Métodos Multivariados: Tarea 4

Aldo Carmona, Diego Arellano, Mateo De La Roche, Victor Contreras

Ejercicio 1

1. Si dos variables X y Y tienen covarianza $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, entonces mostrar que si $c \neq 0$ entonces la primera componente principal está dada por:

$$\sqrt{\frac{c^2}{c^2 + (V_1 - d)^2}} X + \frac{c}{|c|} \sqrt{\frac{(V_i - d)^2}{c^2 + (V_1 - d)^2}} Y,$$

donde V_1 es la varianza explicada por la primera componente principal. ¿Cuál es el valor de V_1?

Consideremos la matriz de covarianza S para dos variables X y Y:

$$S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Para encontrar los valores propios λ, resolvemos la ecuación característica:

$$\det(S - \lambda I) = 0.$$

Desarrollando el determinante, obtenemos:

$$\begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc) = 0.$$

Las soluciones de esta ecuación cuadrática son los valores propios λ_1 y λ_2, que se calculan como:

$$\lambda_{1,2} = \frac{1}{2} \left(a + d \pm \sqrt{(a - d)^2 + 4bc} \right).$$

Dado que V_1 es el mayor valor propio, se corresponde con:

$$V_1 = \frac{1}{2}\left(a + d + \sqrt{(a-d)^2 + 4bc}\right).$$

Ahora, para encontrar el vector propio asociado a V_1, resolvemos $(S - V_1 I)\vec{v} = 0$. Suponiendo que el vector propio es $\vec{v} = (x, y)^T$, tenemos:

$$\begin{pmatrix} a - V_1 & b \\ c & d - V_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dado $c \neq 0$, podemos expresar y en términos de x:

$$cx + (d - V_1)y = 0 \quad \Rightarrow \quad y = -\frac{c}{d - V_1}x.$$

Para que el vector propio tenga longitud unitaria, $x^2 + y^2 = 1$, sustituimos y en términos de x:

$$x^2 + \left(-\frac{c}{d - V_1}x\right)^2 = 1.$$

Resolviendo para x, obtenemos:

$$x = \sqrt{\frac{1}{1 + \left(\frac{c}{d-V_1}\right)^2}} = \sqrt{\frac{(d - V_1)^2}{c^2 + (d - V_1)^2}}.$$

Y para y, usando la relación $y = -\frac{c}{d-V_1}x$, obtenemos:

$$y = \frac{c}{\sqrt{c^2 + (V_1 - d)^2}}.$$

Así, la primera componente principal está dada por:

$$\sqrt{\frac{c^2}{c^2 + (V_1 - d)^2}}X + \frac{c}{|c|}\sqrt{\frac{(V_1 - d)^2}{c^2 + (V_1 - d)^2}}Y,$$

donde V_1 es la varianza explicada por la primera componente principal.

Ejercicio 2

Considerar los datos en el archivo `T8-5.DAT` correspondientes a un tramo censal. Suponer que las observaciones en la variable X_5 = valor de la mediana de hogares fue registrada en unidades de diez miles más que de cientos de miles de dólares (es decir, multipliquen todos los datos listados en la quinta columna por 10).

```
data_ej2 <- read.table("T8-5.DAT", header = TRUE)
data_ej2_adj <- data_ej2
data_ej2_adj[, 5] <- data_ej2_adj[, 5] * 10
```

a. Comparación de estimaciones.

Comparamos las estimaciones con los datos en diez miles y cientos de miles (son dos matrices de covarianzas) para las componentes principales en cada caso.

```
# covarianzas
cov_data_ej2 <- cov(data_ej2)
cov_data_ej2_adj <- cov(data_ej2_adj)

# pca
# scale = FALSE porque nos interesa ver como cambian al ser unidades diferentes
pc_dataej2 <- prcomp(data_ej2, scale = FALSE)
pc_dataej2_adj <- prcomp(data_ej2_adj, scale = FALSE)
```

Matriz de covarianza (datos sin ajustar)

```
             X2.67     X5.71       X69.02       X30.3       X1.48
X2.67    3.3987041 -1.066641   4.30413542  -2.0085678  0.02284220
X5.71   -1.0666408  9.784086  -1.46661133  11.0369308  1.22813980
X69.02   4.3041354 -1.466611  56.46948302 -29.2879393 -0.05073082
X30.3   -2.0085678 11.036931 -29.28793927  90.3787429  0.98260028
X1.48    0.0228422  1.228140  -0.05073082   0.9826003  0.32360845
```

Matriz de covarianza (datos ajustados)

```
            X2.67     X5.71      X69.02      X30.3      X1.48
X2.67    3.398704 -1.066641   4.3041354  -2.008568  0.2284220
X5.71   -1.066641  9.784086  -1.4666113  11.036931 12.2813980
X69.02   4.304135 -1.466611  56.4694830 -29.287939 -0.5073082
```

```
X30.3  -2.008568 11.036931 -29.2879393  90.378743  9.8260028
X1.48   0.228422 12.281398  -0.5073082   9.826003 32.3608446
```

Componentes principales (datos sin ajustar)

```r
summary(pc_dataej2)
```

```
Importance of components:
                           PC1    PC2     PC3     PC4     PC5
Standard deviation     10.4164 6.3516 2.91021 1.69841 0.39487
Proportion of Variance  0.6766 0.2516 0.05282 0.01799 0.00097
Cumulative Proportion   0.6766 0.9282 0.98104 0.99903 1.00000
```

