

Segunda práctica de Probabilidad y Estadística I  
Curso 2017-2018

Grupo A3 formado por:

1. Mark Dervishaj Memi
2. Angelika Krolikowska
3. Lucía Soler García
4. Sergio Ramos Alende

## 1. Modelo Probabilístico:

La variable *tiempos* recoge el tiempo de espera de 50 clientes en una pizzería hasta que llega la pizza a la mesa. El modelo probabilístico con el que vamos a trabajar es el modelo **normal**. Los parámetros de los que consta dicho modelo son **media (291.43)** y **desviación típica (36.73615)**. Sus estimadores puntuales toman el valor, calculados a partir de los datos, de **6.707158** y **4.742677**.

```
> library(MASS)
warning message:
package 'MASS' was built under R version 3.4.4
> fitdistr(mis_tiempos, "normal")
      mean      sd
291.433333  36.736615
( 6.707158) ( 4.742677)
```

A continuación comprobamos con el test de Kolmogorov-Smirnov que el modelo asignado es correcto. El p-valor obtenido para este contraste es **0.6981** y al ser mayor que 0.2 no rechazamos la hipótesis nula, con lo que no rechazamos el modelo asignado.

$H_0$ : Mis datos provienen de un modelo **binomial**.

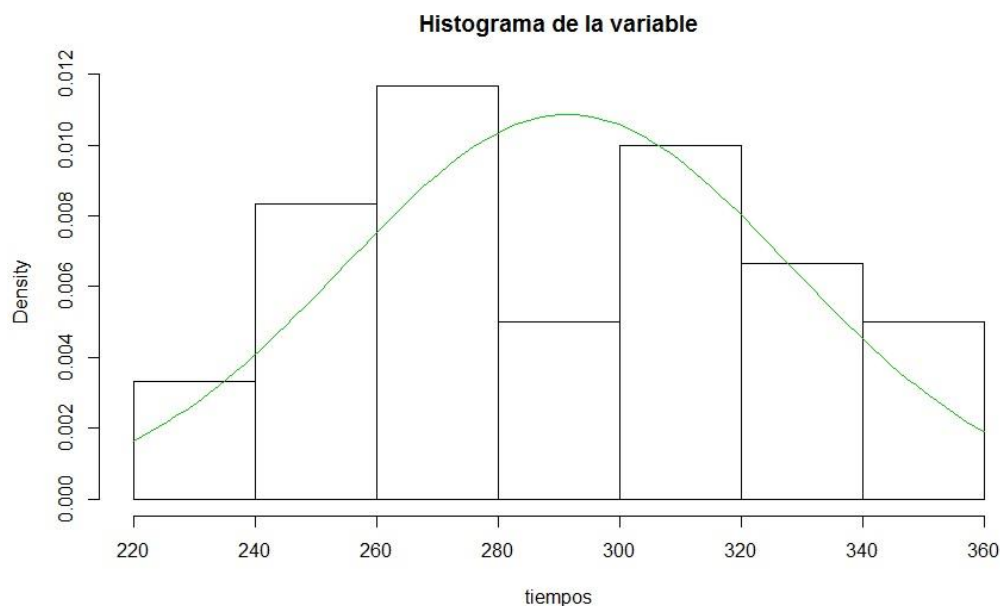
```
> ks.test(tiempos_1, "pnorm", 291.43, 36.74)

one-sample kolmogorov-smirnov test

data: tiempos_1
D = 0.12924, p-value = 0.6981
alternative hypothesis: two-sided

warning message:
In ks.test(tiempos_1, "pnorm", 291.43, 36.74) :
ties should not be present for the kolmogorov-smirnov test
```

En el gráfico siguiente comprobamos visualmente el ajuste de nuestros datos (histograma) a la función de densidad elegida.



2. Calculamos los intervalos de confianza al 98% para la media y para la varianza.

```
> install.packages("TeachingDemos")
Installing package into 'C:/Users/RESKOM326/Documents/R/win-library/3.4'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.4/TeachingDemos_2.10.zip'
Content type 'application/zip' length 1690056 bytes (1.6 MB)
downloaded 1.6 MB
```

package 'TeachingDemos' successfully unpacked and MD5 sums checked

The downloaded binary packages are in  
C:\Users\RESKOM326\AppData\Local\Temp\RtmpUXA7pF\downloaded\_packages

```
> library("TeachingDemos")
Warning message:
package 'TeachingDemos' was built under R version 3.4.4
> t.test(tiempos_1, conf.level = 0.98)
```

One Sample t-test

```
data: tiempos_1
t = 42.721, df = 29, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
98 percent confidence interval:
 274.6379 308.2288
sample estimates:
mean of x
 291.4333
```

```
> sigma.test(tiempos_1, conf.level=0.98)
```

One sample chi-squared test for variance

```
data: tiempos_1
X-squared = 40487, df = 29, p-value < 2.2e-16
alternative hypothesis: true variance is not equal to 1
98 percent confidence interval:
 816.477 2839.932
sample estimates:
var of tiempos_1
 1396.116
```

Podemos afirmar que la media poblacional se encuentra entre los valores 274.6379 y 308.2288 con una confianza del 98%. Con la misma confianza podemos afirmar que la varianza poblacional se encuentra entre 816.477 y 2839.932.

