

# Analisis discriminante

Irma Eunice Martínez de la Cruz

2022-05-02

## ANALISIS DISCRIMINANTE

### 1.-Se carga la libreria MASS

```
library(MASS)
```

### 2.- Se cargan los datos “cats”

```
Z<-as.data.frame(cats)
head(Z)
```

```
##      Sex Bwt Hwt
## 1    F 2.0 7.0
## 2    F 2.0 7.4
## 3    F 2.0 9.5
## 4    F 2.1 7.2
## 5    F 2.1 7.3
## 6    F 2.1 7.6
```

### 3.- Se define la matriz de datos de “cats” y la variable

```
x<-Z[,2:3]
y<-Z[,1]
```

### 4.- Definir como n y p el numero de gatos y variables

```
n<-nrow(x)
p<-ncol(x)
```

### 5.- Se aplica el Analisis discriminante lineal (LDA)

Cross validation (cv): clasificacion optima

```
lda.cats<-lda(y~.,data=x, CV=TRUE)
```

### 6.- lda.cats\$class contiene las clasificaciones hechas por CV usando LDA.

```
lda.cats$class
```

```
##      [1] F F F F F F F F F F F F F F F F F F F F F F F F M M M F M M M
##     [38] M M M M M M M M M F F F F F F F F F F F M M M F F M M M M M M M M M
```

```
## [75] M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M
## [112] M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M M
## Levels: F M
```

## 7.-Creacion de la tabla de clasificaciones buenas y malas

```
table.cats<-table(y,lda.cats$class)
table.cats
```

```
##
## y      F  M
##   F 31 16
##   M 14 83
```

## 8.-Proporcion de errores

```
mis.cats<- n-sum(y==lda.cats$class)
mis.cats/n
```

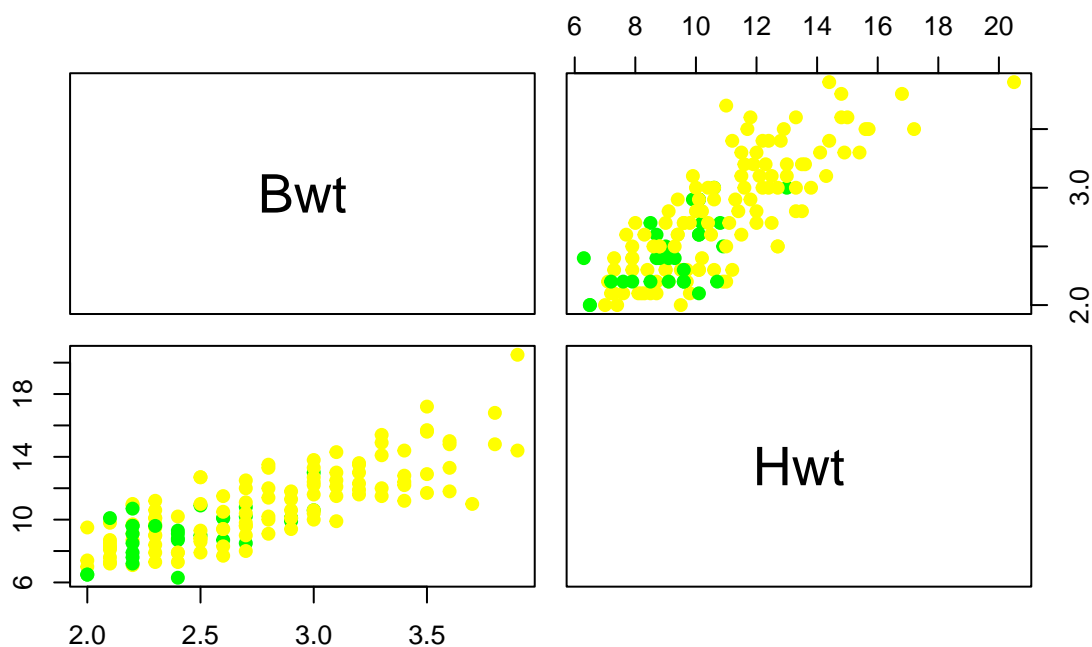
```
## [1] 0.2083333
```

## Scater plot

Buenas clasificaciones en amarillo y malas en verde

```
col.lda.cats<-c("green","yellow")[1*(y==lda.cats$class)+1]
pairs(x,main="Buena Clasificacion (Amarillo), Mala Clasificacion (verde)",
      pch=19,col=col.lda.cats)
```

## Buena Clasificacion (Amarillo), Mala Clasificacion (verde)



## Probabilidad de pertenencia a uno de los dos grupos

```
head(lda.cats$posterior)
```

```
##           F           M
## 1 0.7690302 0.2309698
## 2 0.7723092 0.2276908
## 3 0.7832209 0.2167791
## 4 0.7071594 0.2928406
## 5 0.7082562 0.2917438
## 6 0.7113342 0.2886658
```

## Grafico de probabilidades

```
plot(1:n, lda.cats$posterior[,1],
     main="Probabilidades a posterior",
     pch=20, col="cyan",
     xlab="Número de observaciones", ylab="Probabilidades")
points(1:n, lda.cats$posterior[,2],
       pch=20, col="green")
```