Componentes principales (datos ajustados)

```r
summary(pc_dataej2_adj)
```

```
Importance of components:
                           PC1    PC2    PC3     PC4    PC5
Standard deviation     10.4792 6.6216 5.6375 2.13681 1.5448
Proportion of Variance  0.5708 0.2279 0.1652 0.02373 0.0124
Cumulative Proportion   0.5708 0.7987 0.9639 0.98760 1.0000
```

b. Interpretación de componentes

Con los datos no ajustados la primera componente principal nos indica el negativo de X_4 y la segunda nos indica el negativo de X_3

```r
pc_dataej2$rotation
```

```
                PC1         PC2        PC3         PC4          PC5
X2.67   0.037707731 -0.07089065 -0.1815041 -0.97855412 -0.055142824
X5.71  -0.104364055 -0.12997156  0.9620717 -0.16521675 -0.139057569
X69.02  0.492060576 -0.86458419 -0.0472423  0.09008046  0.004923427
X30.3  -0.863410155 -0.47995778 -0.1524947  0.02941177  0.006613599
X1.48  -0.009250078 -0.01471779  0.1264429 -0.07845842  0.988713448
```

Con los datos ajustados, la primera componente principal nos indica X_4 y la segunda compente princial nos indica X_3 y ahora también tiene también una fuerte componente de X_5

```r
pc_dataej2_adj$rotation
```

```
              PC1        PC2         PC3         PC4         PC5
X2.67  -0.03649865 0.06157149 -0.04082083 -0.55151043  0.83008837
X5.71   0.11825906 0.24967972  0.26195956  0.77039545  0.51141263
X69.02 -0.47911558 0.76107456 -0.42889373  0.02649873 -0.08100467
X30.3   0.85908666 0.31585118 -0.39396783 -0.06726957 -0.04972228
X1.48   0.13077086 0.50484902  0.76847275 -0.31160828 -0.20093869
```

c. Efectos en el cambio de escala.

Podemos ver como los efectos de la variable X_5 en verdad erán más importantes, sin embargo como no estaba bien reescalada, el primer análisis de ocmponentes principales, no lo tomó en cuenta.

Extra (utilizando Scale = True)

```r
pc_dataej2_extra <- prcomp(data_ej2, scale = TRUE)
pc_dataej2_extra$rotation
```

```
             PC1        PC2         PC3        PC4        PC5
X2.67   0.2558707  0.4641923 -0.78553998  0.2262478 -0.2253765
X5.71  -0.5949383  0.3238268  0.15923983 -0.1403305 -0.7043682
X69.02  0.3209996  0.6108437  0.21980378 -0.6624874  0.1913790
X30.3  -0.4769352 -0.2576230 -0.55099788 -0.5723341  0.2738580
X1.48  -0.5000707  0.4900633  0.07521475  0.4033155  0.5843344
```

Al estandarizar las variables vemos como todas las cargas de las variables en las componentes principales cambian por completo.

Ejercicio 3

Considerar los datos sobre toros en el conjunto de datos T1 – 10.DAT sobre toros. Estos datos contienen las características medidas de 76 toros jóvenes (menores a dos años) vendidos en una subasta. Los datos que se incluyen corresponden a las siguientes variables:

- `Raza`: 1= Angus, 5= Hereford, 8= Simmental
- `PVenta`: precio de venta

- `YrHgt`: medición al hombro al año (pulgadas)
- `FtFreBody`: Cuerpo libre de grasa (libras)
- `PrctFFB`: Porcentaje del cuerpo libre de grasa
- `Frame`: Cornamenta. Escala de 1 (pequeña) a 8 (grande)
- `BkFat`: Grasa trasera (en pulgadas)
- `SaleHt`: medición al hombro en el momento de venta (pulgadas)
- `SaleWt`: peso de venta (libras)

Utilizando las 7 últimas variables dadas, hacer un análisis de componentes principales usando la matriz de covarianzas de los datos y la matriz de correlación. El análisis debe incluir lo siguiente:

