

# INFERENCIA ESTADÍSTICA

## MÁSTER BIOINFORMÁTICA

## Y BIOESTADÍSTICA

Universitat de Barcelona - Universitat Oberta de Catalunya

### PRUEBA DE EVALUACIÓN CONTÍNUA 2

*Santiago Royuela Samit*

*Junio de 2022*

#### 1.-Enunciados teóricos

Indica que tipo de análisis o que pruebas estadísticas utilizarías en cada uno de los apartados y si fuera necesario algún tipo de prueba adicional, ¿Qué harías para llevar a cabo el análisis? Formula la hipótesis a contrastar de acuerdo con las preguntas a responder. Justifica la elección. (La solución puede no ser única)

- a) **Se dispone de 500 pacientes asignados aleatoriamente a un a dieta baja en calorías y otros 500 pacientes a un grupo de dieta mediterránea. Se desea comparar, al cabo de un tiempo, la aparición de algún evento cardiovascular (infarto, accidente cerebrovascular, enfermedad arterial o muerte cardiovascular). Indica la prueba más adecuada para comparar la aparición de un evento cardiovascular según el grupo de dieta aleatorizado.**

Por un lado, tenemos **2 grupos distintos**, según su **dieta**, y que además son del mismo tamaño **N=500**. Hemos de comparar entre ellos la aparición de algún evento cardiovascular, que son 4 (Infarto, Accidente Cerebrovascular, Enfermedad Arterial o Muerte Cardiovascular). Podríamos limitar el análisis comparativo de grupos al hecho simple de **si aparece o no** "un evento cardiovascular", sin atender a su tipología.

De otro lado, podemos analizar si existe una **dependencia** entre la "**variable dieta**" y la "**variable evento cardiovascular**" mediante un **test de independencia Ji-Cuadrado**. En caso de existir, analizaríamos las proporciones por cada categoría de las variables y determinaríamos **si existe o no diferencia de proporciones en cada uno de los tipos de dolencia cardiaca**.

Estamos ante un **estudio de cohortes**, es decir, dos grupos que se dividen por un **factor de exposición**, que es la **dieta** en nuestro caso. Vamos a elaborar una

**tabla de contingencia**, que es una tabla de doble entrada de frecuencias y bidimensional.

OCURRE EVENTO?	SI	NO
DIETA BAJA (A)	Frecuencias absolutas: $(n_{1,1})$	Frecuencias absolutas: $(n_{1,2})$
DIETA MEDITERRÁNEA (B)	Frecuencias absolutas: $(n_{2,1})$	Frecuencias absolutas: $(n_{2,2})$

Analizaremos el **riesgo relativo**, que es el de **padecer un evento cardiovascular según el grupo**, es decir, comparara el riesgo en los dos grupos, donde es mayor, menor o si es igual.

(Mirar si distribución Poisson y aproximación a Ji-Cuadrado. Mirar también que, al aproximar la proporción y obtener un estadístico tipo Z –más adecuado para contrastes unilaterales–, si lo elevamos al cuadrado coincide con el estadístico Ji-Cuadrado de independencia – más adecuado a contraste bilaterales-. Se asume normalidad en el estadístico de contraste dado el tamaño de las muestras y el T.L.C.).

Supongamos que tenemos las frecuencias absolutas obtenidas tras el estudio:  $n_{i,j}$  (i, fila: Tipo Dieta ; j, columna: Si hay o no evento). Ahora nos centraríamos en los casos en los que se ha producido un **evento cardiovascular**, tanto en un **grupo** como en el otro, para calcular la **proporción de casos de eventos cardiovasculares** en cada uno de ellos:

$\hat{p}_a = \frac{n_{1,1}}{n_{1.}}$ , que sería la proporción de casos de evento cardiovascular en la “**Dieta Baja**”. De otro lado:  
 $\hat{p}_b = \frac{n_{2,1}}{n_{2.}}$  será la proporción de casos de evento cardiovascular en la “**Dieta Mediterránea**”.

$$\text{Siendo } n_{i.} = \sum_{j=1}^2 n_{i,j}$$

Desconocemos el modelo probabilístico que rige cada grupo de dieta, siendo que, en el ensayo, precisamente se miden los posibles efectos de las dietas sobre los eventos cardiovasculares. De esta manera, de existir una probabilidad determinada bajo una dieta concreta de presentar o no un “**evento cardiovascular**”, nos preguntaremos si es la misma bajo cada una de las dietas, o si son diferentes.

