

Actividad Integradora 1

Facundo Colasurdo Caldironi

2024-08-20

Punto 1. Análisis descriptivo de la variable

Analiza una de las siguientes variables en cuanto a sus datos atípicos y normalidad. La variable que te corresponde analizar te será asignada por tu profesora al inicio de la actividad:

6. Sodio

Datos Atípicos

Graficar el diagrama de caja y bigote. Calcula las principales medidas que te ayuden a identificar datos atípicos (utilizar summary te puede abreviar el cálculo): Cuartil 1, Cuartil 3, Media, Cuartil 3, Rango intercuartílico y Desviación estándar. Identifica la cota de 1.5 rangos intercuartílicos para datos atípicos, ¿hay datos atípicos de acuerdo con este criterio? ¿cuántos son? Identifica la cota de 3 desviaciones estándar alrededor de la media, ¿hay datos atípicos de acuerdo con este criterio? ¿cuántos son? Identifica la cota de 3 rangos intercuartílicos para datos extremos, ¿hay datos extremos de acuerdo con este criterio? ¿cuántos son? Interpreta los resultados obtenidos y argumenta sobre el comportamiento de los datos atípicos y extremos en la variable seleccionada.

```
M=read.csv("file:///Users/facundocolasurdocaldironi/Downloads/food_data_g.csv") #leer la base de datos
M$Sodium
```

```
## [1] 0.016 0.300 0.000 0.017 0.046 0.100 0.200 0.200 0.200 0.008 0.043 0.700
## [13] 0.057 0.076 0.200 0.000 0.200 1.400 0.300 0.100 0.200 0.700 0.300 1.000
## [25] 1.500 0.800 1.800 0.100 0.100 0.800 0.200 0.800 0.900 0.200 2.000 0.900
## [37] 0.300 0.800 1.200 1.500 0.011 0.003 0.052 0.099 0.085 0.028 0.400 1.200
## [49] 0.098 0.016 0.300 0.004 0.030 0.200 0.000 0.300 0.500 1.000 0.400 0.200
## [61] 1.300 1.000 1.200 1.300 1.300 0.200 2.400 0.400 0.900 0.500 0.700 0.700
## [73] 0.300 0.056 0.600 0.500 0.100 0.200 1.300 0.400 0.400 0.300 1.100 1.500
## [85] 0.600 0.600 1.300 0.003 0.500 0.200 0.800 0.200 0.013 0.600 0.800 0.700
## [97] 1.100 1.200 0.100 0.043 1.400 0.400 0.700 1.100 0.600 0.900 1.000 1.200
## [109] 0.900 0.600 0.500 1.700 1.400 0.016 0.200 0.800 0.200 1.000 0.800 1.300
## [121] 0.500 0.200 0.900 0.400 1.400 0.600 0.500 1.400 0.300 0.900 0.300 1.000
## [133] 0.400 0.000 0.800 0.200 0.100 0.900 0.500 0.700 0.300 1.700 0.400 0.800
## [145] 0.600 0.300 0.500 0.012 0.081 0.600 0.200 1.000 1.200 2.300 1.000 0.400
## [157] 1.100 1.900 0.200 0.300 0.200 0.200 0.100 0.081 0.036 1.000 0.061 0.000
## [169] 0.700 0.700 0.400 0.700 0.500 0.900 1.000 0.068 0.700 1.000 0.004 0.075
## [181] 0.500 1.800 0.500 0.900 2.100 1.900 1.200 1.300 1.300 1.600 0.800 0.800
## [193] 0.800 2.500 1.000 1.700 1.700 1.600 1.800 1.800 1.700 0.900 0.900 0.500
## [205] 1.700 1.100 2.200 0.900 0.900 1.300 2.100 0.900 1.100 0.015 0.900 2.100
## [217] 0.900 1.700 1.500 0.900 0.400 1.700 0.900 1.300 0.600 0.500 0.900 1.300
```

```
## [229] 1.800 0.300 0.400 1.300 1.000 1.600 1.100 1.800 1.800 1.700 2.100 1.000
## [241] 1.700 1.600 1.700 1.700 0.900 0.900 0.400 2.300 0.300 1.800 1.100 1.300
## [253] 1.700 0.600 0.800 0.000 0.000 0.800 1.000 0.200 0.400 0.500 0.031 0.200
## [265] 1.100 0.200 0.300 0.500 0.300 0.000 0.014 0.800 0.800 0.300 0.300 0.700
## [277] 1.000 0.800 0.015 0.700 0.200 0.600 0.500 0.500 0.000 1.000 0.040 0.800
## [289] 0.500 0.900 0.400 0.200 0.800 0.900 0.009 0.000 0.700 0.064 0.600 0.900
## [301] 0.700 1.200 0.800 0.800 0.700 0.900 1.100 0.800 1.000 1.000 0.700 1.000
## [313] 0.400 1.200 0.043 0.600 1.000 1.500 1.000 2.200 0.600 0.000 0.000 0.800
## [325] 0.500 0.700 1.000 1.600 0.500 1.300 1.400 1.300 1.300 1.100 1.200 0.500
## [337] 0.700 0.700 0.800 0.800 0.700 0.600 0.600 0.800 0.000 0.200 0.400 0.100
## [349] 0.300 0.019 0.100 0.011 0.100 0.300 0.066 0.055 0.012 0.200 0.100 0.009
## [361] 0.018 0.100 0.200 0.200 0.043 0.068 0.027 0.100 0.041 0.600 5.600 0.098
## [373] 0.300 0.100 0.079 0.042 0.050 0.100 1.300 0.400 0.200 0.200 0.100 0.100
## [385] 0.500 0.022 0.069 0.037 0.060 6.100 0.064 0.011 0.100 0.300 0.046 0.300
## [397] 0.020 0.061 0.200 0.028 0.100 0.078 0.200 0.200 0.100 0.200 0.082 0.200
## [409] 0.800 0.060 0.052 0.084 0.200 0.300 0.100 0.300 0.200 0.200 0.600 0.042
## [421] 0.200 0.001 0.029 0.072 0.058 0.300 0.300 0.019 0.500 0.022 0.200 0.047
## [433] 0.059 0.051 0.003 0.100 0.200 0.058 0.200 0.043 0.086 0.013 0.076 0.100
## [445] 0.007 0.063 0.200 0.100 0.055 0.500 0.098 0.100 0.200 0.100 0.054 0.004
## [457] 0.100 0.011 0.300 0.069 0.200 0.005 0.100 0.042 0.200 0.300 0.200 0.016
## [469] 0.200 0.049 0.100 0.600 0.047 0.014 0.019 0.022 0.100 0.024 0.200 0.045
## [481] 0.100 0.200 0.400 0.100 0.200 0.300 0.200 0.081 0.100 0.100 0.400 0.100
## [493] 0.001 0.500 0.100 1.800 0.400 1.400 0.300 1.700 0.700 0.200 0.300 0.700
## [505] 0.000 1.100 1.400 0.200 0.012 0.068 1.400 0.060 0.200 0.059 0.300 0.200
## [517] 0.100 0.066 0.038 0.700 0.064 0.400 0.400 0.600 0.500 0.500 0.200 0.083
## [529] 0.300 0.200 0.200 0.090 0.057 0.600 0.200 0.079 0.073 0.400 0.200 0.700
## [541] 0.015 1.400 0.400 0.021 0.100 0.200 1.100 0.100 0.200 1.400 0.065
```

```
# Load the 'Sodium' column
Y <- M$Sodium

# Calculate the first (Q1) and third (Q3) quantiles
q1 <- quantile(Y, 0.25) # 1q
q2 <- quantile(Y, 0.50) # 3q
q3 <- quantile(Y, 0.75) # 3q
ri <- q3 - q1 #Rango intercuartílico de X

# Calcular la media y la desviación estándar
mean_Sodium = mean(Y)
sd_Sodium = sd(Y)

print(q1)
```

```
## 25%
## 0.1
```

```
print(q2)
```

```
## 50%
## 0.4
```

```
print(q3)
```

```
## 75%  
## 0.9
```

```
print(ri)
```

```
## 75%  
## 0.8
```

```
print(sd_Sodium)
```

```
## [1] 0.6361261
```

```
par(mfrow = c(2, 1)) #Matriz de gráficos de 2x1  
boxplot(Y, horizontal = TRUE, ylim = c(min(Y), max(Y)))  
abline(v = q3 + 1.5 * ri, col = "red") #línea vertical en el límite de los datos atípicos o extremos  
Y1 <- M[M$Carbohydrates < (q3 + 1.5 * ri), "Sodium"] #En la matriz M, quitar datos más allá de 3 rangos  
summary(Y)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
## 0.0000  0.1000  0.4000  0.5732  0.9000  6.1000
```

```
summary(Y1)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
## 0.0000  0.0495  0.1000  0.2887  0.3000  6.1000
```

```
# Cotas para 1.5 rangos intercuartílicos  
CotaUNO = q1 - 1.5 * ri  
print(CotaUNO)
```

```
## 25%  
## -1.1
```

```
CotaDOS = q3 + 1.5 * ri  
print(CotaDOS)
```

```
## 75%  
## 2.1
```

```
# Cotas para 3 desviaciones estándar  
Inferior_sd = mean_Sodium - 3 * sd_Sodium  
Superior_sd = mean_Sodium + 3 * sd_Sodium  
  
print(Inferior_sd)
```

```
## [1] -1.335173
```

```
print(Superior_sd)
```

```
## [1] 2.481584
```

```
# Cotas para 3 rangos intercuartílicos
```

```
CotaTRES= q1 - 3 * ri
```

```
print(CotaTRES)
```

```
## 25%
```

```
## -2.3
```

```
CotaCUATRO = q3 + 3 * ri
```

```
print(CotaCUATRO)
```

```
## 75%
```

```
## 3.3
```

```
# Identificar datos atípicos según 1.5 rangos intercuartílicos
```

```
Atipicos_Q1 = M[M$Sodium < CotaUNO | M$Sodium > CotaDOS, ]
```

```
print(Atipicos_Q1)
```

```
##      X Unnamed..0      food Caloric.Value  Fat Saturated.Fats
## 67  66          66      kung pao chicken      779 42.2          8.2
## 154 153          153  general tsos chicken      1578 87.5         14.8
## 194 193          193    black bean soup        234  3.4          0.9
## 207 206          206    beef mushroom soup      186  7.3          3.7
## 248 247          247      escarole soup          61  4.0          1.2
## 320 319          319 sandwich with cold cuts      833 39.3         13.9
## 371 370          370    jellyfish dried          21  0.8          0.2
## 390 389          389    salt mackerel           415 34.1          9.7
##      Monounsaturated.Fats Polyunsaturated.Fats Carbohydrates Sugars Protein
## 67      13.1          18.2          41.5      18.3      59.0
## 154      20.8          40.1          128.3     62.1     69.0
## 194       1.2           1.1          39.6       6.4     12.4
## 207       3.1           0.3          15.9       0.0     14.0
## 248       1.8           0.8           4.0       0.0       3.4
## 320      14.5           7.5          79.9     12.4     41.1
## 371       0.1           0.3           0.0       0.0       3.2
## 390      11.3           8.4           0.0       0.0     25.2
##      Dietary.Fiber Cholesterol Sodium Water Vitamin.A Vitamin.B1 Vitamin.B11
## 67       9.1          157.0      2.4 451.7      0.000      0.200      0.000
## 154       4.8          283.6      2.3 241.6      0.081      0.100      0.049
## 194      17.5           0.0      2.5 193.6      0.000      0.100      0.200
## 207       0.6          15.3      2.2 260.8      0.000      0.071      0.044
## 248       0.0           5.5      2.3 535.8      0.000      0.200      0.058
## 320       4.7          105.6      2.2 221.0      0.500      0.004      1.300
## 371       0.0           2.9      5.6  39.4      0.024      0.082      0.033
## 390       0.0          129.2      6.1  58.5      0.031      0.050      0.061
##      Vitamin.B12 Vitamin.B2 Vitamin.B3 Vitamin.B5 Vitamin.B6 Vitamin.C Vitamin.D
## 67      0.034      0.300      16.7      3.0      1.500      42.9      0.000
## 154      0.080      0.600      15.3      3.3      1.100       8.6      0.015
```

```
## 194      0.000      0.100      1.1      0.4      0.200      0.5      0.000
## 207      0.035      0.200      2.7      0.6      0.100      0.0      0.000
## 248      0.060      0.100      5.1      0.4      0.500      10.0     0.000
## 320      0.400      0.051      1.0      16.9     1.500      0.8      29.300
## 371      0.000      0.050      0.1      0.0      0.054      0.0      0.000
## 390      0.065      0.300      4.5      0.0      0.600      0.0      0.030
##      Vitamin.E Vitamin.K Calcium Copper Iron Magnesium Manganese Phosphorus
## 67      6.200      0.012 120.800 0.40 4.6      145.0      1.5      567.8
## 154     6.500      0.200 64.200 0.20 6.2      96.3       0.4      679.5
## 194     0.900      0.035 90.000 0.80 3.9      84.8       1.3      192.8
## 207     0.000      0.000 12.200 0.60 2.1      21.4       0.9      88.5
## 248     0.000      0.000 71.900 0.80 1.7      11.1       2.8      177.0
## 320     0.012      1.600 0.023 668.60 0.5      7.1       78.2      0.9
## 371     0.050      0.059 1.200 0.08 1.3      1.2        0.0      11.6
## 390     3.200      0.043 89.800 0.10 1.9      81.6       0.0      345.4
##      Potassium Selenium Zinc Nutrition.Density
## 67      1316.7      0.097 4.500      320.100
## 154     1075.4      0.065 7.000      368.673
## 194     642.5      0.000 2.800      167.300
## 207     384.3      0.000 3.400      52.100
## 248     591.7      0.000 5.000      95.000
## 320     437.9 1102.600 0.065      166.871
## 371      1.7      0.014 0.200      6.517
## 390     707.2      0.082 1.500      151.025
```

