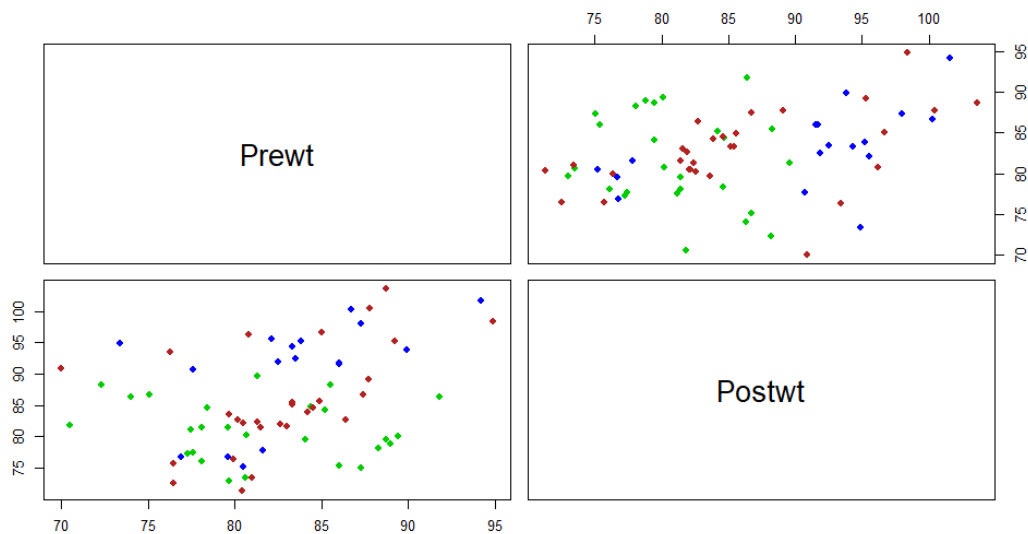


MÉTODOS ESTADÍSTICOS

Nombre(s): Equipo 4: <ul style="list-style-type: none">Diana Zepeda MartínezJosé Juan García Romero	
Nº	Realizar los ejercicios en Rstudio aplicando la teoría de LDA
24	Incluir la descripción de nuevas funciones y una interpretación general.
# BASE DE DATOS anorexia # Peso de adolescentes que padecen de anorexia, antes y después de # recibir uno de los tres posibles tratamientos: # CBT = cognitivo conductual, FT = terapia familiar, Cont = control	
data("anorexia")	
<div><div>Actividad 25.R x</div><div><div>Source on Save</div><div></div></div><div><pre>1 # Diana Zepeda Martínez 2 # José Juan García Romero 3 4 # Practica 25 - LDA 5 6 library(MASS) 7 data("anorexia") 8 9 library(ggplot2) 10 library(ggpubr) 11</pre></div><div><div>ConsoleTerminal xJobs x</div><div>R 4.1.3 · ~/</div><div><pre>> # Diana Zepeda Martínez > # José Juan García Romero > > # Practica 25 - LDA > > library(MASS) > data("anorexia") > library(ggplot2) > library(ggpubr) > </pre></div></div></div>	
a) verificar traslape individual con gráfico	
<pre>12 #a) verificar traslape individual con gráfico 13 14 p1 <- ggplot(data = anorexia, aes(x = Prewt, fill = Treat)) + 15 geom_histogram(position = "identity", alpha = 0.5) 16 p2 <- ggplot(data = anorexia, aes(x = Postwt, fill = Treat)) + 17 geom_histogram(position = "identity", alpha = 0.5) 18 19 ggarrange(p1, p2, nrow = 2, common.legend = TRUE) 20</pre>	
<pre>> #a) verificar traslape individual con gráfico > > p1 <- ggplot(data = anorexia, aes(x = Prewt, fill = Treat)) + + geom_histogram(position = "identity", alpha = 0.5) > p2 <- ggplot(data = anorexia, aes(x = Postwt, fill = Treat)) + + geom_histogram(position = "identity", alpha = 0.5) > ggarrange(p1, p2, nrow = 2, common.legend = TRUE) `stat_bin()` using `bins = 30`. Pick better value with `binwidth`. `stat_bin()` using `bins = 30`. Pick better value with `binwidth`. `stat_bin()` using `bins = 30`. Pick better value with `binwidth`. > </pre>	