a. Número de componentes

Para determinar el número de componentes apropiadas utilizamos un Scree plot

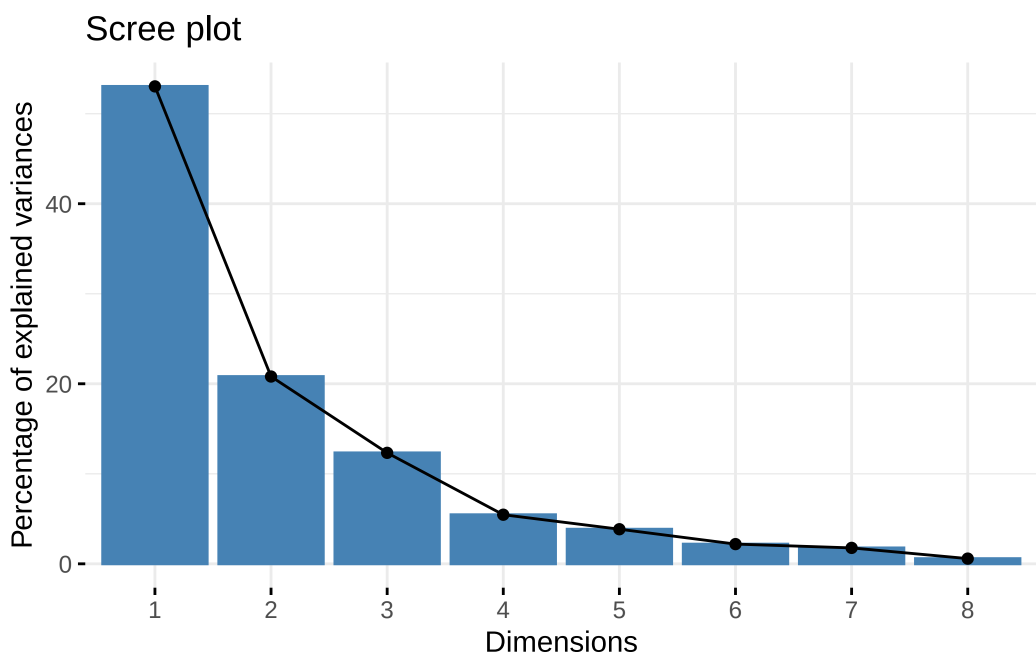
```
dataej3 <- read.table("T1-10.DAT", header=TRUE)
colnames(dataej3) <- c("Raza", "PVenta", "YrHgt", "FtFreBody", "PrctFFB",
                       "Frame", "BkFat", "SaleHt", "SaleWt")

#quitamos datos categóricos
dataej3_pca <- dataej3[, c("PVenta", "YrHgt", "FtFreBody", "PrctFFB",
                           "Frame", "BkFat", "SaleHt", "SaleWt")]


pca_dataej3 <- prcomp(dataej3_pca, scale = TRUE)
summary(pca_dataej3)
```

```
Importance of components:
                          PC1    PC2    PC3     PC4     PC5    PC6     PC7
Standard deviation     2.0598 1.2901 0.9930 0.66049 0.55454 0.4186 0.37620
Proportion of Variance 0.5303 0.2081 0.1233 0.05453 0.03844 0.0219 0.01769
Cumulative Proportion  0.5303 0.7384 0.8617 0.91618 0.95462 0.9765 0.99421
                           PC8
Standard deviation     0.21517
Proportion of Variance 0.00579
Cumulative Proportion  1.00000
```

```
#plot(pca_dataej3, type = "l")

library(factoextra)
```

```
Warning: package 'factoextra' was built under R version 4.2.3

Loading required package: ggplot2

Warning: package 'ggplot2' was built under R version 4.2.3

Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
fviz_screeplot(pca_dataej3,ncp=11)
```



Parece ser que con tres componentes tenemos una buena explicabilidad de la varianza.

b. Interpretación de componentes

Interpretación de las componentes principales.

```
pca_dataej3$rotation[, c("PC1", "PC2", "PC3")]
```

```
                PC1         PC2         PC3
PVenta    -0.1881335  0.56899730 -0.41423552
YrHgt     -0.4494822  0.01639424 -0.28131231
FtFreBody -0.3973579 -0.08907817  0.43005470
PrctFFB   -0.3308616 -0.41369032  0.22448498
Frame     -0.4363243  0.07286536 -0.30165981
BkFat      0.1636283  0.60320520  0.32684981
SaleHt    -0.4509534  0.06350433 -0.02048673
SaleWt    -0.2733070  0.35171013  0.56185640
```

La primera componente principal nos indica que tan corta es la distancia al hombro en el momento de la venta, que tan chica es la cornamenta y que tan pequeña fue la medición al hombro luego de un año. Es decir que mientras mayor sea la primera componente principal de un toro, menor serán esas tres características.

La segunda componente principal, nos indica el el precio de venta del toro, la cantidad de grasa trasera y el negativo el porcentaje del cuerpo libre de grasa. Mientras mayor sea esta componente, más caro se vendió el toro, mayor grasa trasera tenía y menos porcentaje del cuerpo libre de grasa tenía.

La tercera componente principal nos indica el peso de venta y el cuerpo libre de grasa.

