga destas suposições portanto não seria razoável supor normalidade e homocedasticidade multivariada neste caso, o que também fica evidente na figura (X+8), sendo assim, o modelo de análise de variância multivariada não apresentou um ajuste adequado aos dados aqui analisados e se é necessário procurar técnicas alternativas para realizar uma análise adequada ao banco de dados. Dado o nosso contexto acadêmico, iremos continuar com as análises dos resultados para elaborar a conclusão do presente trabalho.