Cada grupo de dieta, A o B, consta de 500 sujetos, siendo que, el hecho de que presenten un “**evento cardiovascular**”, vendrá dado, en cada grupo, por una distribución de probabilidad Binomial (también sería la suma de 500 distribuciones de Bernoulli):

$$X_A \sim \text{Bin}(p_a, N = 500) \text{ , Dieta baja en calorías.}$$

$$X_B \sim \text{Bin}(p_b, N = 500) \text{ , Dieta Mediterránea.}$$

Como para ambos grupos tenemos que  $N=500$ ,  $n_T=1000$ , entonces, por el **Teorema del Límite Central** podemos aproximar dichas binomiales por distribuciones normales: *Notación:  $N(\mu, \sigma^2)$*

$$X_A \sim N(500 \cdot p_A, 500 \cdot p_A(1 - p_A))$$

$$X_B \sim N(500 \cdot p_B, 500 \cdot p_B(1 - p_B))$$

Como tomamos proporciones y dividimos los casos de evento cardiovascular por  $N=500$  para estimar las probabilidades o frecuencias  $p$ 's (proporciones), éstas seguirán una distribución normal dada por: Notación:  $N(\mu, \sqrt{\sigma})$

$\hat{p}_A \sim N\left(p_A, \sqrt{\frac{p_A(1-p_A)}{500}}\right)$ , distribución para la **proporción de casos de evento cardiovascular** en el grupo de **Dieta baja en calorías**.

$\hat{p}_B \sim N\left(p_B, \sqrt{\frac{p_B(1-p_B)}{500}}\right)$ , distribución para la **proporción de casos de evento cardiovascular** en el grupo de **Dieta Mediterránea**.

Ahora podemos hacer un **contraste de hipótesis** sobre la **diferencia de proporciones** en ambos grupos según los “**eventos cardiovasculares**”. También podríamos hallar un **intervalo de confianza** para la **diferencia de proporciones** en los grupos con un **nivel de confianza  $(1-\alpha)$  determinado**.

$$H_0: p_A = p_B$$

Como hipótesis alternativa podríamos plantear diversos casos, bien un contraste bilateral, o 2 unilaterales, según si consideramos que  $p_A > p_B$  o  $p_A < p_B$

$H_1: p_A \neq p_B$  Bilateral.

$H_1: p_A > p_B$  Unilateral.

$H_1: p_A < p_B$  Unilateral.

Si la hipótesis nula es cierta, entonces tenemos el siguiente estadístico de contraste, pues  $p_A = p_B$ :

$$z = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

$$\hat{p} = \frac{n_1\hat{p}_A + n_2\hat{p}_B}{n_1 + n_2}$$

Se asume que la probabilidad es igual en ambos grupos y, por tanto, también lo serán sus varianzas, siendo así que no procede un contraste de varianzas.

Con el error estándar:  $s_{\hat{p}} = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

Ahora, una vez tenemos el estadístico y según el tipo de contraste, bilateral o unilateral, procederíamos a calcular el  $p$ -valor para ver si aceptamos o rechazamos la hipótesis nula:

$H_1: p_A \neq p_B$  ;  $p = 2P(N(0,1) > |z|)$

$H_1: p_A > p_B$  ;  $p = P(N(0,1) > z)$

$H_1: p_A < p_B$  ;  $p = P(N(0,1) < z)$

Fijaríamos un nivel de significación  $\alpha$  y lo compararíamos con el  **$p$ -valor**. Si el  **$p$ -valor** es mayor, no se rechaza la hipótesis nula, y sí se hace en caso contrario, aceptando la hipótesis alternativa.