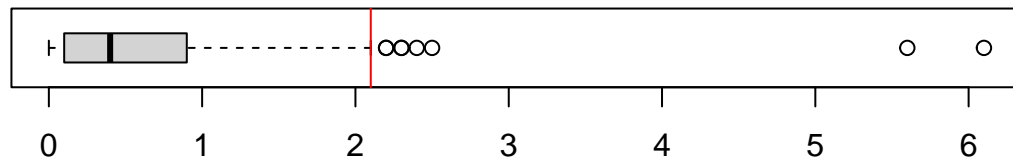
```
# Identificar datos atípicos según 3 rangos intercuartílicos
Atipicos_Q2 = M[M$Sodium < CotaTRES | M$Sodium > CotaCUATRO, ]
print(Atipicos_Q2)
```

```
##      X Unnamed..0      food Caloric.Value Fat Saturated.Fats
## 371 370      370 jellyfish dried      21 0.8      0.2
## 390 389      389 salt mackerel      415 34.1      9.7
##      Monounsaturated.Fats Polyunsaturated.Fats Carbohydrates Sugars Protein
## 371      0.1      0.3      0      0      3.2
## 390     11.3      8.4      0      0     25.2
##      Dietary.Fiber Cholesterol Sodium Water Vitamin.A Vitamin.B1 Vitamin.B11
## 371      0      2.9      5.6 39.4      0.024      0.082      0.033
## 390      0     129.2      6.1 58.5      0.031      0.050      0.061
##      Vitamin.B12 Vitamin.B2 Vitamin.B3 Vitamin.B5 Vitamin.B6 Vitamin.C Vitamin.D
## 371      0.000      0.05      0.1      0      0.054      0      0.00
## 390      0.065      0.30      4.5      0      0.600      0      0.03
##      Vitamin.E Vitamin.K Calcium Copper Iron Magnesium Manganese Phosphorus
## 371      0.05      0.059      1.2 0.08 1.3      1.2      0      11.6
## 390      3.20      0.043      89.8 0.10 1.9      81.6      0      345.4
##      Potassium Selenium Zinc Nutrition.Density
## 371      1.7      0.014 0.2      6.517
## 390     707.2      0.082 1.5      151.025
```

```
# Identificar datos atípicos según 3 desviaciones estándar
Atipicos_Q3 = M[M$Sodium < Inferior_sd | M$Sodium > Superior_sd, ]
print(Atipicos_Q3)
```

```
##      X Unnamed..0      food Caloric.Value Fat Saturated.Fats
```