c. Índice de tamaño de cuerpo

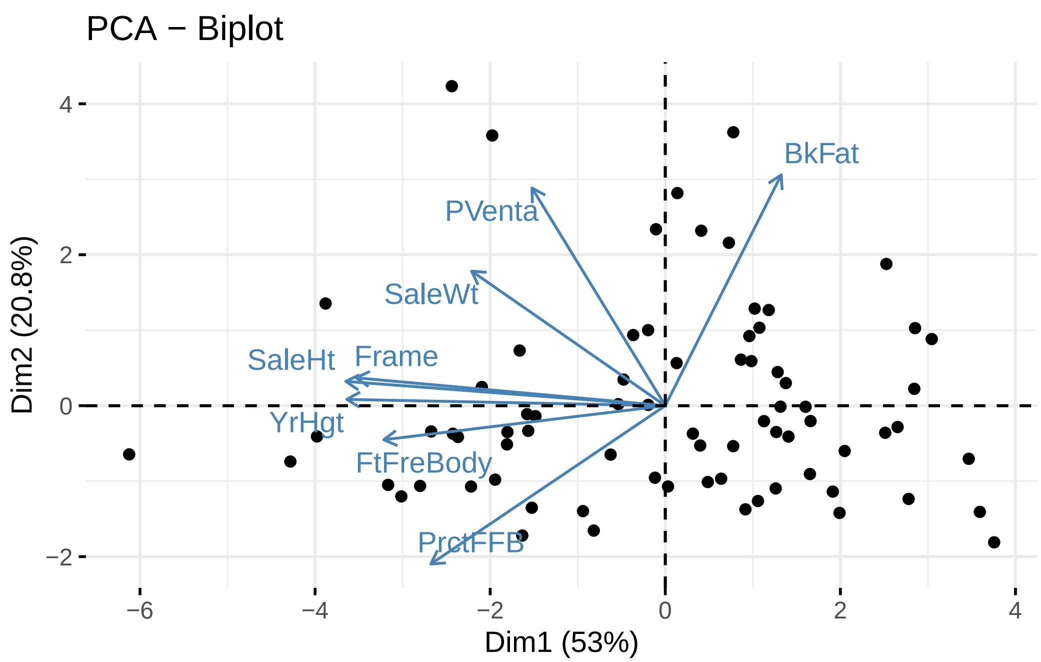
¿Será posible desarrollar un índice 'Tamaño de cuerpo' o 'configuración de cuerpo' basado en las 7 variables consideradas?

El negativo de la primera componente principal parece servir como ese indice, ya que las cargas de las variables del tamaño del cuerpo del toro son las más altas.

d. Gráfica de componentes

Hacer una gráfica de las dos primeras componentes. ¿Hay outliers? Si los hay, hacer una sustitución de la matriz de covarianzas con una matriz de covarianzas estimada de manera robusta.

```r
fviz_pca_biplot(pca_dataej3,repel = T,geom.ind = "point")
```



e. Normalidad de datos originales

Evalúen si los datos originales son normales. Si no lo son, buscar las transformaciones que los acerquen a normalidad. Repetir el análisis con los datos transformados y probar la significancia de la varianza de las componentes principales con el resultado de Anderson.

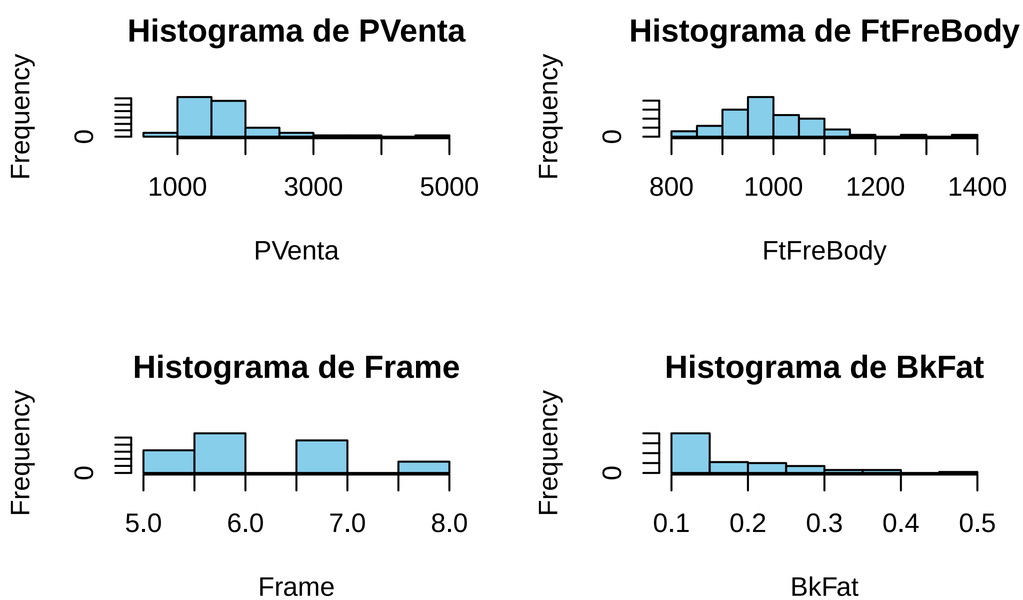
```
apply(dataej3_pca, 2, shapiro.test)
```

```
$PVenta

    Shapiro-Wilk normality test

data:  newX[, i]
W = 0.78455, p-value = 4.064e-09


$YrHgt

    Shapiro-Wilk normality test

data:  newX[, i]
```

```
W = 0.97939, p-value = 0.2605


$FtFreBody

    Shapiro-Wilk normality test

data:  newX[, i]
W = 0.92569, p-value = 0.000289


$PrctFFB

    Shapiro-Wilk normality test

data:  newX[, i]
W = 0.97047, p-value = 0.07677


$Frame

    Shapiro-Wilk normality test

data:  newX[, i]
W = 0.87622, p-value = 2.61e-06


$BkFat

    Shapiro-Wilk normality test

data:  newX[, i]
W = 0.87396, p-value = 2.161e-06


$SaleHt

    Shapiro-Wilk normality test

data:  newX[, i]
W = 0.9914, p-value = 0.8972
```

