MASTER EN DATA SCIENCE PARA FINANZAS

RIESGO OPERACIONAL

Practica 1

CUNEF
COLEGIO UNIVERSITARIO DE ESTUDIOS FINANCIEROS

FINANCIAL [THINKING]

Prof. Sonia de Paz Cobo

```
################    Práctica 1_fraude1     ##########################



#PAquetes que vamos a necesitar
library(moments) #Package para calculo de momentos (asimetría, curtosis, ...)
library(actuar) #Package para análisis actuarial
library(fitdistrplus) #Package para ajuste de distribuciones
library(ggplot2) #Package para visulación



x<-read.csv("pract1.csv")
names(x)[1]=paste("Fraudes")

################## Información sobre los datos

head(x$Fraudes,10) #vemos los 10 primeros elementos
summary(x$Fraudes)
table(x$Fraudes) #tabla de frecuencias

skewness(x$Fraudes)  #coef. de asimetria - Asimetria positiva o la dcha
kurtosis(x$Fraudes)  #coef. de curtosis  - Leprtocúrtica o mas apuntada Normal
```

```
> summary(x$Fraudes)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.0000  0.0000  0.0000  0.4862  1.0000  5.0000

> table(x$Fraudes) #tabla de frecuencias

   0    1    2    3    4    5
6131 3022  715  120   10    2
> skewness(x$Fraudes)  #coef. de asimetria - Asimetria positiva o la dcha
[1] 1.412611
> kurtosis(x$Fraudes)  #coef. de curtosis  - Leprtocúrtica o mas apuntada Normal
[1] 4.954833
>
```

```
> names(x)[1]=paste("Fraudes")
> mean(x$Fraudes)
[1] 0.4862
> var(x$Fraudes)
[1] 0.4808576
> median(x$Fraudes)
[1] 0
> quantile(x$Fraudes,probs=c(0.05, 0.95))
 5% 95%
  0   2
> quantile(x$Fraudes,seq(0,1, 0.20))
  0%  20%  40%  60%  80% 100%
   0    0    0    0    1    5
> quantile(x$Fraudes,seq(0.9,1, 0.01))
 90%  91%  92%  93%  94%  95%  96%  97%  98%  99% 100%
   1    1    2    2    2    2    2    2    2    3    5
```