## 194	193	193 black bean soup	234	3.4	0.9		
## 371	370	370 jellyfish dried	21	0.8	0.2		
## 390	389	389 salt mackerel	415	34.1	9.7		
##		Monounsaturated.Fats	Polyunsaturated.Fats	Carbohydrates	Sugars	Protein	
## 194		1.2	1.1	39.6	6.4	12.4	
## 371		0.1	0.3	0.0	0.0	3.2	
## 390		11.3	8.4	0.0	0.0	25.2	
##		Dietary.Fiber	Cholesterol	Sodium	Water	Vitamin.A	Vitamin.B1
## 194		17.5	0.0	2.5	193.6	0.000	0.100
## 371		0.0	2.9	5.6	39.4	0.024	0.082
## 390		0.0	129.2	6.1	58.5	0.031	0.050
##		Vitamin.B12	Vitamin.B2	Vitamin.B3	Vitamin.B5	Vitamin.B6	Vitamin.C
## 194		0.000	0.10	1.1	0.4	0.200	0.5
## 371		0.000	0.05	0.1	0.0	0.054	0.0
## 390		0.065	0.30	4.5	0.0	0.600	0.0
##		Vitamin.E	Vitamin.K	Calcium	Copper	Iron	Magnesium
## 194		0.90	0.035	90.0	0.80	3.9	84.8
## 371		0.05	0.059	1.2	0.08	1.3	1.2
## 390		3.20	0.043	89.8	0.10	1.9	81.6
##		Manganese	Phosphorus				
## 194		1.3	192.8				
## 371		0.0	11.6				
## 390		0.0	345.4				
##		Potassium	Selenium	Zinc	Nutrition.Density		
## 194		642.5	0.000	2.8	167.300		
## 371		1.7	0.014	0.2	6.517		
## 390		707.2	0.082	1.5	151.025		



Interpreta los resultados obtenidos y argumenta sobre el comportamiento de los datos atípicos y extremos en la variable seleccionada

Normalidad

Realiza pruebas de normalidad univariada para la variable (utiliza las pruebas de Anderson-Darling y de Jarque Bera). No olvides incluir H0 y H1 para la prueba de normalidad. Grafica los datos y su respectivo QQPlot: `qqnorm(datos)` y `qqline(datos)` Calcula el coeficiente de sesgo y el coeficiente de curtosis Compara las medidas de media, mediana y rango medio de cada variable Realiza el gráfico de densidad empírica y teórica suponiendo normalidad en la variable. Adapta el código: `hist(datos,freq=FALSE) lines(density(datos),col="red") curve(dnorm(x,mean=mean(datos,sd=sd(datos)),from=-6, to=6, add=TRUE, col="blue",lwd=2)` Interpreta los gráficos y los resultados obtenidos en cada punto con vías a indicar si hay normalidad de los datos Comenta las características encontradas: Considera alejamientos de normalidad por simetría, curtosis Comenta si hay aparente influencia de los datos atípicos en la normalidad de los datos Emite una conclusión sobre la normalidad de los datos. Se debe argumentar en términos de los 3 puntos analizados: las pruebas de normalidad, los gráficos y las medidas.

```
# Calorías
```

```
library(nortest)
library(moments)
library(tseries)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method      from
##   as.zoo.data.frame zoo
```

```
X <- M$Sodium

ad_test = ad.test(X)
print(ad_test)
```

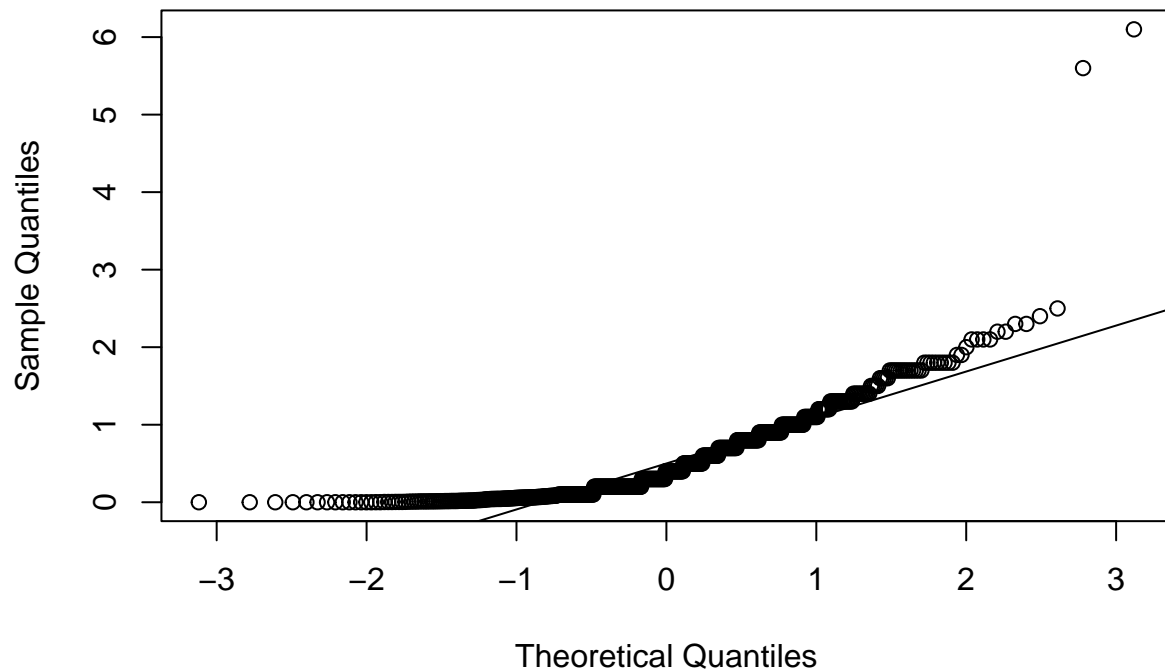
```
##
## Anderson-Darling normality test
##
## data:  X
## A = 24.827, p-value < 2.2e-16
```

```
jb_test_N = jarque.bera.test(X)
print(jb_test_N)
```

```
##
## Jarque Bera Test
##
## data:  X
## X-squared = 6834.2, df = 2, p-value < 2.2e-16
```

```
qqnorm(X)
qqline(X)
```

