

Evaluación estadística

Ivonne Yáñez Mendoza

Módulo: Estadística

Profesor: Conrado Manuel García

Master big data & business analytics, Universidad Complutense de Madrid

31 de mayo de 2022

Ejercicio 1

Para resolver este ejercicio se han creado dos dataframes a partir del libro excel que contiene los datos tanto para el grupo predinástico temprano y el predinástico tardío.

Con estos datos por separado se han analizado por separado cada uno de estos dos periodos históricos, generando las respectivas mediciones y gráficos pertinentes.

Ejercicio 1.a

Se deben obtener las diferentes medidas y graficos de caja y bigotes para cada submuestra.

Apartado periodo predinástico temprano

Medidas de centralizacion, dispersión, asimetría y curtosis para el periodo predinástico temprano

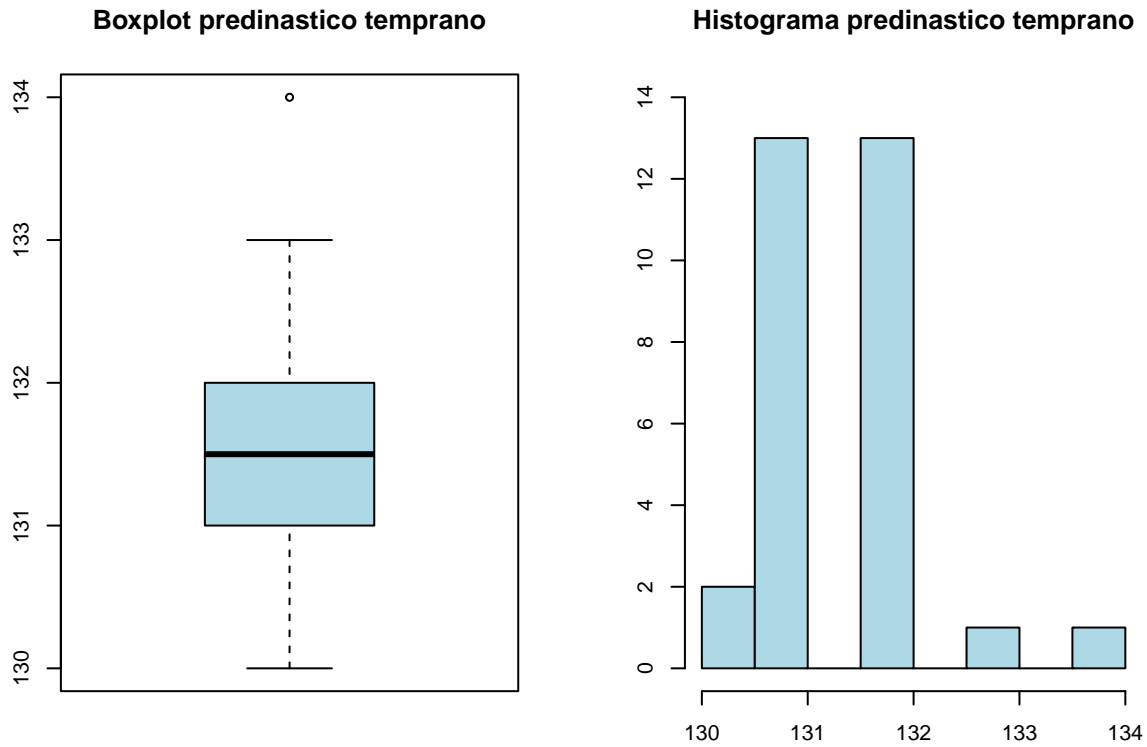
Table 1: Anchura de craneo periodo predinastico temprano

Epoca	Casos	Media	Mediana	Moda	Rango	Desviacion	Pearson	Fisher	Curtosis
1	30	131.5	131.5	131	4	0.82	0.0062	0.624	1.0222

Table 2: Cuartiles periodo predinástico temprano

Epoca	Casos	0%	25%	50%	75%	100%
1	30	130	131	131.5	132	134

Boxplot e histograma para periodo predinástico temprano



Observaciones apartado predinastico temprano

Apartado periodo predinástico tardio

Medidas de centralizacion, dispersión, asimetría y curtosis para el periodo predinástico tardio.