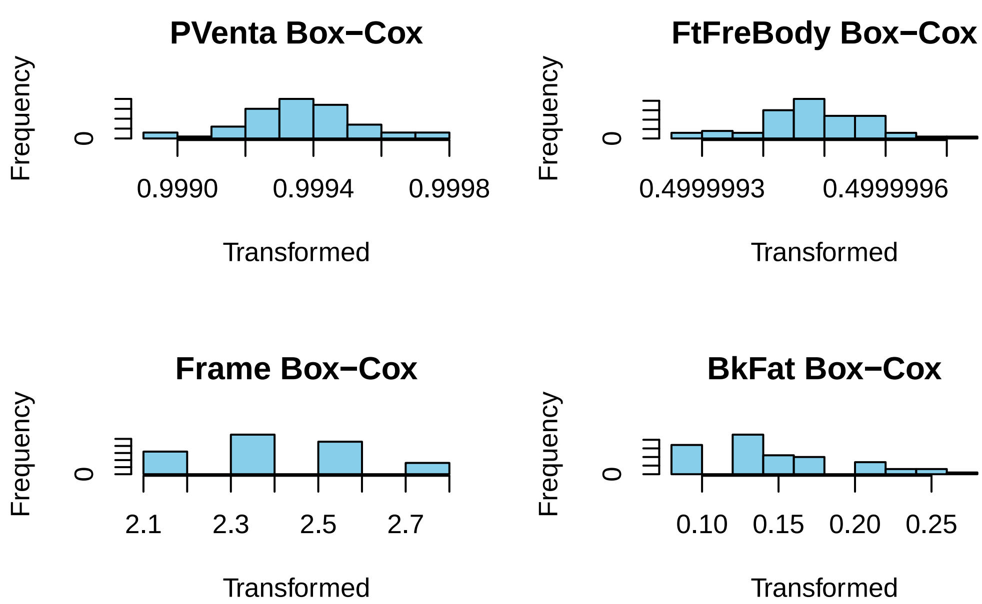
```
$SaleWt

    Shapiro-Wilk normality test

data:  newX[, i]
W = 0.98558, p-value = 0.5545
```

podemos ver que las columnas Pventa, FtFreBody (a pesar de tener una W alta, tenemos un p value bajo lo que indica que es poco probale que sean normales), Frame y BkFat no siguen distribución normal.




```r
library(MASS)

# PVenta
bc_result_PVenta <- boxcox(dataej3_pca$PVenta ~ 1, plotit = FALSE)
lambda_optimo_PVenta <- bc_result_PVenta$x[which.max(bc_result_PVenta$y)]
dataej3_pca$PVenta_bc <-
  (dataej3_pca$PVenta^lambda_optimo_PVenta - 1) / lambda_optimo_PVenta

# FtFreBody
bc_result_FtFreBody <- boxcox(dataej3_pca$FtFreBody ~ 1, plotit = FALSE)
lambda_optimo_FtFreBody <- bc_result_FtFreBody$x[which.max(bc_result_FtFreBody$y)]
dataej3_pca$FtFreBody_bc <-
```

```
  (dataej3_pca$FtFreBody^lambda_optimo_FtFreBody - 1) / lambda_optimo_FtFreBody

# Frame
dataej3_pca$Frame_adj <- dataej3_pca$Frame + (min(dataej3_pca$Frame, 0)) + 1
bc_result_Frame <- boxcox(dataej3_pca$Frame_adj ~ 1, plotit = FALSE)
lambda_optimo_Frame <- bc_result_Frame$x[which.max(bc_result_Frame$y)]
dataej3_pca$Frame_bc <-
  (dataej3_pca$Frame_adj^lambda_optimo_Frame - 1) / lambda_optimo_Frame

#  BkFat
dataej3_pca$BkFat_adj <- dataej3_pca$BkFat + (min(dataej3_pca$BkFat, 0)) + 1
bc_result_BkFat <- boxcox(dataej3_pca$BkFat_adj ~ 1, plotit = FALSE)
lambda_optimo_BkFat <- bc_result_BkFat$x[which.max(bc_result_BkFat$y)]
dataej3_pca$BkFat_bc <-
  (dataej3_pca$BkFat_adj^lambda_optimo_BkFat - 1) / lambda_optimo_BkFat
```




```
lapply(dataej3_pca[, c("PVenta_bc", "FtFreBody_bc", "Frame_bc", "BkFat_bc")],
       shapiro.test)
```

```
$PVenta_bc

    Shapiro-Wilk normality test
```

```
data:  X[[i]]
W = 0.98303, p-value = 0.4138


$FtFreBody_bc

	Shapiro-Wilk normality test

data:  X[[i]]
W = 0.98585, p-value = 0.5708


$Frame_bc

	Shapiro-Wilk normality test

data:  X[[i]]
W = 0.87557, p-value = 2.472e-06


$BkFat_bc

	Shapiro-Wilk normality test

data:  X[[i]]
W = 0.90639, p-value = 3.975e-05
```

```r
library(nortest)
ad.test(dataej3_pca$PVenta_bc)
```

```

	Anderson-Darling normality test

data:  dataej3_pca$PVenta_bc
A = 0.38438, p-value = 0.3858
```

```r
ad.test(dataej3_pca$FtFreBody_bc)
```

```
	Anderson-Darling normality test

data:  dataej3_pca$FtFreBody_bc
A = 0.31677, p-value = 0.5326
```