Normal Q-Q Plot



```
print("Curtosis")
```

```
## [1] "Curtosis"
```

```
kurtosis(X)
```

```
## [1] 19.3626
```

```
print("Sesgo")
```

```
## [1] "Sesgo"
```

```
skewness(X)
```

```
## [1] 2.735999
```

```
media <- mean(M$Sodium)
mediana <- median(M$Sodium)
rango_medio <- (min(M$Sodium) + max(M$Sodium)) / 2

cat("Media:", media, "\n")
```

```
## Media: 0.5732051
```



```
cat("Mediana:", mediana, "\n")
```

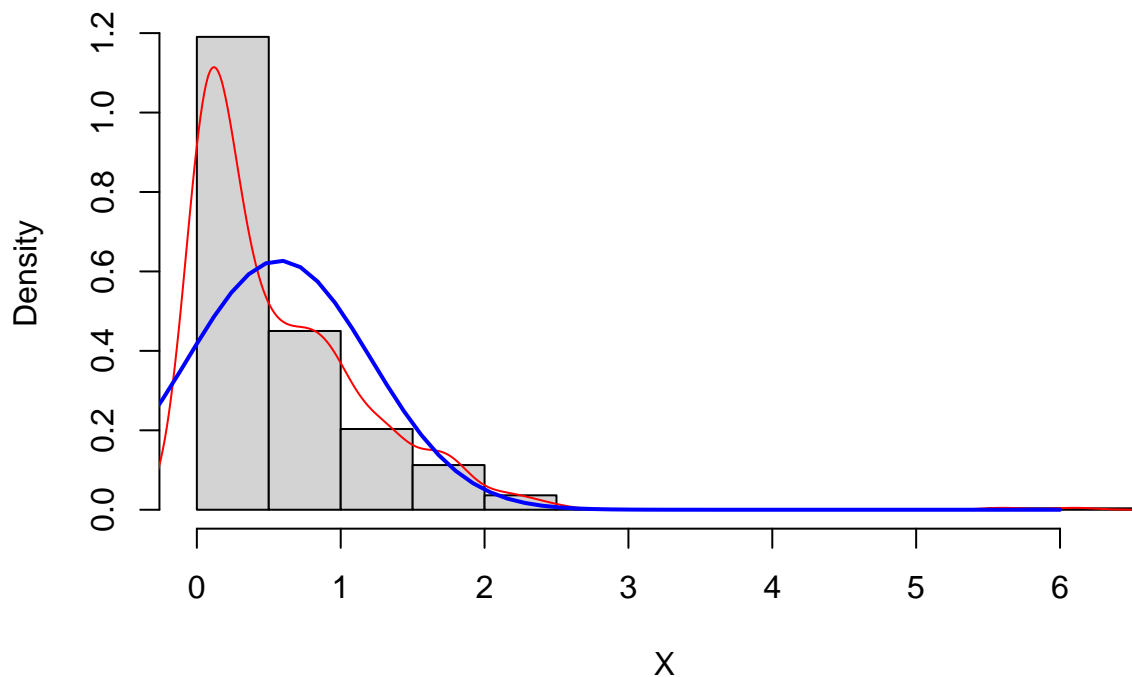
```
## Mediana: 0.4
```

```
cat("Rango Medio:", rango_medio, "\n")
```

```
## Rango Medio: 3.05
```

```
hist(X,freq=FALSE)  
lines(density(X),col="red")  
curve(dnorm(x, mean=mean(X), sd=sd(X)), from=-6, to=6, add=TRUE, col="blue", lwd=2)
```

Histogram of X



Interpreta los gráficos y los resultados obtenidos en cada punto con vías a indicar si hay normalidad de los datos

Considera alejamientos de normalidad por simetría, curtosis

Comenta si hay aparente influencia de los datos atípicos en la normalidad de los datos

Emite una conclusión sobre la normalidad de los datos. Se debe argumentar en términos de los 3 puntos analizados: las pruebas de normalidad, los gráficos y las medidas.

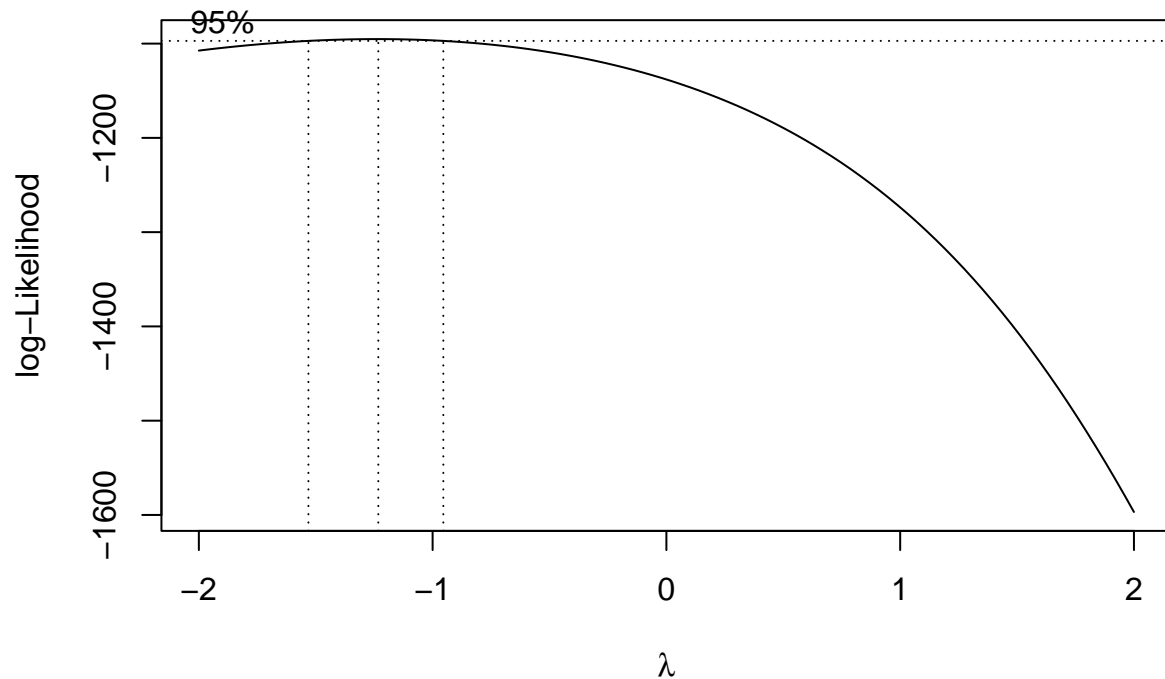
Punto 2. Transformación a normalidad

Encuentra la mejor transformación de los datos para lograr normalidad. Puedes hacer uso de la transformación Box-Cox o de Yeo Johnson o el comando powerTransform para encontrar la mejor lambda para la

transformación. Utiliza el modelo exacto y el aproximado de acuerdo con las sugerencias de Box y Cox para la transformación.

```
library(MASS)

Sodium <- M$Sodium
bc <- boxcox((Sodium + 1) ~ 1)
```



```
l <- bc$x[which.max(bc$y)]
print(l)
```