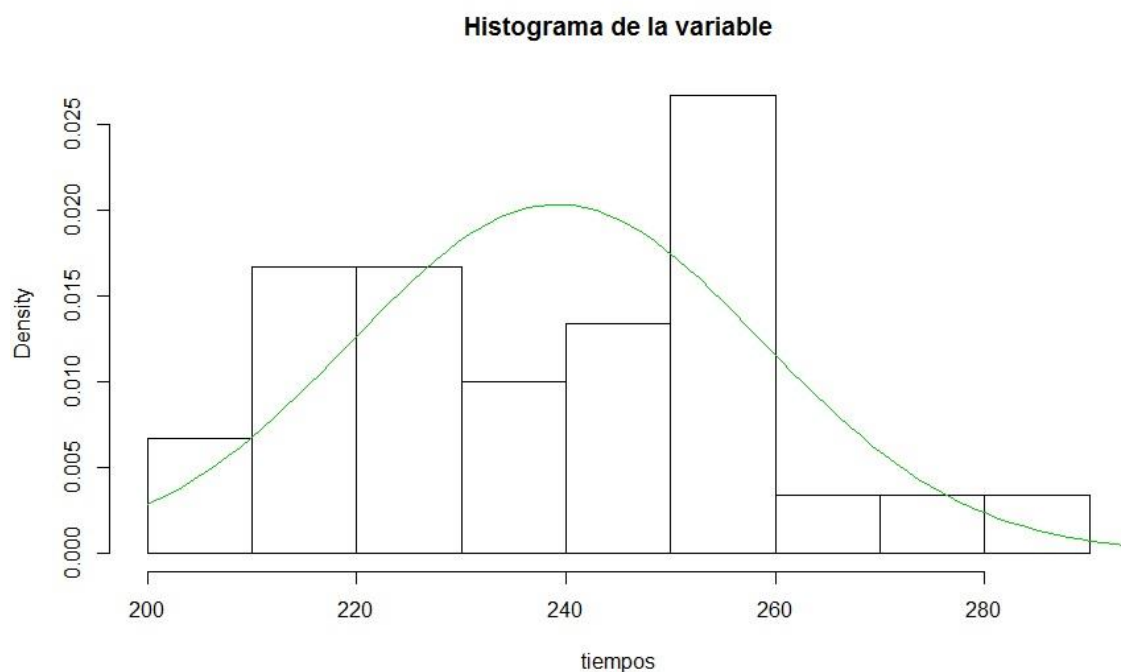
3. Generamos una nueva muestra de una distribución binomial con parámetros media (239.13) y desviación típica (19.63).

```
> library(MASS)
> fitdistr(mis_tiempos2, "normal")
      mean      sd
239.133333 19.629456
( 3.583832) ( 2.534152)
> hist(tiempos_2)
> ks.test(tiempos_2, "pnorm", 239.13, 19.63)

      one-sample kolmogorov-smirnov test

data:  tiempos_2
D = 0.16418, p-value = 0.3938
alternative hypothesis: two-sided

warning message:
In ks.test(tiempos_2, "pnorm", 239.13, 19.63) :
  ties should not be present for the kolmogorov-smirnov test
> summary(tiempos_2)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 208.0   222.5   239.5   239.1   255.5   281.0
> hist(tiempos_2, freq=F, main="Histograma de la variable", xlab="tiempos")
> curve(dnorm(x, 239.13, 19.63), from=200, to=310, add=TRUE, col=3)
```



Comparamos las varianzas y medias de la variable original y de la nueva:

a) Comparamos sus varianzas:

```
> var.test(tiempos_1,tiempos_2, ratio=1, alternative="two.sided",conf.level=0.98)

      F test to compare two variances

data:  tiempos_1 and tiempos_2
F = 3.5025, num df = 29, denom df = 29, p-value = 0.001169
alternative hypothesis: true ratio of variances is not equal to 1
98 percent confidence interval:
 1.445272 8.488165
sample estimates:
ratio of variances
      3.502529
```

En el contraste de comparación de varianzas, la hipótesis nula es **que las varianzas son iguales** y la hipótesis alternativa es **que son diferentes**. Se ha obtenido un valor  $\hat{d}$  de **3.502529** y un p-valor de **0.001169**. Por tanto, **se rechaza** la hipótesis nula y las muestras proceden de poblaciones con **diferente** varianza.

El IC al 95% para la razón de varianzas  $\frac{\sigma_1^2}{\sigma_2^2}$  es **(1.445272, 8.488165)**. En base a este intervalo, como el **valor 1 no está contenido** en él, debemos **rechazar** la hipótesis de igualdad de varianzas.

b) Comparamos sus medias:

```
> t.test(tiempos_1,tiempos_2, mu=0, alternative="two.sided",paired=FALSE,var.equal=FALSE,conf.level=0.98)

      welch Two Sample t-test

data:  tiempos_1 and tiempos_2
t = 6.7618, df = 44.311, p-value = 2.473e-08
alternative hypothesis: true difference in means is not equal to 0
98 percent confidence interval:
 33.63259 70.96741
sample estimates:
mean of x mean of y
 291.4333  239.1333
```

En el contraste de comparación de medias, la hipótesis nula es que **son iguales** y la hipótesis alternativa es que **son distintas**. Se ha obtenido un valor  $\hat{d}$  de **6.7618** y un p-valor de **2.473e-08**. Por tanto, **rechazamos** la hipótesis nula y las muestras proceden de poblaciones con **distinta** media.

El IC al 95% para la diferencia de medias  $\mu_1 - \mu_2$  es **(33.63259, 70.96741)**. En base a este intervalo, como el **valor 0 no está contenido** en él, debemos **rechazar** la hipótesis de igualdad de medias.