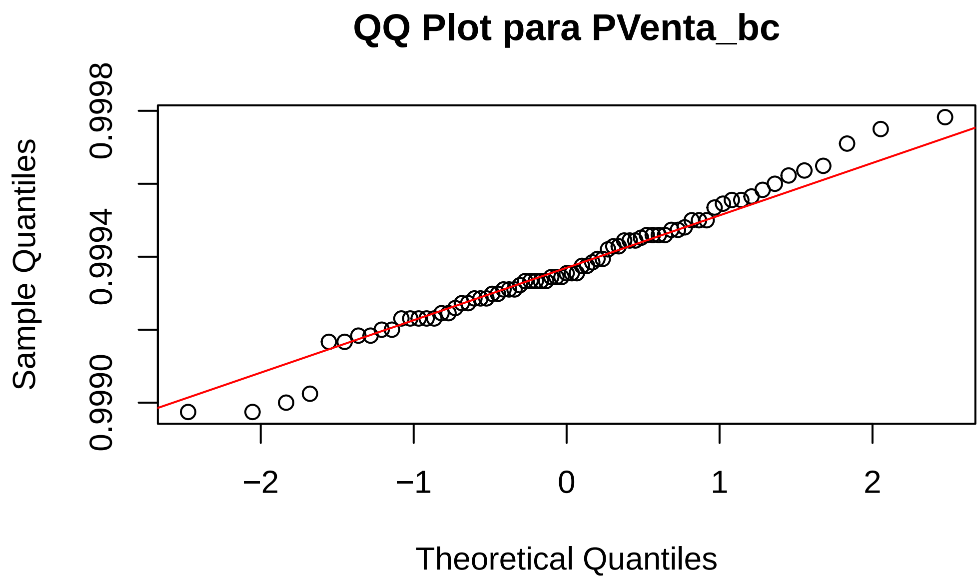
```
ad.test(dataej3_pca$Frame_bc)
```

```
	Anderson-Darling normality test

data:  dataej3_pca$Frame_bc
A = 3.7443, p-value = 1.984e-09
```

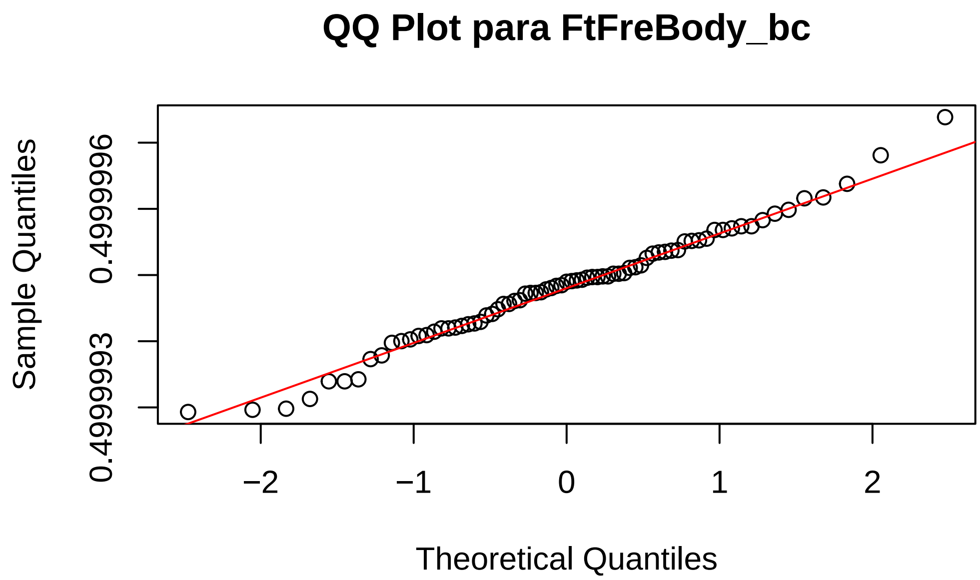
```
ad.test(dataej3_pca$BkFat_bc)
```

```
	Anderson-Darling normality test

data:  dataej3_pca$BkFat_bc
A = 2.4532, p-value = 2.923e-06
```

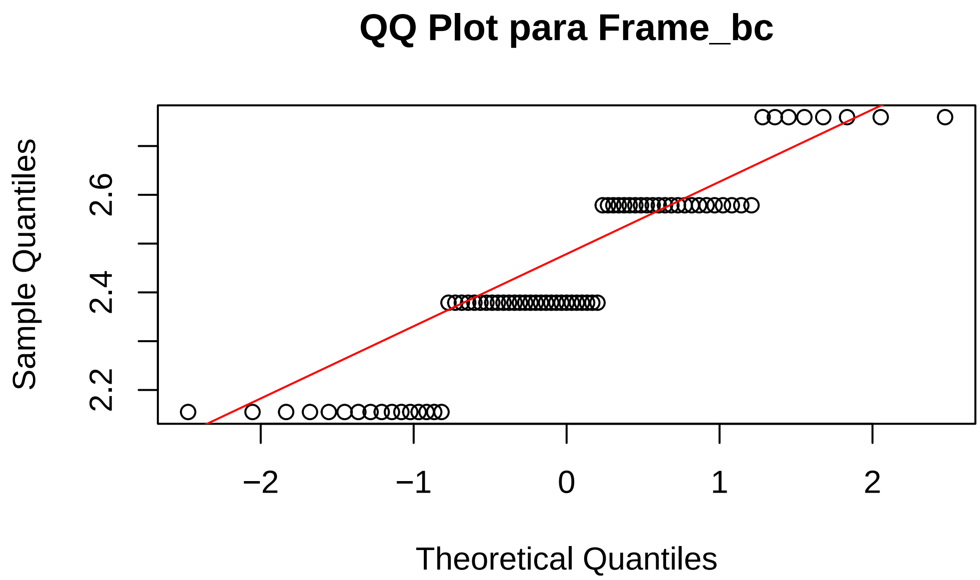
```
# QQ plot para PVenta_bc
qqnorm(dataej3_pca$PVenta_bc, main="QQ Plot para PVenta_bc")
qqline(dataej3_pca$PVenta_bc, col="red")
```