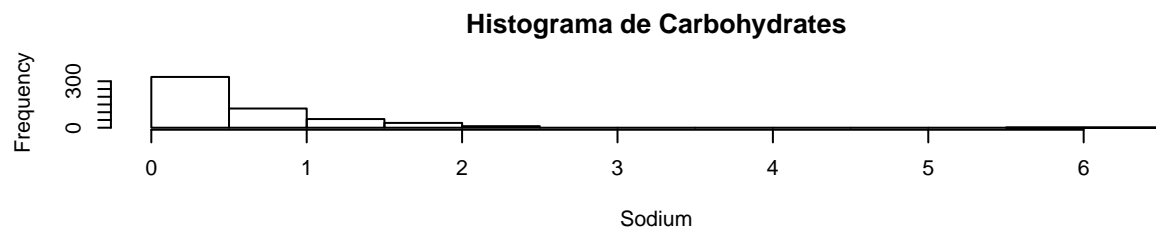
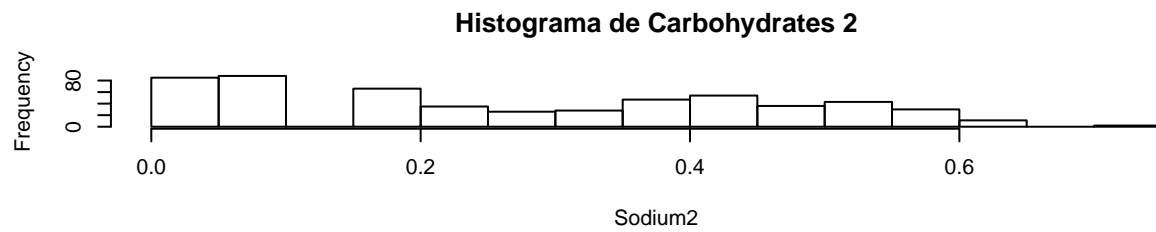
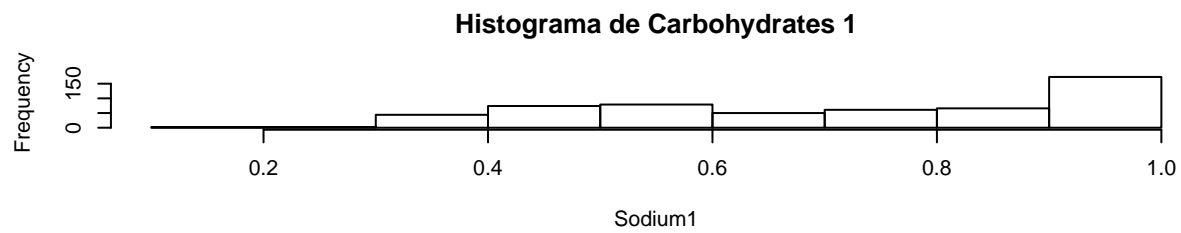
```
## [1] -1.232323
```

Escribe las ecuaciones de los modelos encontrados.

$1/X$

$\text{frac}(x + 1)^{-1.232323} - 1/^{-1.232323}$

```
Sodium1= 1/(Sodium + 1)
Sodium2=((Sodium+1)^1-1)/1
par(mfrow=c(3,1))
hist(Sodium1,col=0,main="Histograma de Carbohydrates 1")
hist(Sodium2,col=0,main="Histograma de Carbohydrates 2")
hist(Sodium,col=0,main="Histograma de Carbohydrates")
```



```
# Transformación 1
```

```
cat("Transformacion 1")
```

```
## Transformacion 1
```

```
min_trans1 <- min(Sodium1)
cat("minimo:", min_trans1, "\n")
```

```
## minimo: 0.1408451
```

```
max_trans1 <- max(Sodium1)
cat("maximo:", max_trans1, "\n")
```

```
## maximo: 1
```

```
mean_trans1 <- mean(Sodium1)
cat("media:", mean_trans1, "\n")
```

```
## media: 0.712022
```

```
median_trans1 <- median(Sodium1)
cat("mediana:", median_trans1, "\n")
```

```
## mediana: 0.7142857
```

```
q1_trans1 <- quantile(Sodium1, 0.25,)
cat("Cuartil 1:", q1_trans1, "\n")
```

```
## Cuartil 1: 0.5263158
```

```
q3_trans1 <- quantile(Sodium1, 0.75)
cat("Cuartil 3:", q3_trans1, "\n")
```

```
## Cuartil 3: 0.9090909
```

```
skewness_trans1 <- skewness(Sodium1)
cat("Sesgo:", skewness_trans1, "\n")
```

```
## Sesgo: -0.2674914
```

```
kurtosis_trans1 <- kurtosis(Sodium1)
cat("Curtosis:", kurtosis_trans1, "\n")
```

```
## Curtosis: 1.815712
```

```
# Transformación 2
cat("Transformacion 2")
```

```
## Transformacion 2
```

```
min_trans2 <- min(Sodium2)
cat("minimo:", min_trans2, "\n")
```

```
## minimo: 0
```

```
max_trans2 <- max(Sodium2)
cat("maximo:", max_trans2, "\n")
```

```
## maximo: 0.7389901
```

```
mean_trans2 <- mean(Sodium2)
cat("media:", mean_trans2, "\n")
```

```
## media: 0.2705392
```

```
median_trans2 <- median(Sodium2)
cat("mediana:", median_trans2, "\n")
```

```
## mediana: 0.2754339
```

```
q1_trans2 <- quantile(Sodium2, 0.25,)
cat("Cuartil 1:", q1_trans2, "\n")
```

```
## Cuartil 1: 0.0899258
```

```
q3_trans2 <- quantile(Sodium2, 0.75)
cat("Cuartil 3:", q3_trans2, "\n")
```

```
## Cuartil 3: 0.4435491
```

```
skewness_trans2 <- skewness(Sodium2)
cat("Sesgo:", skewness_trans2, "\n")
```