Podemos construir un intervalo de confianza con un nivel de confianza para la diferencia de proporciones. Para ello:

- Primero fijamos el **nivel de confianza (1-α)**.
- Calculamos el error estándar de la diferencia de proporciones:

$$s_{\hat{p}} = \sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$\hat{p} = \frac{n_1\hat{p}_A + n_2\hat{p}_B}{n_1 + n_2}$$

- Calculamos el valor crítico de la N(0,1), según suponemos se distribuye el estadístico bajo la hipótesis nula, y para el nivel de confianza fijado.
- Calculamos el margen de error a partir del valor crítico y el error estándar.

El **intervalo de confianza** vendrá dado por:

$$(\hat{p}_A - \hat{p}_B) \pm z_{\alpha/2} \cdot s_{\hat{p}}$$

De esta manera, obtendríamos un intervalo con un nivel de confianza dado en el que, bajo la hipótesis nula, en un **(1-α)%** de muestreos, el estadístico se hallaría en dicho intervalo.

### Test de independencia Ji-Cuadrado

Podemos realizar un **test de independencia Ji-Cuadrado**, basado en el estadístico de Pearson, para evaluar la hipótesis con respecto a una posible relación de dependencia entre variables categóricas, los dos tipos de dieta. La hipótesis nula establece que las dos variables aleatorias categóricas son independientes, y la alternativa que no lo son.

Para ello elaboramos las siguientes tablas en base a las frecuencias absolutas obtenidas.

**Tabla 1. TABLA DE OBSERVADOS**

TIPO EVENTO	INFARTO	CEREBROVASCULAR	ARTERIAL	MUERTE	TOTALES
DIETA BAJA (A)	Frecuencias absolutas: (n <sub>1,1</sub> )	Frecuencias absolutas: (n <sub>1,2</sub> )	Frecuencias absolutas: (n <sub>1,3</sub> )	Frecuencias absolutas: (n <sub>1,4</sub> )	n <sub>1.</sub>
DIETA MEDITERRÁNEA (B)	Frecuencias absolutas: (n <sub>2,1</sub> )	Frecuencias absolutas: (n <sub>2,2</sub> )	Frecuencias absolutas: (n <sub>2,3</sub> )	Frecuencias absolutas: (n <sub>2,4</sub> )	n <sub>2.</sub>
<b>TOTALES</b>	n <sub>.1</sub>	n <sub>.2</sub>	n <sub>.3</sub>	n <sub>.4</sub>	n <sub>..</sub>

**Tabla 2. TABLA DE ESPERADOS**

TIPO EVENTO	INFARTO	CEREBROVASCULAR	ARTERIAL	MUERTE	TOTALES
DIETA BAJA (A)	Frecuencias $\frac{(n_{1.}) \cdot n_{.1}}{n_{.}}$	Frecuencias $\frac{(n_{1.}) \cdot n_{.2}}{n_{.}}$	Frecuencias $\frac{(n_{1.}) \cdot n_{.3}}{n_{.}}$	Frecuencias $\frac{(n_{1.}) \cdot n_{.4}}{n_{.}}$	n <sub>1.</sub>
DIETA MEDITERRÁNEA (B)	Frecuencias $\frac{(n_{2.}) \cdot n_{.1}}{n_{.}}$	Frecuencias $\frac{(n_{2.}) \cdot n_{.2}}{n_{.}}$	Frecuencias $\frac{(n_{2.}) \cdot n_{.3}}{n_{.}}$	Frecuencias $\frac{(n_{2.}) \cdot n_{.4}}{n_{.}}$	n <sub>2.</sub>
<b>TOTALES</b>	n <sub>.1</sub>	n <sub>.2</sub>	n <sub>.3</sub>	n <sub>.4</sub>	n <sub>..</sub>

Si siguiendo con el test de **independencia Ji-Cuadrado**, basado en el hecho de que **la probabilidad conjunta para 2 variables independientes es igual al producto de sus probabilidades individuales**, y utilizaremos esta regla para encontrar las frecuencias esperadas.