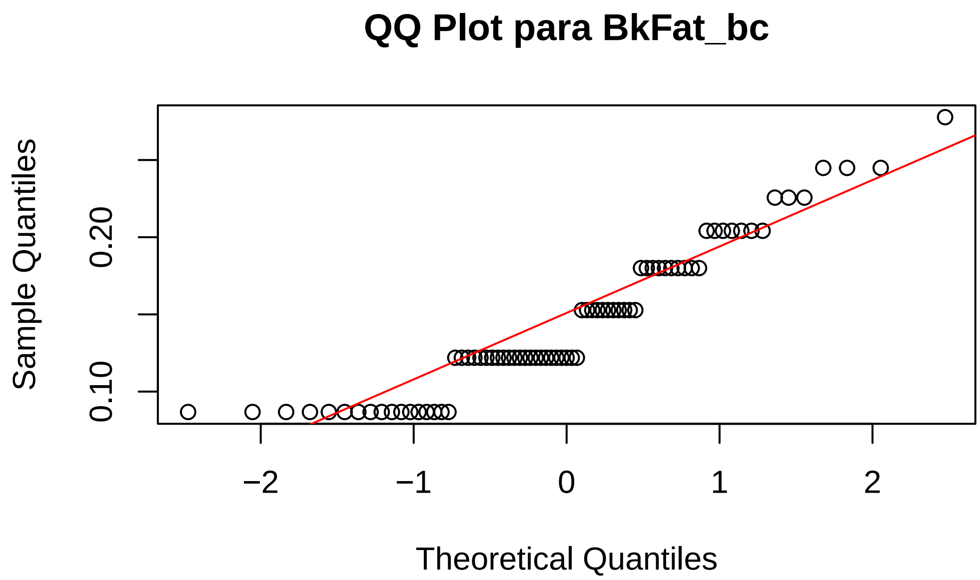

```
# QQ plot para FtFreBody_bc
qqnorm(dataej3_pca$FtFreBody_bc, main="QQ Plot para FtFreBody_bc")
qqline(dataej3_pca$FtFreBody_bc, col="red")
```




```r
# QQ plot para Frame_bc
qqnorm(dataej3_pca$Frame_bc, main="QQ Plot para Frame_bc")
qqline(dataej3_pca$Frame_bc, col="red")
```




```
# QQ plot para BkFat_bc
qqnorm(dataej3_pca$BkFat_bc, main="QQ Plot para BkFat_bc")
qqline(dataej3_pca$BkFat_bc, col="red")
```



Ejercicio 4

Consideren la matriz de correlaciones siguiente. Los datos originales corresponden a las mediciones de 8 variables de química sanguínea de 72 pacientes en un estudio clínico. (Jolliffe, 2002). La matriz de correlaciones de las variables rblood, plate, wblood, neut, lymph, bilir, sodium y potass, en ese orden, es la siguiente:

```
ej4_corr <- c(
  1.000, 0.290, -0.202, -0.055, -0.105, -0.252, -0.229,  0.058,
  0.290, 1.000,  0.415,  0.285, -0.376, -0.349, -0.164, -0.129,
 -0.202, 0.415,  1.000,  0.419, -0.521, -0.441, -0.145, -0.076,
 -0.055, 0.285,  0.419,  1.000, -0.877, -0.076,  0.023, -0.131,
 -0.105, -0.376, -0.521, -0.877,  1.000,  0.206,  0.034,  0.151,
 -0.252, -0.349, -0.441, -0.076,  0.206,  1.000,  0.192,  0.077,
 -0.229, -0.164, -0.145,  0.023,  0.034,  0.192,  1.000,  0.423,
  0.058, -0.129, -0.076, -0.131,  0.151,  0.077,  0.423,  1.000
)
```

y las desviaciones estándar, que tienen considerables diferencias, son:

rblood plate wblood neut lymph bilir sodium potass

```r
ej4_desv <- c(0.371, 41.253, 1.935, 0.077, 0.071, 4.037, 2.732, 0.297)
```

```r
# reconstruimos la matriz de correlaciones
ej4_corr_matrix <- matrix(ej4_corr, nrow = 8, ncol = 8)
# Calculamos la matriz de covarianzas
ej4_cov_matrix <- ej4_corr_matrix * (ej4_desv %*% t(ej4_desv))
```