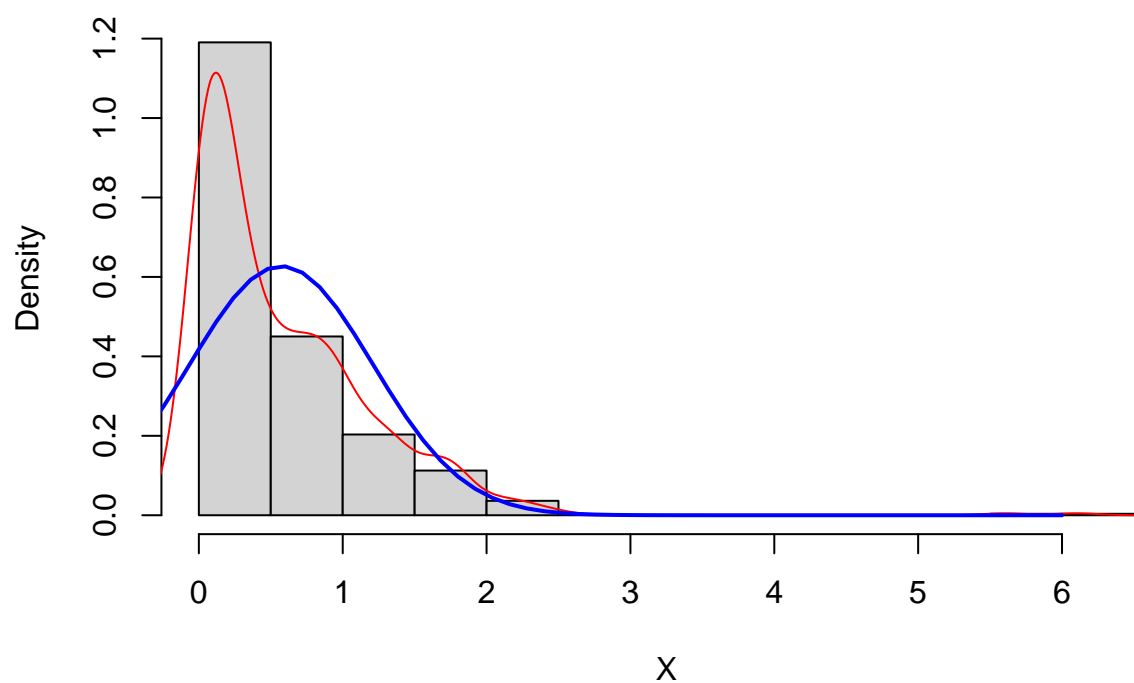
```
## Sesgo: 0.1827484
```

```
kurtosis_trans2 <- kurtosis(Sodium2)
cat("Curtosis:", kurtosis_trans2, "\n")
```

```
## Curtosis: 1.713775
```

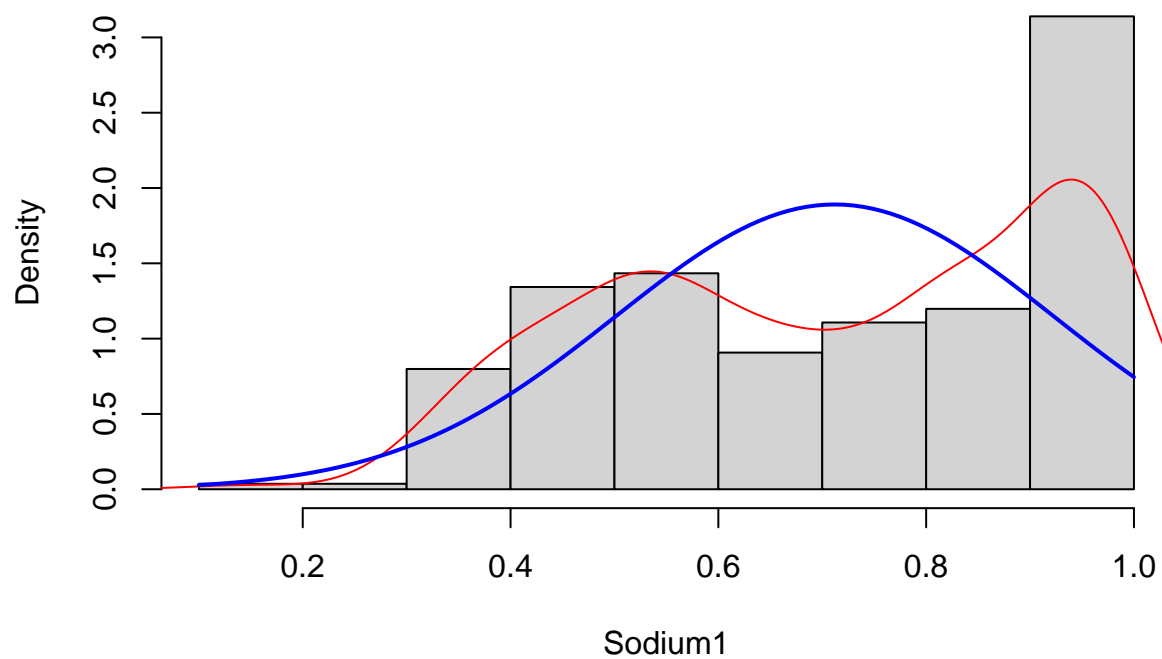
```
hist(X,freq=FALSE)
lines(density(X),col="red")
curve(dnorm(x, mean=mean(X), sd=sd(X)), from=-6, to=6, add=TRUE, col="blue", lwd=2)
```

Histogram of X



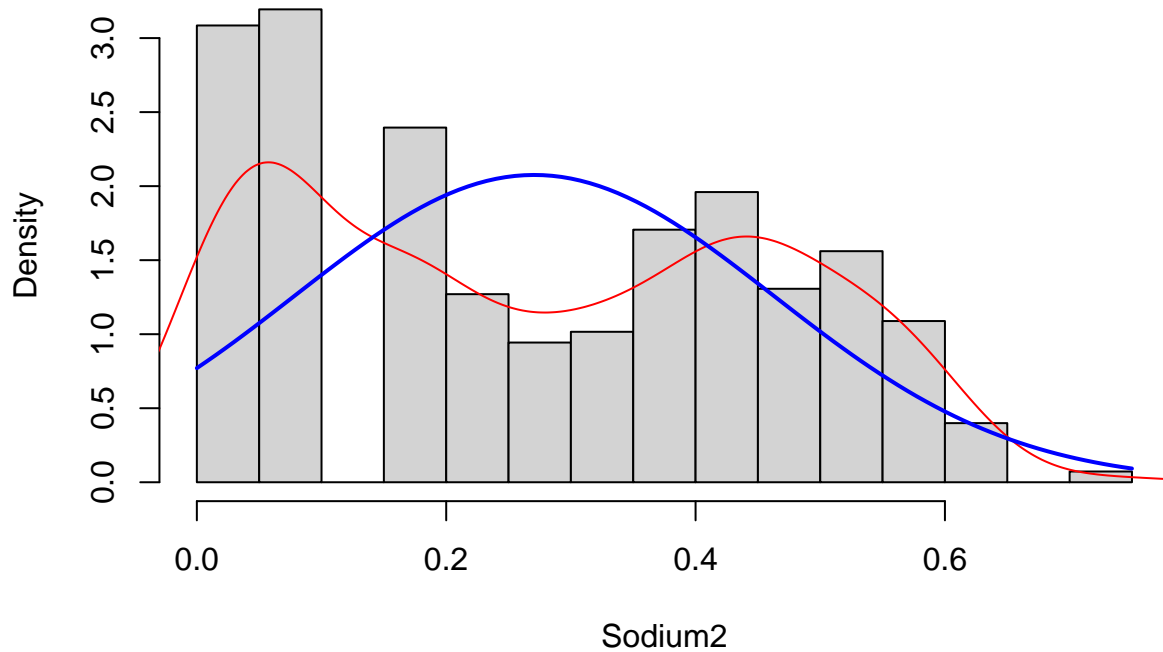
```
hist(Sodium1, freq=FALSE)
lines(density(Sodium1), col="red")
curve(dnorm(x, mean=mean(Sodium1), sd=sd(Sodium1)), add=TRUE, col="blue", lwd=2)
```

Histogram of Sodium1



```
hist(Sodium2, freq=FALSE)
lines(density(Sodium2), col="red")
curve(dnorm(x, mean=mean(Sodium2), sd=sd(Sodium2)), add=TRUE, col="blue", lwd=2)
```