A partir de las tablas podemos construir el estadístico  $\chi^2$  de Person:

$$Q = \sum_{i=1}^2 \sum_{j=1}^4 \left( \frac{O_{i,j} - E_{i,j}}{E_{i,j}} \right)^2$$

$$O_{i,j} = \frac{n_{i,j}}{n_{i,\cdot}}$$

$$E_{i,j} = \frac{(n_{i,\cdot}) \cdot n_{\cdot,j}}{n_{\cdot,\cdot}}$$

El **estadístico Q**, bajo la hipótesis nula de independencia en las 2 variables categóricas o grupos, se distribuirá como una **Ji-Cuadrado** con  $(2-1) \times (4-1) = 3$  **grados de libertad**:  $\chi^2_3$

Podemos observar el nivel de significación observado buscando la cola superior de probabilidad del valor observado para Q, que denotaremos por q, basándose en la distribución  $\chi^2_3$ . El **test Ji-Cuadrado** nos dirá **si existe una dependencia** entre el **tipo de dieta asignada y el evento cardiovascular**, pero **no con el tipo de evento en concreto**. En caso de existir dependencia en el tipo de dieta y el hecho de un **evento cardiovascular**, se procede a un análisis de **contraste de proporciones entre los dos grupos** para cada **tipo de accidente** o, **dentro de una dieta**, mirar **contrastos de proporciones entre tipos de evento de accidente cardiovascular**, siempre trabajando con las muestras de los grupos a comparar; utilizando el método visto al principio del ejercicio entre los grupos que queramos contrastar.

Ejemplo: Hacemos un contraste de proporciones para los eventos cardiovasculares de tipo Infarto en los grupos de dieta, hipocalórica y mediterránea:

$$\hat{p}_C \sim N \left( \frac{n_{1,1}}{n_{\cdot,1}}, \sqrt{\frac{\frac{n_{1,1}}{n_{\cdot,1}} \left( 1 - \frac{n_{1,1}}{n_{\cdot,1}} \right)}{n_{\cdot,1}}} \right), \text{ distribución para la proporción de casos de evento cardiovascular INFARTO en el grupo de Dieta baja en calorías.}$$

$$\hat{p}_D \sim N \left( \frac{n_{2,1}}{n_{\cdot,1}}, \sqrt{\frac{\frac{n_{2,1}}{n_{\cdot,1}} \left( 1 - \frac{n_{2,1}}{n_{\cdot,1}} \right)}{n_{\cdot,1}}} \right), \text{ distribución para la proporción de casos de evento cardiovascular INFARTO en el grupo de Dieta Mediterránea}$$

El estadístico estandarizado será:

$$z = \frac{\hat{p}_C - \hat{p}_D}{\sqrt{\hat{p}(1 - \hat{p}) \left( \frac{1}{n_{1,1}} + \frac{1}{n_{2,1}} \right)}} \sim N(0,1)$$

$$\hat{p} = \frac{n_{1,1}\hat{p}_C + n_{2,1}\hat{p}_D}{n_{\cdot,1}}$$

Calculamos el error estándar de la diferencia de proporciones:

$$s_{\hat{p}} = \sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_{1,1}} + \frac{1}{n_{2,1}}\right)}$$

El **intervalo de confianza** vendrá dado por:

$$(\hat{p}_C - \hat{p}_D) \pm z_{\alpha/2} \cdot s_{\hat{p}}$$

De esta manera, obtendríamos un intervalo con un nivel de confianza dado en el que, bajo la hipótesis nula, en un **(1- $\alpha$ )%** de muestreos, el estadístico se hallaría en dicho intervalo. Así podríamos ir comparando proporciones entre los grupos pertinentes que deseemos contrastar.

- b) A los 90 sujetos con evento cardiovascular asignados al grupo de dieta mediterránea se les midió un biomarcador cuantitativo basado en la expresión de algunos genes en el momento del diagnóstico y a los 6 meses. Indica la prueba más adecuada para evaluar si el biomarcador ha cambiado en los seis meses.**

En este caso, estamos ante un **diseño de dos muestras relacionadas** (apareadas), ya que todos los sujetos pasan por las **dos condiciones** (**pre**=momento diagnóstico y **post**=6 meses). La **variable dependiente** es el **biomarcador cuantitativo**. Las opciones que tenemos son: aplicar **pruebas paramétricas** o **no paramétricas**.