b) Gráfico de correlación

```
21 #b) Gráfico de correlación
22
23 pairs(x = anorexia[, c("Prewt", "Postwt")],
24       col = c("firebrick", "green3", "blue")[anorexia$Treat],
25       pch = 19)
26
> #b) Gráfico de correlación
>
> pairs(x = anorexia[, c("Prewt", "Postwt")],
+       col = c("firebrick", "green3", "blue")[anorexia$Treat],
+       pch = 19)
> |
```



c) Generar el modelo lda

```
27 #c) Generar el modelo LDA
28
29 library(MASS)
30 modelo_lda <- lda(Treat~., anorexia)
31
32 modelo_lda
33
> #c) Generar el modelo LDA
>
> library(MASS)
> modelo_lda <- lda(Treat~., anorexia)
> modelo_lda
Call:
lda(Treat ~ ., data = anorexia)

Prior probabilities of groups:
      CBT      Cont      FT
0.4027778 0.3611111 0.2361111

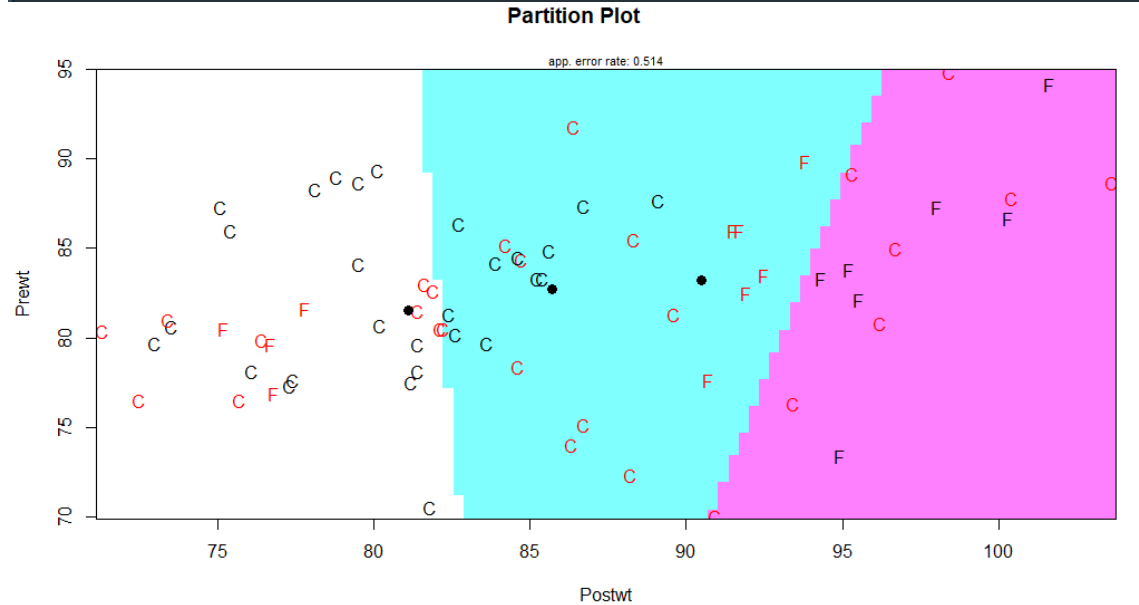
Group means:
      Prewt  Postwt
CBT  82.68966 85.69655
Cont  81.55769 81.10769
FT    83.22941 90.49412

Coefficients of linear discriminants:
      LD1      LD2
Prewt  0.01157317 -0.20152472
Postwt -0.13955521  0.03690958

Proportion of trace:
      LD1      LD2
0.9961 0.0039
> |
```