a. Componentes Pricipales

```r
# PCA usando la matriz de correlaciones
ej4_pca_cor <- prcomp(ej4_corr_matrix, scale = TRUE)
summary(ej4_pca_cor)
```

```
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7
Standard deviation     2.0844 1.2750 0.9958 0.76882 0.50967 0.39249 0.18196
Proportion of Variance 0.5431 0.2032 0.1240 0.07389 0.03247 0.01926 0.00414
Cumulative Proportion  0.5431 0.7463 0.8702 0.94413 0.97660 0.99586 1.00000
                             PC8
Standard deviation     6.281e-17
Proportion of Variance 0.000e+00
Cumulative Proportion  1.000e+00
```

```r
# PCA usando la matriz de covarianzas
ej4_pca_cov <- prcomp(ej4_cov_matrix, scale = FALSE)
summary(ej4_pca_cov)
```

```
Importance of components:
                            PC1     PC2     PC3   PC4    PC5     PC6      PC7
Standard deviation     604.8284 5.28943 2.48393 0.893 0.0278 0.02245 0.002638
Proportion of Variance   0.9999 0.00008 0.00002 0.000 0.0000 0.00000 0.000000
Cumulative Proportion    0.9999 0.99998 1.00000 1.000 1.0000 1.00000 1.000000
                             PC8
Standard deviation     8.629e-17
Proportion of Variance 0.000e+00
Cumulative Proportion  1.000e+00
```

Aplicar componentes principales a la matriz de covarianzas y a la matriz de correlaciones. Explicar las diferencias.

Para el PCA basado en la matriz de correlaciones, vemos que:

PC1 explica el 54.31% de la varianza. PC2 explica el 20.32% de la varianza. Sumando hasta la tercera componente principal, explican el 87.02% de la varianza. Para el PCA basado en la matriz de covarianzas, el primer componente principal explica prácticamente toda la varianza (99.99%)

b. Interpretación para análisis

Basado en la observación anterior, sobre qué debería hacerse el análisis?

Podemos ver que es mejor hacer el pca con la matriz de correlaciones ya que la varianza no se encuentra tan cargada en la primer componente principal como pasa con la matriz de covarianzas. Al usar la matriz de correlaciones permitimos que cada variable aporte de forma significativa al análisis. No dejamos que una sola variable de gran varianza a escala domine todo el análisis.

Ejercicio 5

Encontrar las componentes principales de la siguiente matriz de correlación calculada de las mediciones de 7 características físicas en 3,000 convictos criminales: Las variables son: 1. largo de la cabeza, 2. ancho de la cabeza, 3. ancho de la cara, 4. longitud del dedo pulgar izquierdo, 5. longitud del antebrazo izquierdo, 6. longitud del pie izquierdo, 7. Altura.

$$\begin{bmatrix} 1 & & & & & & \\ 0.402 & 1 & & & & & \\ 0.396 & 0.618 & 1 & & & & \\ 0.301 & 0.150 & 0.321 & 1 & & & \\ 0.305 & 0.135 & 0.289 & 0.846 & 1 & & \\ 0.339 & 0.206 & 0.365 & 0.759 & 0.797 & 1 & \\ 0.340 & 0.183 & 0.345 & 0.661 & 0.800 & 0.736 & 1 \end{bmatrix}$$

```r
var_names <-  c("LargoCabeza", "AnchoCabeza", "AnchoCara", "LongPulgarIzq",
                "LongAntebrazoIzq", "LongPieIzq", "Altura")

rownames(ej5_corr_matrix) <- var_names
colnames(ej5_corr_matrix) <- var_names
# Realizar PCA
```

```r
ej5_pca <- prcomp(ej5_corr_matrix)

summary(ej5_pca)
```

```
Importance of components:
                          PC1    PC2     PC3    PC4     PC5     PC6       PC7
Standard deviation     0.7061 0.2677 0.15380 0.1385 0.09824 0.04686 3.986e-17
Proportion of Variance 0.7978 0.1147 0.03785 0.0307 0.01544 0.00351 0.000e+00
Cumulative Proportion  0.7978 0.9125 0.95034 0.9810 0.99649 1.00000 1.000e+00
```

```r
# cargas
ej5_pca$rotation
```

```
                        PC1         PC2         PC3         PC4         PC5
LargoCabeza      -0.1665810  0.80818753 -0.26192541  0.06009251 -0.09814788
AnchoCabeza      -0.4245453 -0.32256714  0.50643740 -0.12662191 -0.14034146
AnchoCara        -0.2611507 -0.46790217 -0.79355859  0.12799746  0.02333546
LongPulgarIzq     0.4333626 -0.09341607  0.07747033  0.60939351  0.24997097
LongAntebrazoIzq  0.4786835 -0.06054039  0.05709355  0.02909543  0.21776869
LongPieIzq        0.3957672 -0.09885027 -0.07335440  0.02447247 -0.90635088
Altura            0.3876075 -0.04108336 -0.17475356 -0.76887355  0.19684036
                         PC6        PC7
LargoCabeza      -0.07557717 0.48116809
AnchoCabeza      -0.06426870 0.64759058
AnchoCara        -0.10181694 0.23630075
LongPulgarIzq     0.45520500 0.39548840
LongAntebrazoIzq -0.81934401 0.21057766
LongPieIzq        0.01324359 0.07726485
Altura            0.31792978 0.29417985
```