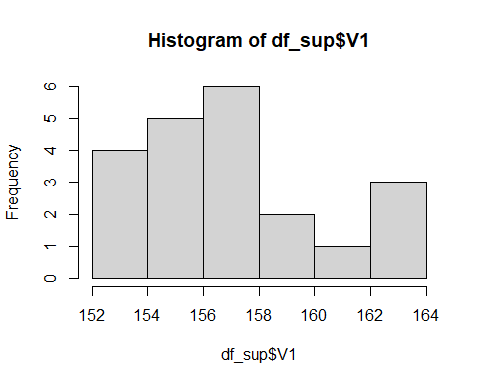
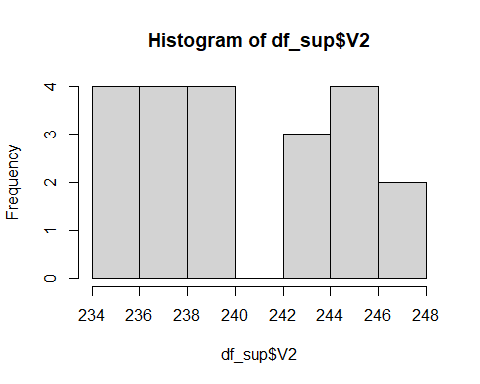
Guia 3.1 ejercicio 4

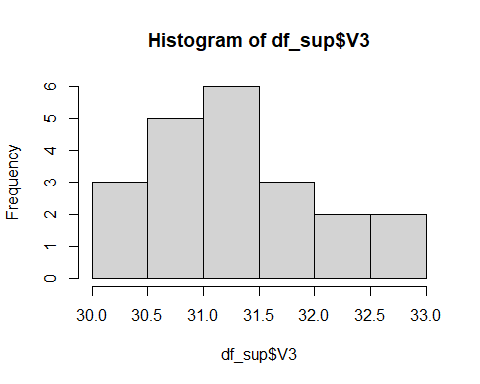
df\_sup = read.table("C:\\Users\\Dell7400\\Documents\\Ale\\Facu\\Multivariado\\datos\\supervivientes.txt")  
df\_no\_sup = read.table("C:\\Users\\Dell7400\\Documents\\Ale\\Facu\\Multivariado\\datos\\no\_supervivientes.txt")  
  
  
# veamos primero la pinta de los datos  
  
hist(df\_sup$V1)



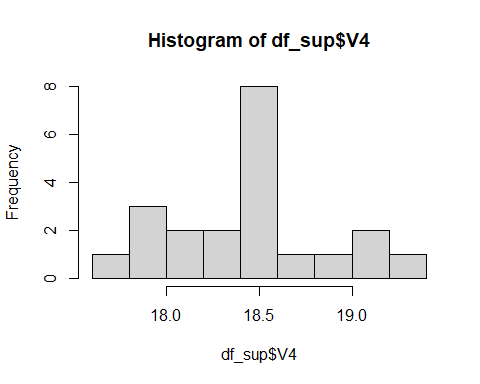
hist(df\_sup$V2)



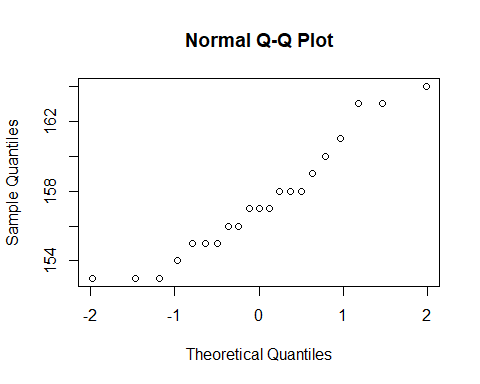
hist(df\_sup$V3)



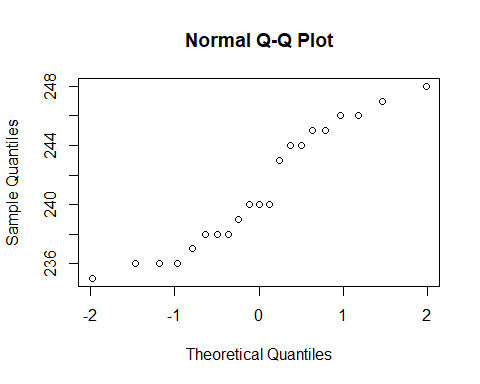
hist(df\_sup$V4)