- **Prueba paramétrica:** Para poder aplicar esta prueba, la variable diferencia (d=pre-post) **debe seguir aproximadamente una distribución normal**. Para comprobar la normalidad se puede aplicar un **test de normalidad** como el **test de Shapiro-Wilks** o el de **Kolmogorov-Smirnoff**. La prueba paramétrica correspondiente, al no conocerse la varianza poblacional, es la **prueba t-student** (desconocemos la varianza) **para datos emparejados**:

- o Hipótesis:

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

- o Estadístico de contraste:

$$t = \frac{\bar{d}}{\frac{\hat{s}_d}{\sqrt{n}}} \quad gl = n - 1 = 89$$

$\hat{s}_d$  Es la **desviación estándar muestral** (dividida por n-1, si fuera la **segada** dividiría en la formula anterior por n-1)

- o *p - valor*

- Si p-valor >  $\alpha$ , mantenemos hipótesis nula
- Si p-valor <  $\alpha$ , rechazamos hipótesis nula y aceptamos hipótesis alternativa

- Probabilidad de las 2 colas para el cálculo del p-valor:

$$p = P(|t_{n-1}| > [t]) = P(t_{n-1} > [t]) + P(t_{n-1} < -[t])$$

- o Intervalo de Confianza con nivel de confianza (1- $\alpha$ ), o el resultado del contraste, vendrá dado por (contraste bilateral):

- Si  $|t| \leq T_{Teo} \rightarrow$  No rechazamos  $H_0$
- Si  $|t| \geq T_{Teo} \rightarrow$  Rechazamos  $H_0$  y aceptamos  $H_1$

$$T_{Teo} = t_{n-1, 1-\alpha/2}$$

- **Prueba no paramétrica:** En caso de que la variable “d” no se distribuya de forma normal podemos optar por **Bootstrap**, prueba de **Wilcoxon de rangos signados** o **prueba de permutación**.

c) Se diseña un estudio de prueba de concepto en el que se disponen de tres grupos de 7 pacientes, uno con el tratamiento experimental a dosis bajas, otro con el tratamiento experimental a dosis altas y un tercero en placebo. Para evaluar la eficacia se mide la reducción del nivel de un biomarcador hepático al final del estudio. Indica el análisis que efectuarías.

En este caso, estamos ante un **diseño de tres muestras independientes** donde el **factor intersujetos de efectos fijos** es el “**Tipo de tratamiento**” con **tres niveles** (Dosis baja, Dosis alta y Placebo). La **variable dependiente** es la “**Reducción en el nivel de biomarcador**” que asumimos es una **variable cuantitativa**. Las opciones que tenemos son: aplicar **pruebas paramétricas** o **no paramétricas**.

- **Prueba paramétrica:** Para poder aplicar esta prueba, la variable dependiente debe seguir aproximadamente una **distribución normal** en **cada una de las poblaciones** y las **varianzas de las mismas deben ser homogéneas** – homocedasticidad- (iguales en la población las 3 varianzas. **3 poblaciones normales con igual varianza y que sus medias pueden o no ser iguales**).

Para **poner a prueba la normalidad** podemos emplear el **test de Shapiro-Wilks** o el de **Kolmogorov-Smirnoff**. Para poner a **prueba la homogeneidad de varianzas** podemos utilizar el **test de Levene**. La **prueba paramétrica a emplear** es el **ANOVA de un factor de efectos fijos** (balanceado o equilibrado porque cada grupo es del igual tamaño), aunque **tenemos muy pocos datos para cada tipo de grupo**, lo que nos llevará a optar por una prueba no paramétrica. Sin embargo, veremos **cómo se plantearía el ANOVA** y luego veremos **unas pruebas no paramétricas**:

