# PRÁCTICA 4

### Álvaro Larraya

## Ejercicio 1

## a)

library(PASWR2)

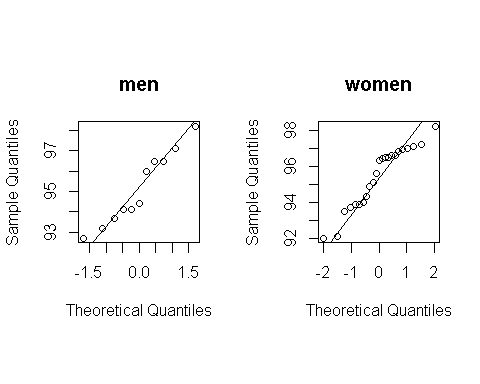
## Loading required package: lattice

## Loading required package: ggplot2

str(STATTEMPS)

## 'data.frame': 34 obs. of 3 variables:  
## $ temperature: num 92.7 94.1 96.5 93.2 97.1 93.7 94.1 96 98.2 96.5 ...  
## $ gender : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 2 2 2 2 ...  
## $ class : Factor w/ 2 levels "8 a.m.","9 a.m.": 1 1 1 1 1 1 2 2 2 2 ...

men<-STATTEMPS$temperature[STATTEMPS$gender=="Male"]  
women<-STATTEMPS$temperature[STATTEMPS$gender=="Female"]  
par(mfrow=c(1,2),pty="s")  
qqnorm(men, main="men")  
qqline(men)  
qqnorm(women, main="women")  
qqline(women)



Podemos ver que se asemejan más o menos a la distribución de una normal teórica, siendo los puntos los cuantiles de la muestra y la línea la distribución normal con la que la estamos comparando. Al haber similitudes notables, las aproximaremos a una normal.  
  
Como no sabemos si las varianzas son iguales o distintas, realizamos el cociente de varianzas  
  
```r  
var.test(x=men, y=women, conf.level = 0.95)$conf

## [1] 0.4041425 3.6985107  
## attr(,"conf.level")  
## [1] 0.95

Luego el cociente de varianzas está entre 0.4041425 y 3.6985107 con una confianza del 95%, al estar el uno incluido en el intervalo, podemos suponer que las varianzas son iguales

t.test(x=men, y=women, var.equal=TRUE, conf.level = 0.95)$conf

## [1] -1.6025708 0.9970372  
## attr(,"conf.level")  
## [1] 0.95

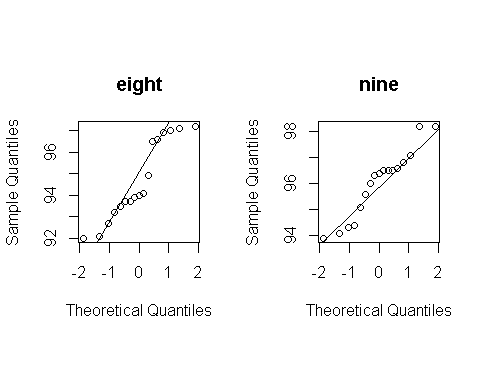
La difirencia de medias entre hombres y mujeres en la temperatura corporal está entre -1.6025708 y 0.9970372 con una confianza del 95%, al estar incluido el cero, concluimos que no hay una diferencia de temperatura entre generos.

## b)

library(PASWR2)  
str(STATTEMPS)

## 'data.frame': 34 obs. of 3 variables:  
## $ temperature: num 92.7 94.1 96.5 93.2 97.1 93.7 94.1 96 98.2 96.5 ...  
## $ gender : Factor w/ 2 levels "Female","Male": 2 2 2 2 2 2 2 2 2 2 ...  
## $ class : Factor w/ 2 levels "8 a.m.","9 a.m.": 1 1 1 1 1 1 2 2 2 2 ...

eight<-STATTEMPS$temperature[STATTEMPS$class=="8 a.m."]  
nine<-STATTEMPS$temperature[STATTEMPS$class=="9 a.m."]  
par(mfrow=c(1,2),pty="s")  
qqnorm(eight, main="eight")  
qqline(eight)  
qqnorm(nine, main="nine")  
qqline(nine)



Podemos aproximar también las temeperaturas medidas a distintas horas a una normal por la misma razón que en el apartado anterior

Como no sabemos si las varianzas son iguales o distintas, realizamos el cociente de varianzas

var.test(x=eight, y=nine, conf.level = 0.95)$conf

## [1] 0.7188687 5.4814489  
## attr(,"conf.level")  
## [1] 0.95

Luego el cociente de varianzas está entre 0.7188687 y 5.4814489 con una confianza del 95%, al estar el uno incluido en el intervalo, podemos suponer que las varianzas son iguales

t.test(x=eight, y=nine, var.equal=TRUE, conf.level = 0.95)$conf

## [1] -2.4917186 -0.2612226  
## attr(,"conf.level")  
## [1] 0.95

La diferencia de medias entre la temperaturas tomadas a las 8 am y a las 9 am está entre -2.4917186 y -0.2612226, al no estar incluido el cero concluimos con una confianza del 95% que la hora a la que se mide la temperatura es un parametro con cierta importancia

## Ejercicio 2

library(PASWR2)  
str(TOE)

## 'data.frame': 12 obs. of 1 variable:  
## $ energy: int 12222 6674 15961 3994 2841 1036 1343 4608 5864 17390 ...

Como desconocemos la varianza de cantidad de TOE^2 de los doce países

t.test(TOE , conf.level = 0.95)$conf

## [1] 3756.425 12788.075  
## attr(,"conf.level")  
## [1] 0.95

La media de cantidad de produccion de energía nuclear está entre 3756.425 y 12788.075 TOE con una confianza del 95%

## Ejercicio 3

## a)

library(PASWR2)  
str(FORMULA1)

## 'data.frame': 10 obs. of 3 variables:  
## $ race : int 1 2 3 4 5 6 7 8 9 10  
## $ team1: num 5.61 6.13 5.42 5.95 5.51 ...  
## $ team2: num 5.93 5.33 5.83 4.82 5.66 ...

team1<-FORMULA1$team1  
team2<-FORMULA1$team2

Realizamos el cociente de varianzas

var.test(x=team1, y=team2, conf.level = 0.95)$conf

## [1] 0.1099935 1.7828442  
## attr(,"conf.level")  
## [1] 0.95

El cociente de varianzas está entre 0.1099935 y 1.7828442 con una confianza del 95%, al estar incluido el 1 asumimos que las varianzas son iguales

## b)

Calculamos la diferencia de medias entre los dos equipos en segundos

t.test(x=team1, y=team2, var.equal=TRUE, conf.level = 0.95)$conf

## [1] -0.2420843 0.4636843  
## attr(,"conf.level")  
## [1] 0.95

La diferencia de medias poblacionales está entre -0.2420843 y 0.4636843 segundos con una confianza del 95%