Histogram of Sodium2



```
library(nortest)
library(tseries)

cat("Prueba de Anderson:\n")
```

```
## Prueba de Anderson:
```

```
ad.test(Sodium)
```

```
##
## Anderson-Darling normality test
##
## data: Sodium
## A = 24.827, p-value < 2.2e-16
```

```
ad.test(Sodium1)
```

```
##
## Anderson-Darling normality test
##
## data: Sodium1
## A = 12.714, p-value < 2.2e-16
```



```
ad.test(Sodium2)
```

```
##  
## Anderson-Darling normality test  
##  
## data: Sodium2  
## A = 12.799, p-value < 2.2e-16
```

```
cat("Prueba de Jarque:\n")
```

```
## Prueba de Jarque:
```

```
jarque.bera.test(Sodium)
```

```
##  
## Jarque Bera Test  
##  
## data: Sodium  
## X-squared = 6834.2, df = 2, p-value < 2.2e-16
```

```
jarque.bera.test(Sodium1)
```

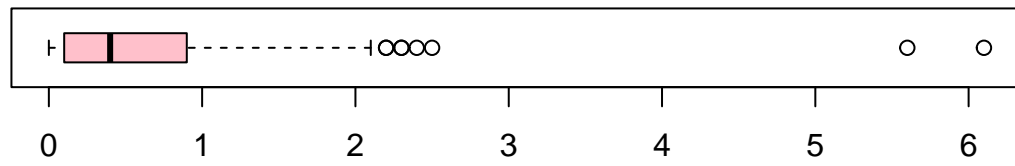
```
##  
## Jarque Bera Test  
##  
## data: Sodium1  
## X-squared = 38.771, df = 2, p-value = 3.811e-09
```

```
jarque.bera.test(Sodium2)
```

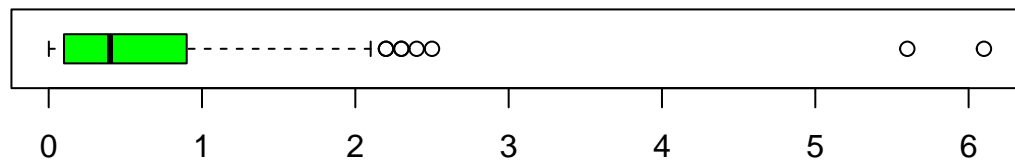
```
##  
## Jarque Bera Test  
##  
## data: Sodium2  
## X-squared = 41.049, df = 2, p-value = 1.22e-09
```

```
M2 <- subset(M, Sodium > 0)  
par(mfrow=c(2,1))  
boxplot(M$Sodium, horizontal = TRUE, col = "pink", main = "Sodium Originales")  
boxplot(M2$Sodium, horizontal = TRUE, col = "green", main = "Sodium sin Ceros")
```