G1: Dosis baja

G2: Dosis alta

G3: Placebo

- Hipótesis nula:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

Hipótesis alternativa:

$$H_1: \text{Al menos dos medias distintas}$$

- Estadístico de contraste (F de Snedecor):

$$F_{k-1, n-k} = \frac{\frac{SCE}{k-1}}{\frac{SCD}{n-k}} \quad \text{Grados de libertad:}$$

$$gl_{\text{numerador}} = k - 1 \quad gl_{\text{denominador}} = n - k$$

SCE es la suma cuadrática entre grupos

$SCE = \sum_{j=1}^3 n_j \cdot (\bar{x}_j - \bar{x})^2$  (Número de cada grupo  $n_j=7$ , media en cada grupo y la media total o gran media)

*SCD es la suma cuadrática dentro de los grupos o de errores o intragrupos.*

$$SCD = \sum_{j=1}^3 \left( \sum_{i=1}^7 (x_{j,i} - \bar{x}_j)^2 \right)$$

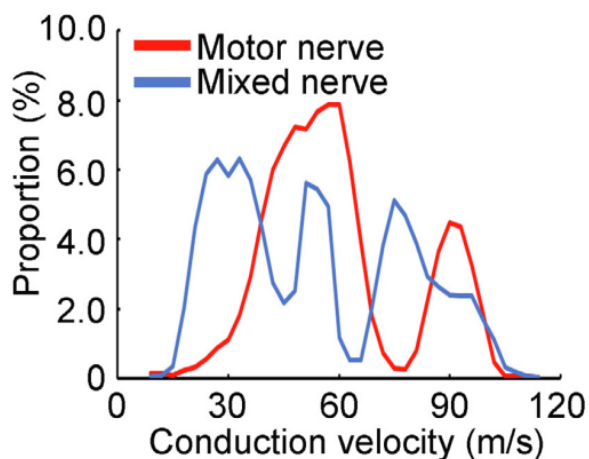
**NOTA:** el ANOVA siempre es un contraste unilateral derecho, la zona crítica siempre se sitúa en la cola derecha de la distribución.

○ *p - valor*

- Si  $p\text{-valor} > \alpha$ , mantenemos hipótesis nula
- Si  $p\text{-valor} < \alpha$ , rechazamos hipótesis nula y aceptamos hipótesis alternativa

- **Prueba no paramétrica:** En caso de que **no se cumpla la normalidad o la homogeneidad de varianzas –solo tenemos 7 datos por grupo–** podemos optar por un **Bootstrap**, prueba de **Kruskal-Wallis** o prueba de **Permutación**.

- c) Se efectúa un ensayo aleatorio doble ciego para evaluar el efecto del trandolapril, sobre la neuropatía diabética, asignando a 23 pacientes el trandolapril y a otros 23 a un placebo. Como respuesta entre otras variables se midió la velocidad de conducción motora del nervio del peroneo(m/s) cuya distribución se muestra en la figura (motor nerve). Indica el análisis para comparar el nivel de dicha variable al finalizar el estudio.



En este caso, tenemos un diseño de **dos muestras independientes**, siendo la variable independiente "**Tipo de tratamiento**" con dos niveles (G1: **Trandolapril**, G2: **Placebo**) y la variable dependiente la velocidad de **conducción motora del nervio del peroneo** (m/s) que es una variable cuantitativa.



Debemos fijarnos en la curva roja, pues la azul hace referencia a otro nervio. En este caso, estamos ante un **diseño de dos muestras independientes** donde el **factor intersujetos de efectos fijos** es el “**Tipo de tratamiento**” con **dos niveles** (Trandolapril y Placebo).

A pesar de que la curva azul no es objeto del estudio, ya nos indica que se forman 3 grupos o una distribución multimodal, con 3 picos o modas. Si nos fijamos en **la curva roja, objeto del estudio**, observamos que nos indica **proporciones en tanto por 1**. En principio, parece existir 2 modas, es decir, una distribución bimodal con **2 grupos diferenciados o tendencias**. Sin embargo, deberemos seguir con nuestro análisis para profundizar en la información que podemos extraer del gráfico, pues observamos que la primera campana roja, la más grande, muestra una ligera inflexión mostrando lo que pudieran ser dos máximos, mientras que la segunda, la más alejada, se distribuye conforme a una normal con una moda en 90 m/s.