d) Gráfica de partición para identificar límites de clasificación

```
34 #d) Gráfica de partición para identificar límites de clasificación
35
36 library(klar)
37
38 partimat(Treat ~., data = anorexia, method = "lda")
39
> #d) Gráfica de partición para identificar límites de clasificación
>
> library(klar)
>
> partimat(Treat ~., data = anorexia, method = "lda")
>
```



e) Predicción de Prewt = 70, Postwt = 90

```
40 #e) Predicción de Prewt = 70, Postwt = 90
41
42 nuevas_observaciones <- data.frame(Prewt = 70, Postwt = 90)
43 prediccion <- predict(object = modelo_lda,
44                       newdata = nuevas_observaciones,
45                       interval = "confidence",
46                       level = 0.95)
47
48 prediccion$class
49
50 # Probabilidad de pertenencia a cada clase:
51
52 prediccion$posterior
53
54
```

```

> #d) Gráfica de partición para identificar límites de clasificación
>
> library(klar)
>
> partimat(Treat ~., data = anorexia, method = "lda")
>
>
> #e) Predicción de Prewt = 70, Postwt = 90
>
> nuevas_observaciones <- data.frame(Prewt = 70, Postwt = 90)
> prediccion <- predict(object = modelo_lda,
+                       newdata = nuevas_observaciones,
+                       interval = "confidence",
+                       level = 0.95)
>
> prediccion$class
[1] CBT
Levels: CBT Cont FT
>
> # Probabilidad de pertenencia a cada clase:
>
> prediccion$posterior
      CBT      Cont      FT
1 0.4053162 0.2186976 0.3759861
>

```

f) Predicción de Prewt = 80, Postwt = 100

```

55 #f) Predicción de Prewt = 80, Postwt = 100
56
57 nuevas_observaciones <- data.frame(Prewt = 80, Postwt = 100)
58
59 prediccion <- predict(object = modelo_lda,
60                       newdata = nuevas_observaciones,
61                       interval = "confidence")
62
63 prediccion$class
64
65 # Probabilidad de pertenencia a cada clase:
66
67 prediccion$posterior
68

```

```

> #d) Gráfica de partición para identificar límites de clasificación
> library(klar)
>
> partimat(Treat ~., data = anorexia, method = "lda")
>
>
> #e) Predicción de Prewt = 70, Postwt = 90
>
> nuevas_observaciones <- data.frame(Prewt = 70, Postwt = 90)
> prediccion <- predict(object = modelo_lda,
+                       newdata = nuevas_observaciones,
+                       interval = "confidence",
+                       level = 0.95)
>
> prediccion$class
[1] CBT
Levels: CBT Cont FT
>
> # Probabilidad de pertenencia a cada clase:
>
> prediccion$posterior
      CBT      Cont      FT
1 0.4053162 0.2186976 0.3759861
>
>
> #f) Predicción de Prewt = 80, Postwt = 100
>
> nuevas_observaciones <- data.frame(Prewt = 80, Postwt = 100)
>
> prediccion <- predict(object = modelo_lda,
+                       newdata = nuevas_observaciones,
+                       interval = "confidence")
>
> prediccion$class
[1] FT
Levels: CBT Cont FT
>
> # Probabilidad de pertenencia a cada clase:
>
> prediccion$posterior
      CBT      Cont      FT
1 0.3167727 0.06952407 0.6137032
>

```

f) Descripción de nuevas funciones

lda: Realiza un análisis discriminante lineal.

ggarrange: Organiza plots para compararlos.

partimat: Crea una matriz con la clasificación del método LDA.

g) Interpretación general de los resultados

La primera gráfica muestra los puntos muy dispersos.

Con el análisis discriminante lineal, tenemos las probabilidades de CBT=40.27%, Cont=36.11% y FT=23.61%.

Por parte del PREWT y POSTWT, queremos obtener un 70 y 90 respectivamente, el cual nos da una probabilidad de pertenencia de CBT=40.53%, Cont=21.86%, FT=37.59%.

Dentro de la predicción de PREWT y POSTWT, con 80 y 100 respectivamente, el cual nos da una probabilidad de pertenencia de CBT=31.67%, Cont=6.95%, FT=61.37%.