

pueyo-joseluis-practica-n-2

Jose Luis Pueyo Viltres

20/11/2021

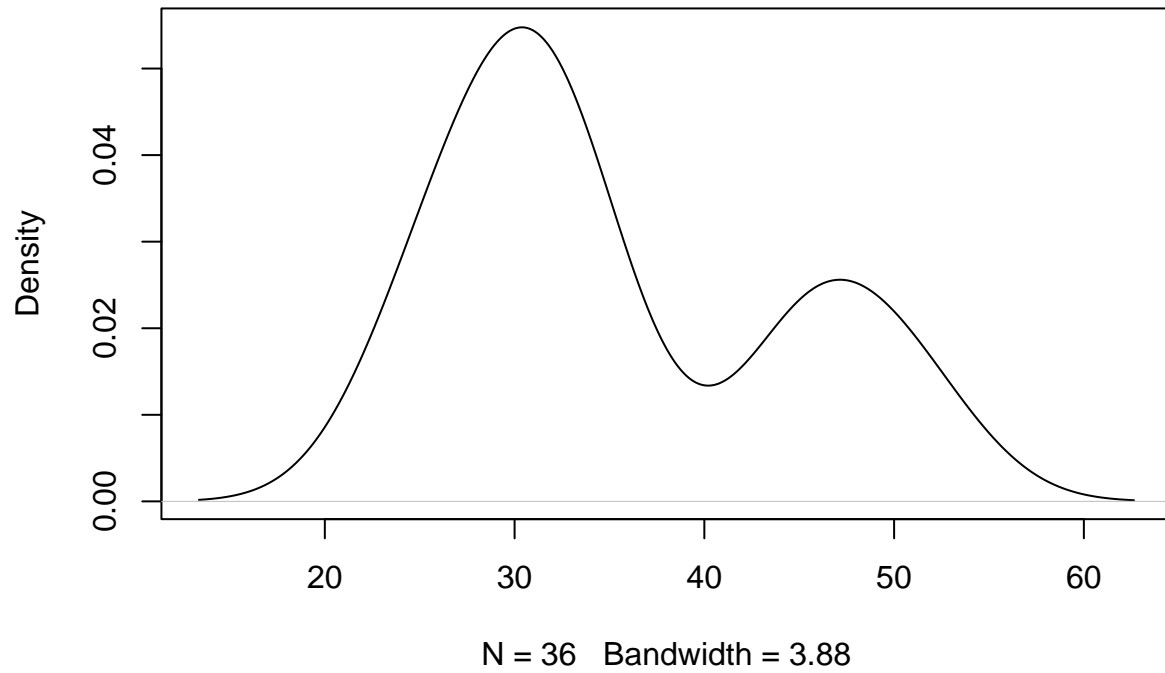
Cargo los datos a utilizar desde un archivo .csv

```
datos <- read.csv2("datos.csv", header = T)
datos
```

```
##   AMB.CO2 AMB.CO2.1 AMB.CO2.2 AMB.CO2.3 SENSE.CO2 SENSE.CO2.1 SENSE.CO2.2
## 1      45      26      30      25      23      25      30
## 2      26      25      45      34      30      23      40
## 3      25      34      26      51      49      30      25
## 4      34      51      25      46      41      49      23
## 5      51      46      34      32      35      41      30
## 6      46      32      51      30      28      35      49
## 7      32      30      46      46      41      28      41
## 8      30      32      32      32      35      35      35
## 9      30      30      30      30      28      28      28
##   SENSE.CO2.3
## 1           49
## 2           41
## 3           35
## 4           28
## 5           30
## 6           49
## 7           41
## 8           35
## 9           28
```

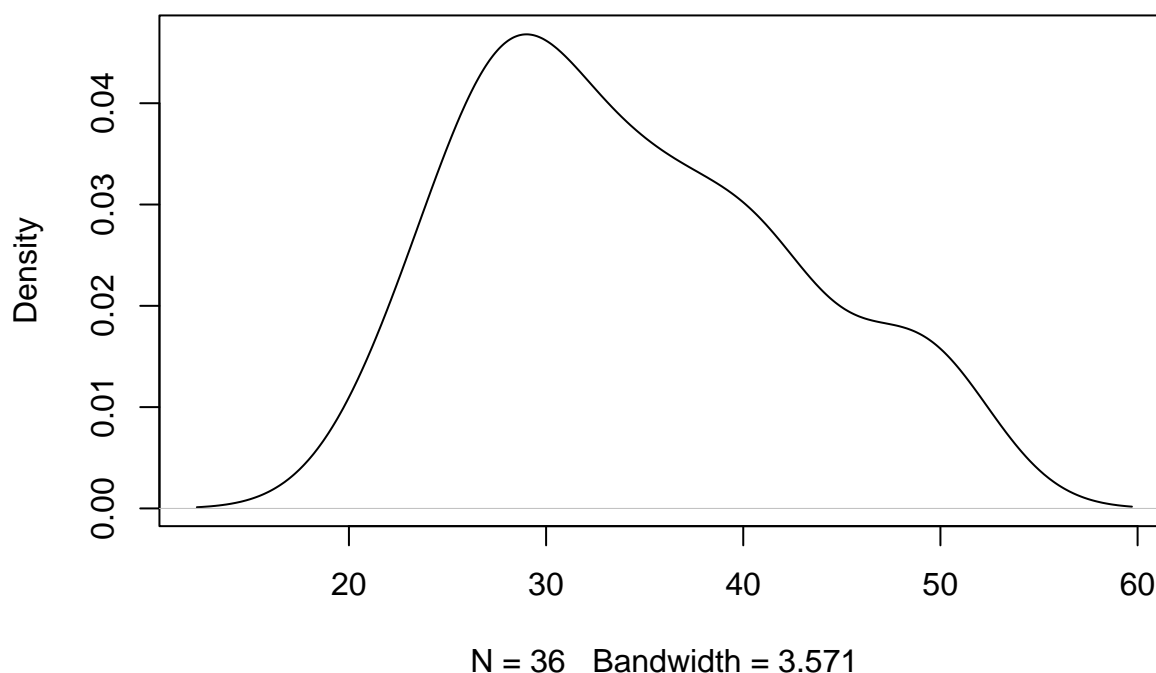
```
datosConC02 <- c(datos[,1],datos[,2],datos[,3],datos[,4])
datosSinC02 <- c(datos[,5],datos[,6],datos[,7],datos[,8])
plot(density(datosConC02))
```