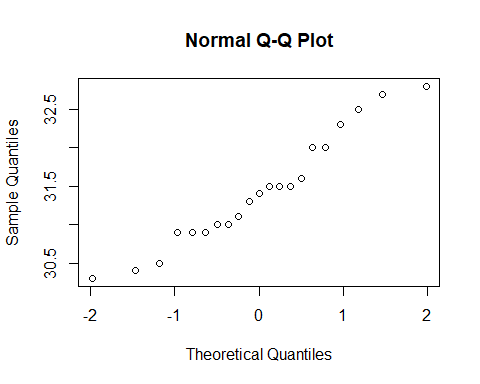
qqnorm(df\_sup$V1)



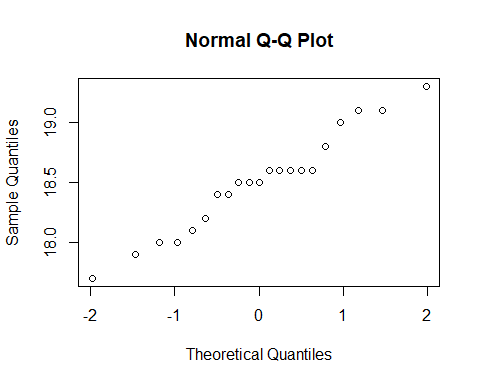
qqnorm(df\_sup$V2)



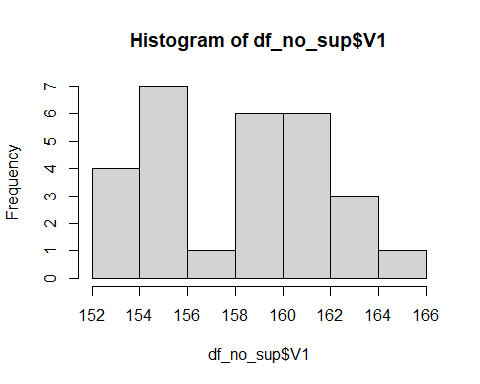
qqnorm(df\_sup$V3)



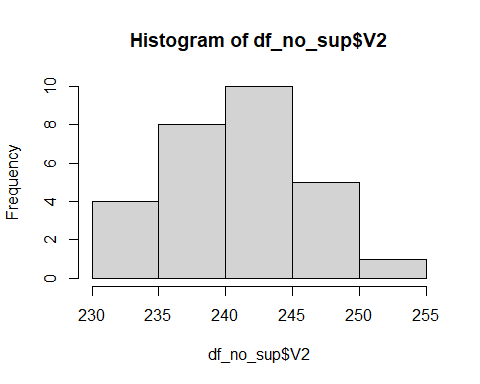
qqnorm(df\_sup$V4)



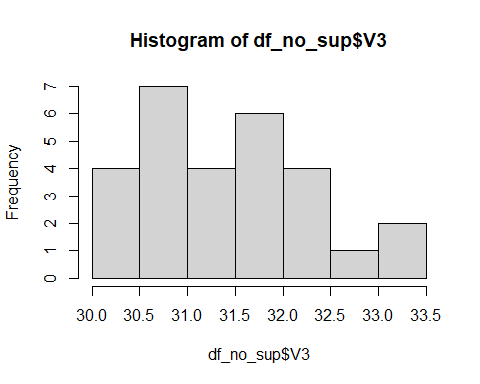
hist(df\_no\_sup$V1)



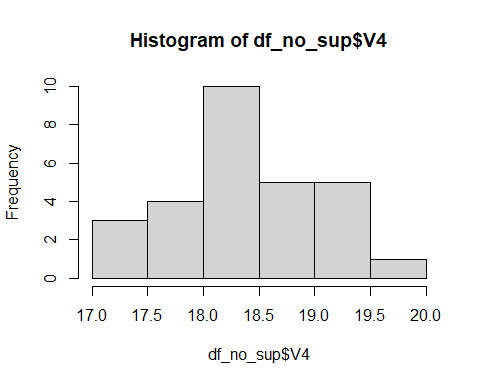
hist(df\_no\_sup$V2)