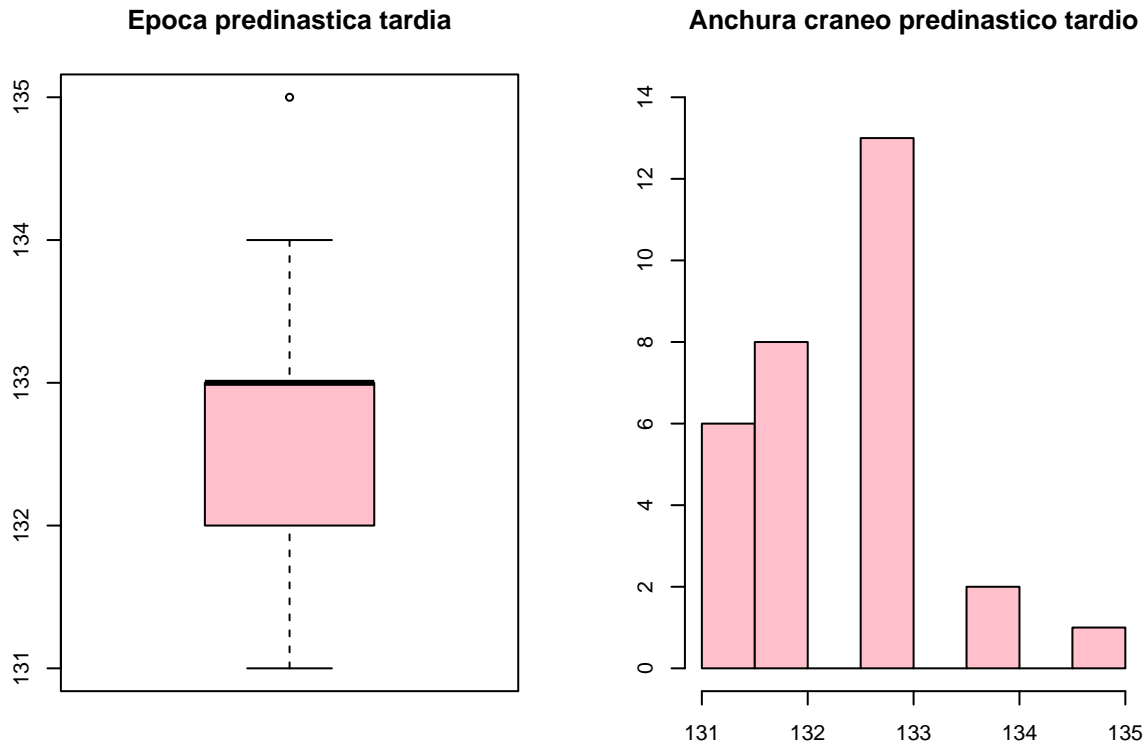
Table 3: Anchura de craneo periodo predinastico tardio

Epoca	Casos	Media	Mediana	Moda	Rango	Desviacion	Pearson	Fisher	Curtosis
2	30	132.5	133	133	4	1.01	0.0076	0.185	-0.3707

Table 4: Cuartiles periodo predinastico tardio

Epoca	Casos	0%	25%	50%	75%	100%
2	30	131	132	133	133	135

Boxplot e histograma para el periodo predinastico tardio



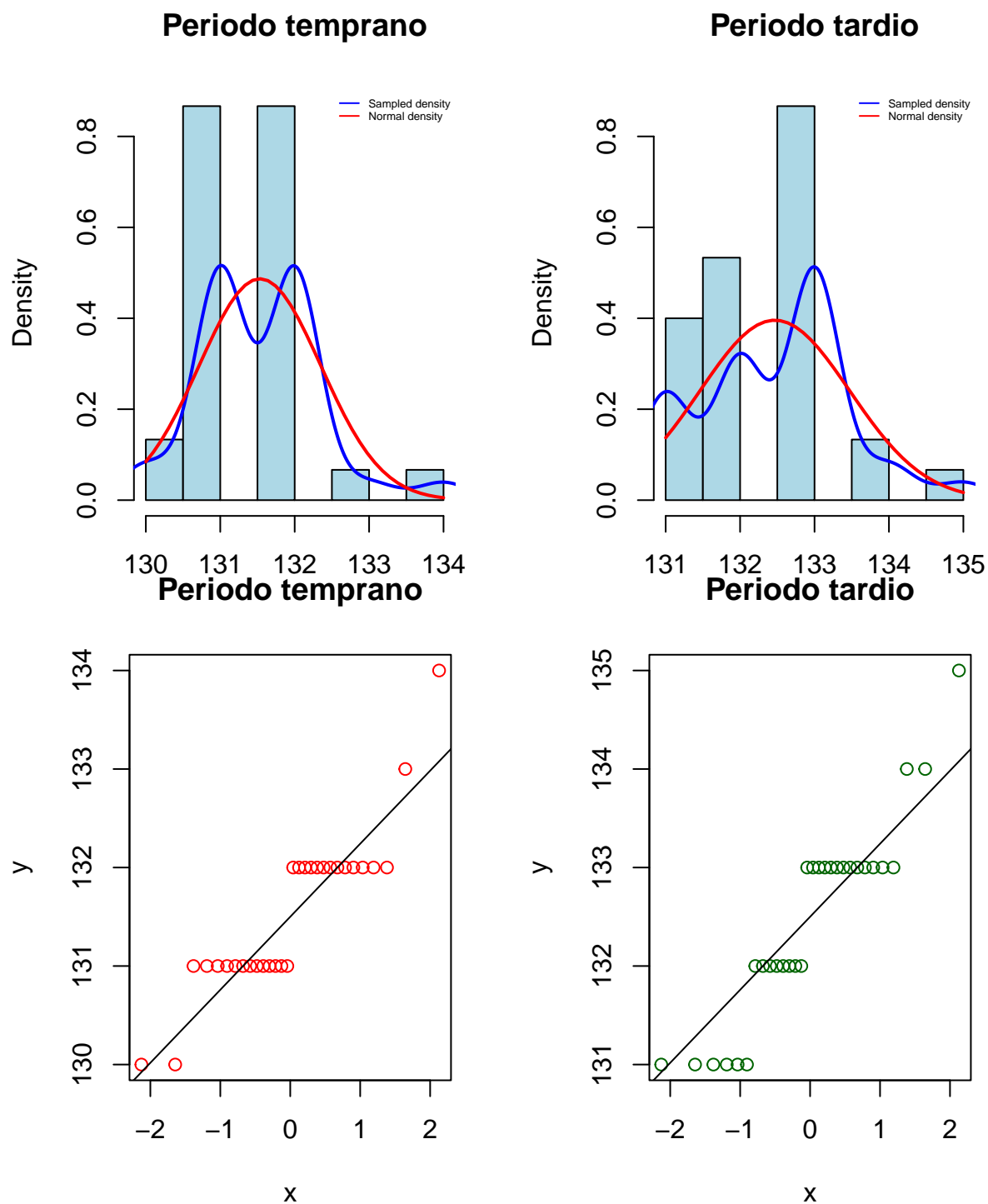
Observaciones apartado predinastico tardio