Sodium Originales

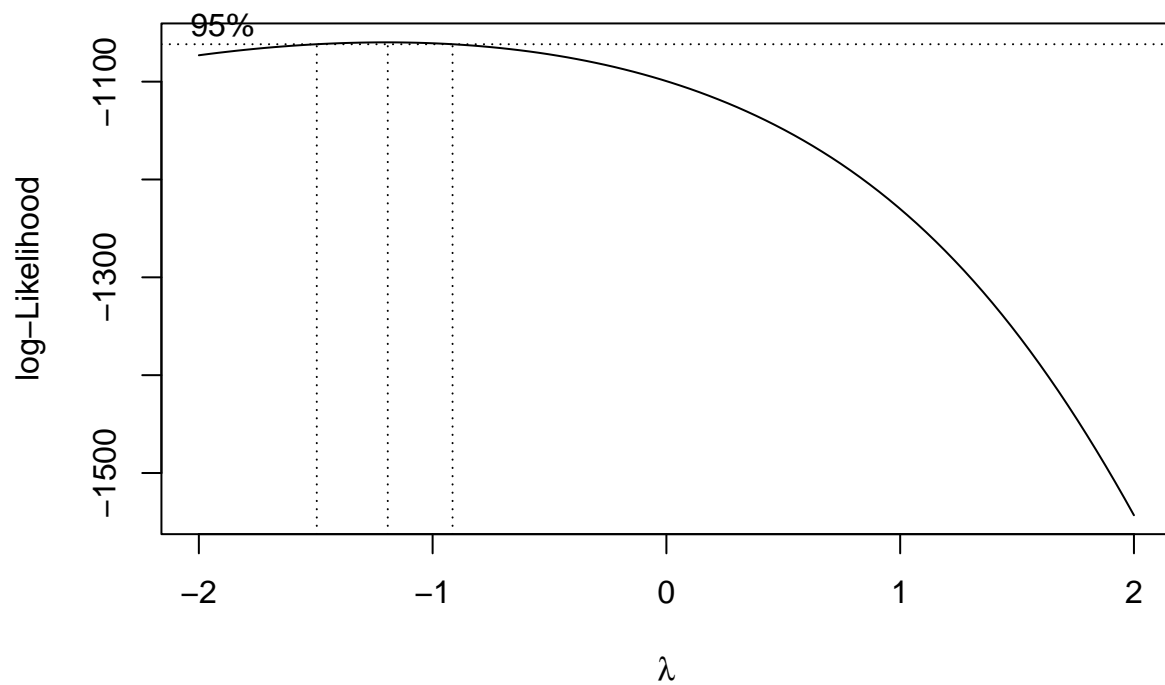


Sodium sin Ceros



```
library(MASS)

M2 <- subset(M, Sodium > 0)
M2_Sodium <- M2$Sodium
bc <- boxcox((M2_Sodium + 1) ~ 1)
```



```

1 <- bc$x[which.max(bc$y)]

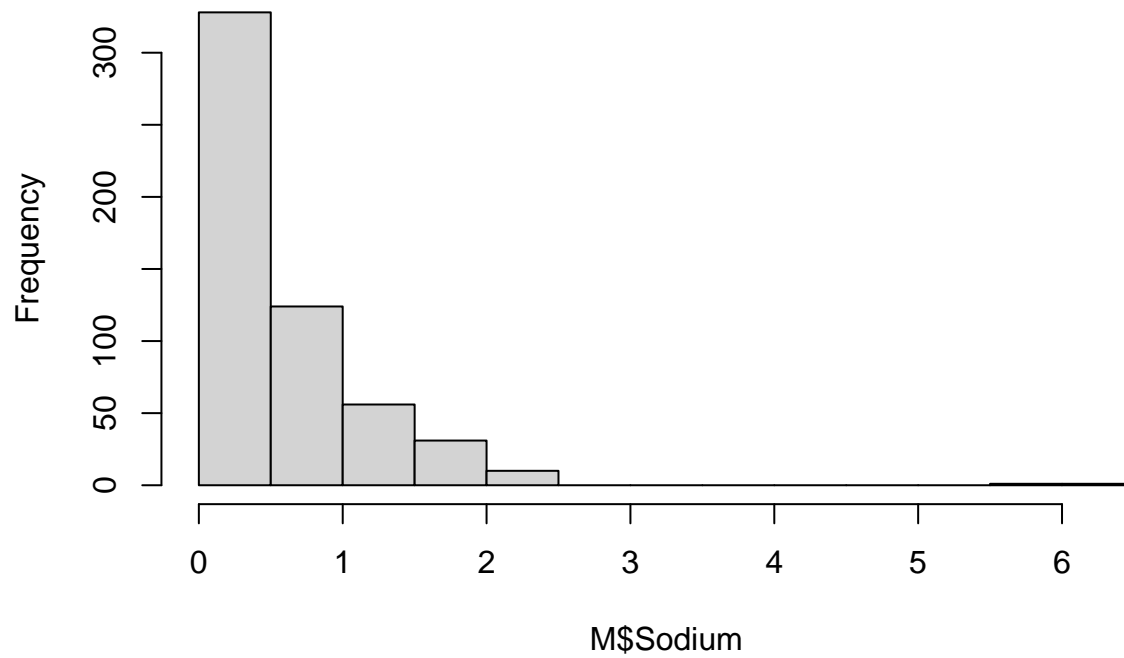
# Aplicar la transformación de Box-Cox
M2$Sodium_exa <- (M2_Sodium + 1)^1 - 1 / 1

# Aplicar la transformación aproximada 1/x
M2$Sodium_aprox <- 1 / M2_Sodium

# Histograma de los datos originales
hist(M$Sodium)

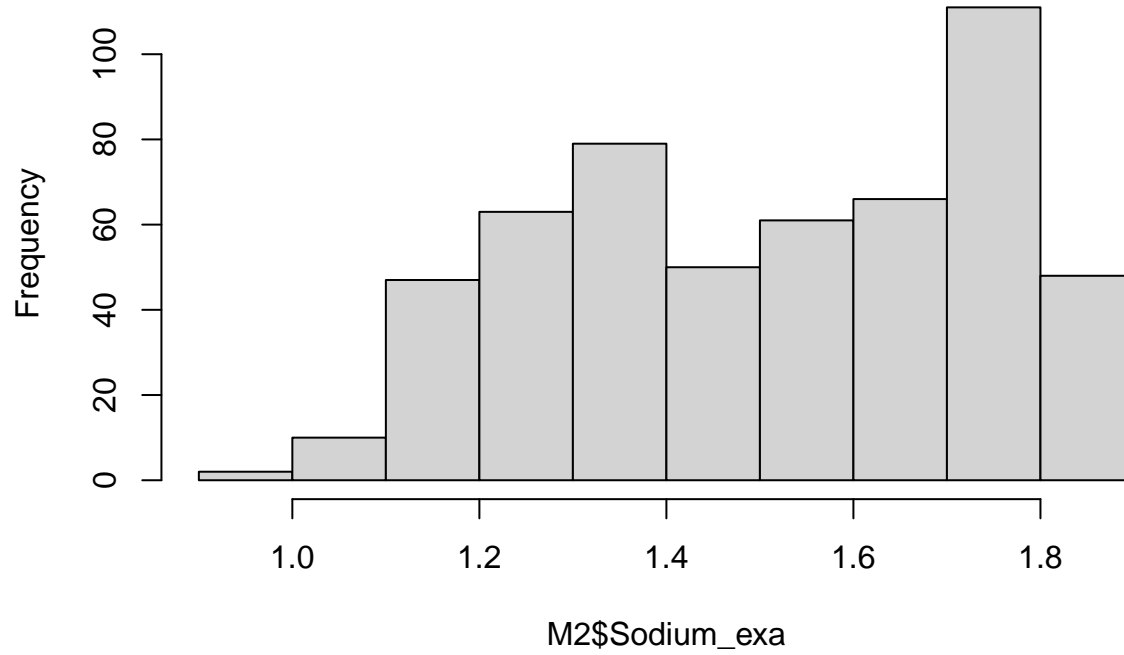
```

Histogram of M\$Sodium



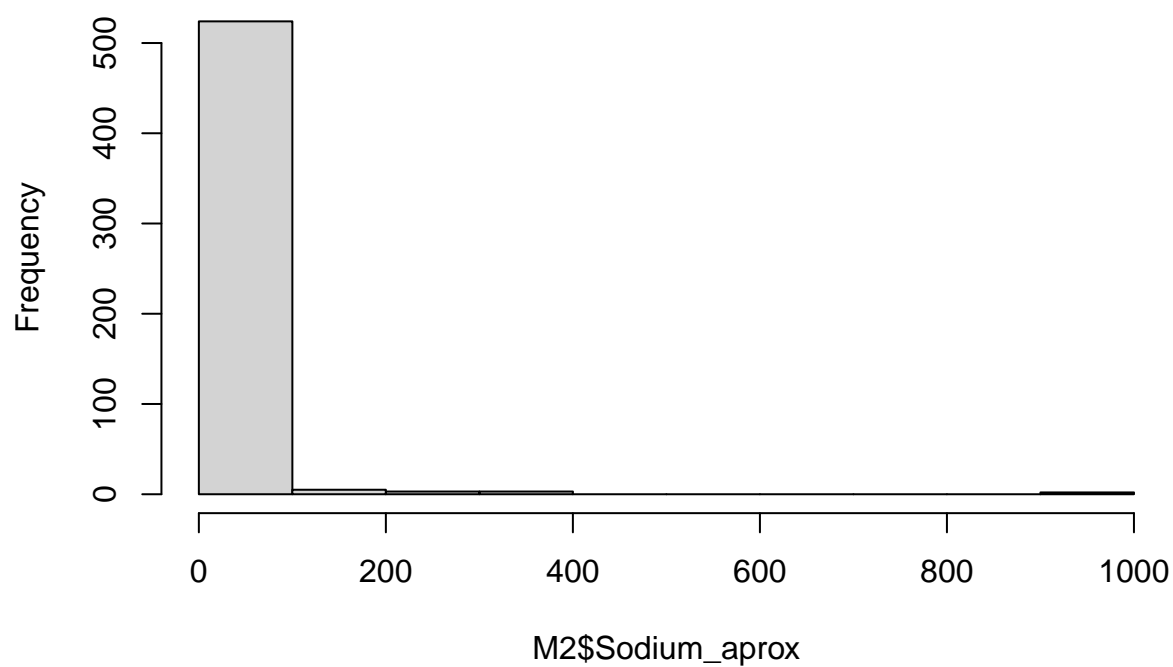
```
# Histograma de la transformación exacta (Box-Cox)  
hist(M2$Sodium_exa)
```

Histogram of M2\$Sodium_exa



```
# Histograma de la transformación aproximada (1/x)  
hist(M2$Sodium_aprox)
```

Histogram of M2\$Sodium_aprox



```
Sodium_original <- M$Sodium  
Sodium_original2 <- M2$Sodium
```

```
min(Sodium_original)
```

```
## [1] 0
```

```
max(Sodium_original)
```

```
## [1] 6.1
```

```
mean(Sodium_original)
```

```
## [1] 0.5732051
```

```
median(Sodium_original)
```

```
## [1] 0.4
```

```
quantile(Sodium_original, 0.25)
```

```
## 25%
```

```
## 0.1
```

```
quantile(Sodium_original, 0.75)
```

```
## 75%  
## 0.9
```

```
skewness(Sodium_original)
```

```
## [1] 2.735999
```

```
kurtosis(Sodium_original)
```

```
## [1] 19.3626
```

```
min(Sodium_original2)
```

```
## [1] 0.001
```

```
max(Sodium_original2)
```

```
## [1] 6.1
```

```
mean(Sodium_original2)
```

```
## [1] 0.588149
```

```
median(Sodium_original2)
```

```
## [1] 0.4
```

```
quantile(Sodium_original2, 0.25)
```

```
## 25%  
## 0.1
```

```
quantile(Sodium_original2, 0.75)
```

```
## 75%  
## 0.9
```

```
skewness(Sodium_original2)
```

```
## [1] 2.737874
```

```
kurtosis(Sodium_original2)
```

```
## [1] 19.41985
```