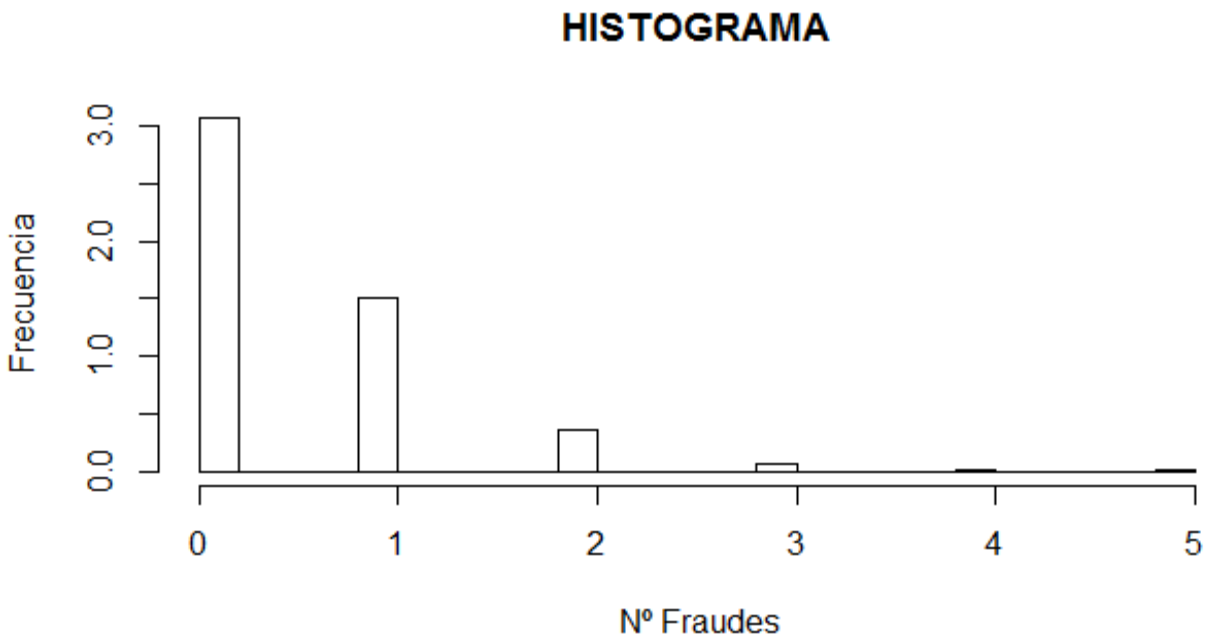
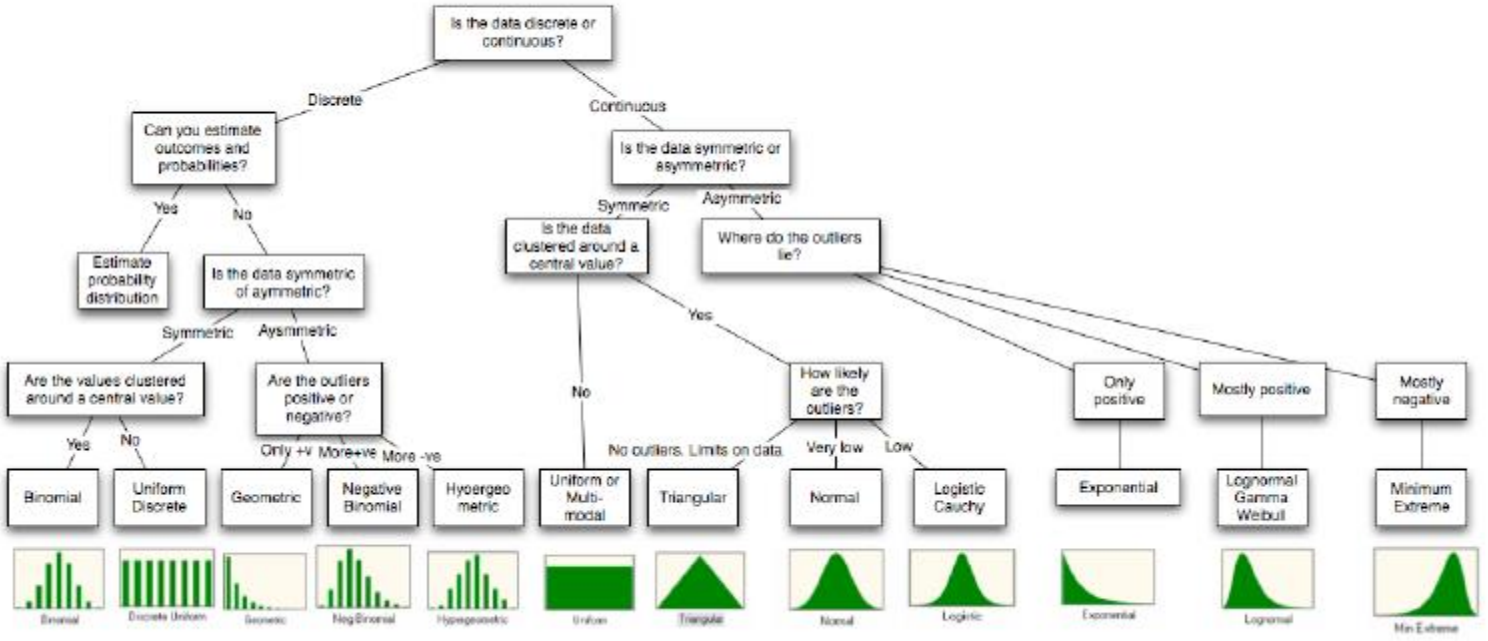


Figure 6A.15: Distributional Choices

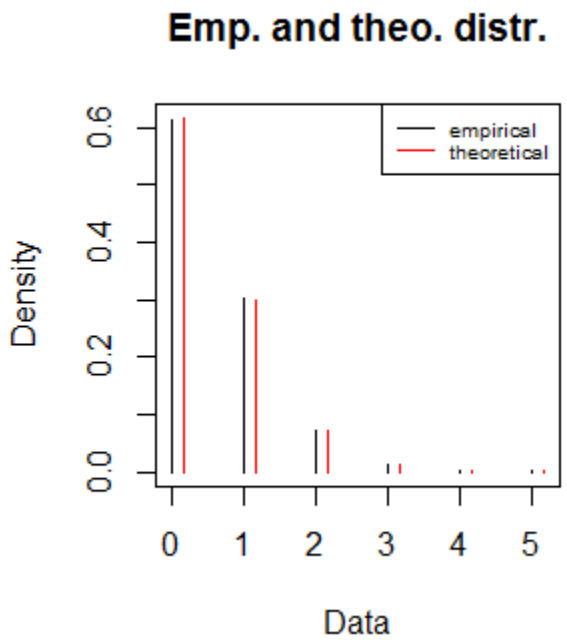


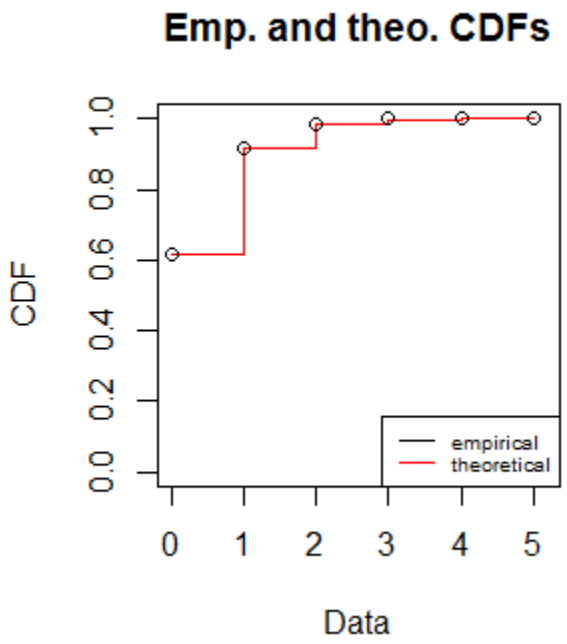
```
#################################################################

#Ajuste máxima verosimilitud de distribuciones univariables


fpois=fitdist(x$Fraudes, "pois")
fpois
plot(fpois)
```