Ejercicio 1.b

Determinar si cada una de las dos submuestras sigue una distribucion normal utilizando el test de Kolmogorov-Smirnov.

En primer lugar se han creado graficos para ver de modo visual si hay distribucion normal, o no, de los datos procesados.

Estudio preliminar de la normalidad



Para este estudio en particular y al ser muestras relativamente pequeñas (30 casos por submuestra) se ha decidido utilizar el test de Shapiro Wilk para normalidad, obteniendo los siguientes resultados:

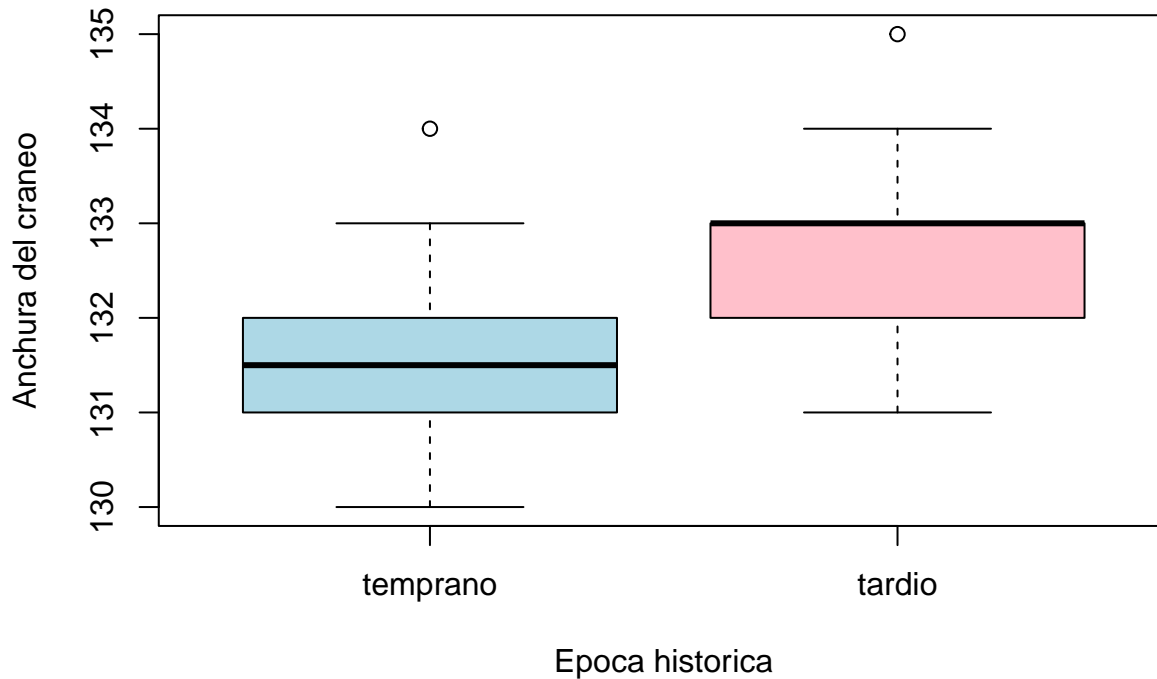
1. Periodo predinastico temprano

Shapiro-Wilk normality test

```
data: pre_temprano$Anchura.del.cráneo
W = 0.83781, p-value = 0.0003481
```

El p value en este caso entrega un valor de 0.0003481 y siendo este valor menor a 0.05 se decide rechazar la hipotesis nula, por lo tanto la muestra para el periodo predinastico temprano **no sigue una distribucion normal**.

igualdad de varianza



```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group 1  0.6195 0.4344
     58
```

->

Two Sample t-test

```
data: pre_temprano$Anchura.del.cráneo and pre_tardio$Anchura.del.cráneo
t = -3.9354, df = 58, p-value = 0.0002248
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.4080621 -0.4586046
sample estimates:
mean of x mean of y
 131.5333  132.4667
```