density.default(x = datosConC02)



```
plot(density(datosSinC02))
```

density.default(x = datosSinC02)



Doneu estimacions de la mitjana de la freqüència respiratòria d'una persona, que està sotmesa a cadascun dels dos tipus d'atmosferes, i de les seves desviacions estàndards.

∅

Entre quins dos valors podeu assegurar que estarà la diferència de les freqüències respiratòries d'una persona en un 99% dels casos?

Para esto calculo el intervalo de confianza de $\mu_1 - \mu_2$

```
diferenciaMu <- abs(mean(x = datosConC02)-mean(x = datosSinC02))
t.test(x=datosSinC02,y=datosConC02,mu=diferenciaMu,conf.level = 0.99)$conf.int
```

```
## [1] -6.101426  4.490315
## attr("conf.level")
## [1] 0.99
```

Para calcular el intervalo de confianza he utilizado la función t.test, con el \$conf.int solo muestro el intervalo de confianza, que son los dos valores resultantes: [-6.101426,4.49.0315]

Podeu assegurar (teòricament i empíricament) que cadascun dels dos conjunts de dades segueixen una distribució de probabilitat normal?

∅

```
chisq.test(datosConC02)$p.value>0.01
```

```
## [1] FALSE
```

```
chisq.test(datosSinC02)$p.value>0.01
```

```
## [1] FALSE
```

Genereu una col·lecció de valors de mida 36 que representi una mostra d'una variable aleatòria amb distribució normal de mitjana 2 unitats més que la de la mostra de freqüències respiratòries sense CO i desviació la mateixa.

Genero la colección de datos con la función rnorm

```
require(stats)
normAleatoria = rnorm(n = 36, mean = mean(datosSinCO2)+2, sd = sd(datosSinCO2))
normAleatoria
```

```
## [1] 29.05394 33.01877 23.36526 35.16572 26.82039 21.58821 45.62430 42.62509
## [9] 31.22727 32.95515 37.37713 53.40078 42.51080 28.46818 46.31928 41.31440
## [17] 41.66578 31.48116 41.93564 34.00987 43.30891 46.97184 36.77580 36.22177
## [25] 29.26669 46.69119 47.58663 20.31244 24.98487 36.55923 31.45207 23.43475
## [33] 49.26488 41.57103 25.36114 38.59109
```

Oblidant-vos ara de com s'ha generat la mostra i dels resultats anteriors, considereu que les tres mostres s'ón independents i tenen una distribució normal amb la mateixa desviació. Podeu assegurar que les tres poblacions tenen la mateixa mitjana?

Para determinar si las 3 poblaciones tienen la misma media hago un test de anova

```
anova_datos = data.frame(datosConCO2, datosSinCO2, normAleatoria)
exp.abc <- stack(anova_datos)
anova(lm(values~ind, data=exp.abc))
```

```
## Analysis of Variance Table
##
## Response: values
##           Df Sum Sq Mean Sq F value Pr(>F)
## ind         2   45.6   22.787   0.3112 0.7332
## Residuals 105 7687.8   73.218
```

Como el $Pr(>F)$ es 0.8879 » 0.01 puedo aceptar con un 99% de confianza la H_0 , es decir, que las medias:

```
mean(normAleatoria)
```

```
## [1] 36.06337
```

```
mean(datosSinCO2)
```

```
## [1] 34.47222
```

```
mean(datosConCO2)
```

```
## [1] 35.27778
```

son iguales.

Considerem ara que totes les dades s'han obtingut del Carles i que cada fila correspon a les dades obtingudes en un mes concret que va del gener al setembre. Per exemple, al mes de gener s'ha mesurat 4 vegades la freqüència respiratòria del Carles amb una alta concentració de CO i s'han obtingut els valors 45, 26, 30 i 25. Contrasteu amb un test ANOVA si la freqüència respiratòria depèn de la concentració de CO? I depèn del mes de l'any? Potser depèn de la combinació dels dos factors?

∅