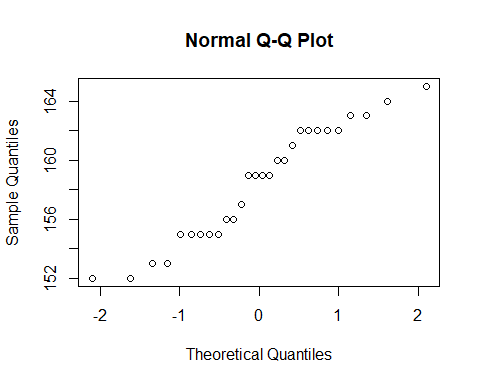
hist(df\_no\_sup$V3)



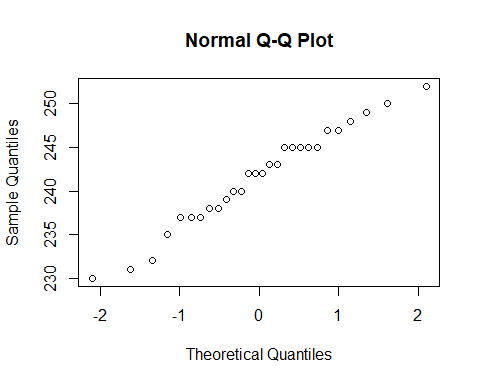
hist(df\_no\_sup$V4)



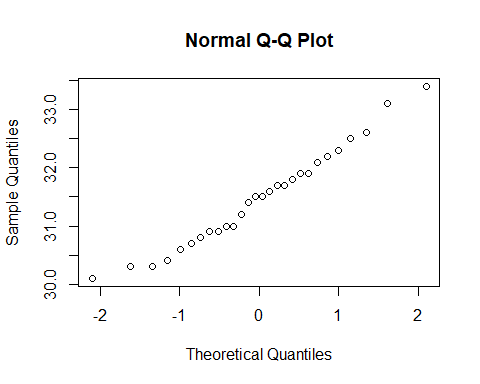
qqnorm(df\_no\_sup$V1)



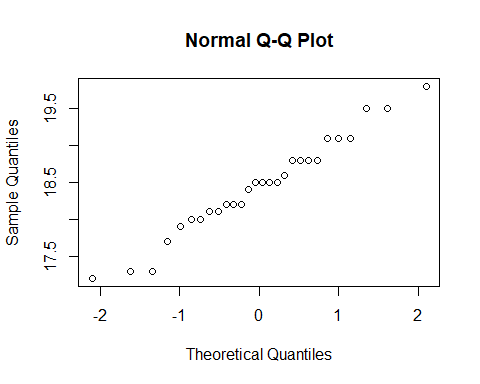
qqnorm(df\_no\_sup$V2)



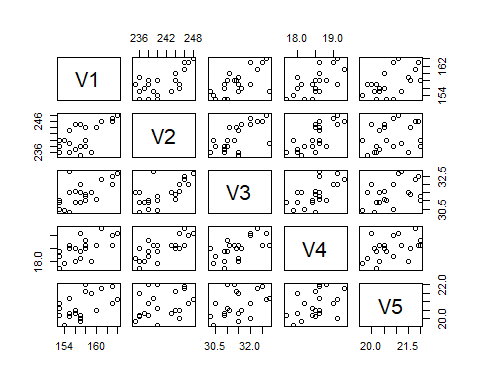
qqnorm(df\_no\_sup$V3)



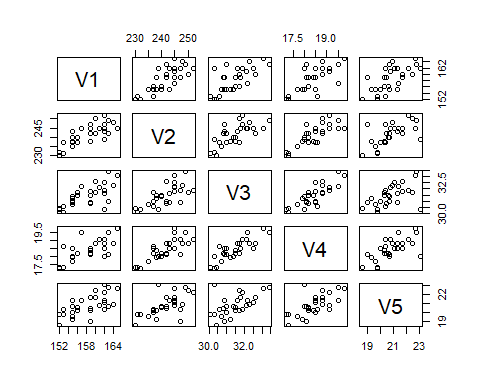
qqnorm(df\_no\_sup$V4)