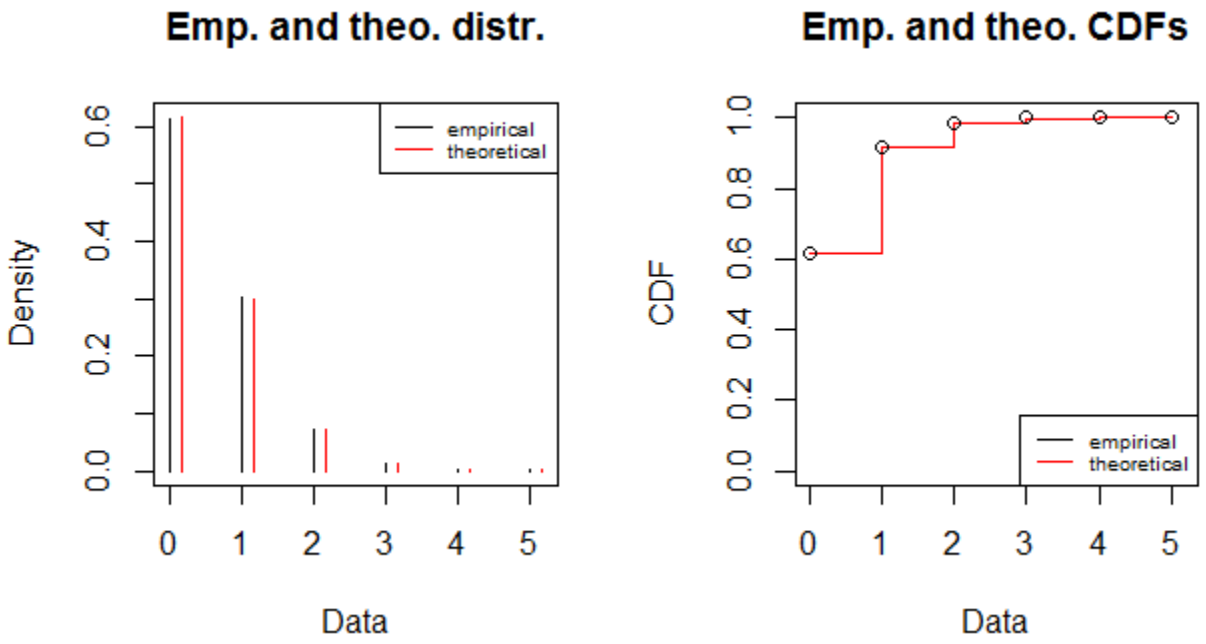
```
Fitting of the distribution ' pois ' by maximum likelihood
Parameters:
       estimate  Std. Error
lambda   0.4862 0.006972775
```





```
fnbinom=fitdist(x$Fraudes, "nbinom")
fnbinom
plot(fnbinom)
```

```
> Fitting of the distribution ' nbinom ' by maximum likelihood
 Parameters:
          estimate  Std. Error
 size 1.714778e+05         NaN
 mu   4.861763e-01 0.006972445
```



Se define en función a los valores que tiene la muestra. Determinante para calcular el estadístico Chi cuadrado.

```
#Ajuste por bondad del ajuste

gofstat(list(fpois, fnbinom), chisqbreaks=c(0:4, 9), discrete=TRUE,
        fitnames=c("Poisson", "Binomial Negativa"))
```

```
Chi-squared statistic:  2.09339 2.093969
Degree of freedom of the Chi-squared distribution:  5 4
Chi-squared p-value:  0.8360777 0.7184805
   the p-value may be wrong with some theoretical counts < 5
Chi-squared table:
     obscounts theo Poisson theo Binomial Negativa
<= 0      6131 6.149588e+03           6.149738e+03
<= 1      3022 2.989930e+03           2.989848e+03
<= 2       715 7.268519e+02           7.267989e+02
<= 3       120 1.177985e+02           1.177852e+02
<= 4        10 1.431840e+01           1.431630e+01
<= 9         2 1.513485e+00           1.513217e+00
> 9           0 1.308547e-06           1.308248e-06
```

Pruebas para la selección del módelo

```
Goodness-of-fit criteria
                                 Poisson Binomial Negativa
Akaike's Information Criterion 18242.25           18244.25
Bayesian Information Criterion 18249.46           18258.67
```