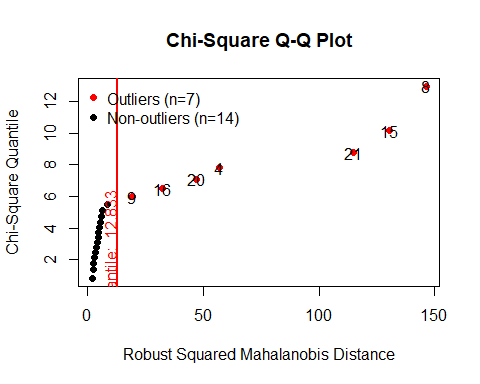
# En forma univariada pocas parecieran ser normales  
  
# de todas formas, no hay que confiarse, son pocos datos  
  
  
# Veamos un grafico bivariado  
  
plot(df\_sup)



# dependiendo el par de variables que vemaos, parecieran tener forma de elipse, posiblemente con mas observaciones   
  
  
plot(df\_no\_sup)  
  
# en este caso, la gran mayoria de par de variables tienen forma de elipse, al menos dos a dos parecieran ser normales  
  
  
# Hagamos un test de shapiro wilks para analizar la normalidad conjunta, tomando alfa 0.05  
  
alfa = 0.05  
  
library(MVN)



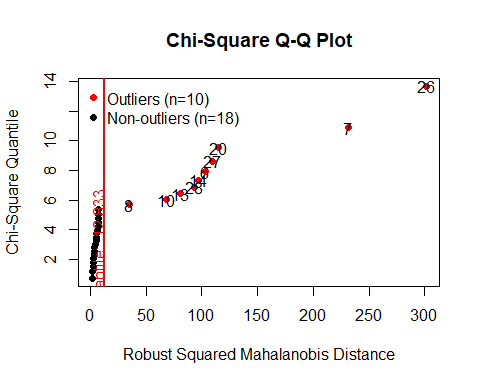
test\_sup = mvn(df\_sup, mvnTest = "hz", multivariateOutlierMethod = "quan")



pvalue\_sup = test\_sup$multivariateNormality[3]  
  
ifelse(pvalue\_sup<alfa,"rechazo, no hay normalidad","no rechazo, hay normalidad multivariada")

## p value   
## [1,] "no rechazo, hay normalidad multivariada"

test\_no\_sup = mvn(df\_no\_sup, mvnTest = "hz", multivariateOutlierMethod = "quan")



pvalue\_no\_sup = test\_no\_sup$multivariateNormality[3]  
  
ifelse(pvalue\_no\_sup<alfa,"rechazo, no hay normalidad","no rechazo, hay normalidad multivariada")

## p value   
## [1,] "no rechazo, hay normalidad multivariada"

# Al parecer hay normalidad multivariada en ambas muestras, procedemos a construir los estadisticos de Hotelling suponiendo  
  
#- Normalidad en los vectores aleatorios e independencia entre muestas  
#- Igualdad de matriz de varianzas y covarianzas  
  
  
n1 = nrow(df\_sup)  
n2 = nrow(df\_no\_sup)  
  
  
df\_1\_medias = apply(df\_sup,2,mean)  
df\_2\_medias = apply(df\_no\_sup,2,mean)  
  
resta = df\_1\_medias-df\_2\_medias  
  
s1 = cov(df\_sup)  
s2 = cov(df\_no\_sup)  
  
s = ((n1-1)\*s1+(n2-1)\*s2)/(n1+n2-2)  
  
# construyo el valor del estadistico To\_2  
  
To\_2 = ((n1\*n2)/(n1+n2))\*t(resta)%\*%solve(s)%\*%resta   
  
p = ncol(df\_sup)  
  
  
Fo= ((n1+n2-p-1)/((n1+n2-2)\*p)) \* To\_2  
  
#busco el fractil de la t student cn p,n1+n2-p-1 grados de libertad  
  
gl1 = p  
gl2 = n1+n2-p-1  
  
Fcritico = qf(0.99, gl1, gl2, lower.tail = T, log.p = F)  
  
ifelse(Fo>Fcritico,"Rechazo Ho, no hay igualdad entre los vectores de medias","No Rechazo Ho, hay igualdad entre los vectores de medias")

## [,1]   
## [1,] "No Rechazo Ho, hay igualdad entre los vectores de medias"

#b  
  
#La combinación lineal del componente de medias es donde se alcanza el supremo, entonces  
  
  
solve(s)%\*%resta

## [,1]  
## V1 -0.15532570  
## V2 -0.02649058  
## V3 -0.09285760  
## V4 1.03247387  
## V5 0.06932512