La **primera campana** parece tener **dos modas o picos**, una en torno a 45 m/s y otra en 60 m/s. De otro lado, observamos que en el grupo de la **primera campana** se agrupan más sujetos que en la **segunda campana**. Observamos que la primera campana, con dos modas o picos, se distribuye de 15 m/s hasta unos 75 m/s, mientras que la **segunda campana** va de 75 m/s (mínimo menor) hasta 105 m/s aproximadamente.

Hemos de pensar que, al tener **2 grupos de igual tamaño**, 23 individuos en cada uno, la campana mayor, la primera, debe agrupar a los sujetos que han recibido placebo, puesto que no tendría sentido que en **trandolapril** hiciera efecto más allá del número de sujeto a los que se les ha suministrado. Lo lógico es pensar que, quizá, no siempre en medicamento hace efecto en un sujeto, cosa que le llevará a un resultado propio del grupo placebo.

De hecho, observando la anchura de las dos campanas y sus alturas o proporciones, podemos decir que la segunda campana, centrada en 90 m/s, representa una tercera parte de la primera campana, unos 7 u 8 sujetos aproximadamente. Está claro, que esta segunda campana estará asociada a sujetos que han tomado **trandolapril** y que, el resto de sujetos que también lo han tomado, se han distribuido dentro de la primera campana junto a los sujetos que han tomado placebo. Pero este efecto nos lleva a concluir que, si bien los sujetos que han tomado placebo se han distribuido normalmente, los que han tomado **tandrolapril no cumplen la condición de normalidad**, puesto que parte de ellos se distribuyen normalmente en la segunda campana y la gran mayoría en la primera campana junto a los sujetos placebo.

Es así que la cola izquierda de la segunda campana (sujetos que seguro han tomado **trandolapril**) conecta con la cola derecha de la **primera campana** sin sumar a esta última y alcanzando un mínimo de 0. Se deduce pues que, la “**distorsión**” observada en la **primera campana** es debido a la **suma de dos distribuciones normales**, la de los **sujetos placebo**, la del pico mayor y centrada en 60 m/s, a la que se ha sumado otra procedente de **sujetos trandolapril** y centrada en 45 m/s, provocando una **campana donde se observan dos máximos o modas** dentro de lo que hemos llamado “**la primera campana**” o **agrupación**”.

Por tanto, los 23 sujetos que han tomado **trandolapril** se han distribuido entorno a dos distribuciones normales, una mayoritaria y centrada en en 45 m/s, y que se suma a la curva normal que forman los sujetos placebos, centrada en 60 m/s; los otros trandolapriles se distribuyen normalmente en “**la segunda campana**” (¡que es la 3ª ahora!) centrada en 90 m/s.

Esto nos lleva a concluir que hay una variable no controlada en los pacientes que han tomado **trandolapril** que tiene un efecto notable sobre la velocidad de conducción del nervio en cuestión. Factor que pudiera ser de carácter cualitativo o mediante el que podemos agrupar a los sujetos en dos conjuntos o subgrupos, en tanto que distribuye a la muestra en **dos grupos claramente diferenciados** y distribuidos normalmente con medias distintas y desviaciones típicas que pudieran ser o no iguales.

Una vez sabemos los sujetos de la campana **última campana**, que es de sólo **trandolapriles**, podemos acudir a la suma de campanas que conforma la primera campana y hacer un estudio de contraste de proporciones en ese grupo que conforma la "**primera campana**" (suma de 2 campanas cercanas), dividiéndolo en los que han tomado placebo y los que han tomado **trandolapril** y efectuando un **contraste de proporciones** para saber si, lo que **parece ser dos campanas que se suman, lo son realmente o no en el sentido estadístico**, pues objetivamente sí sabemos que **unos han sido medicados y otros no**, pero podríamos concluir si el **trandolapril**, en mayoría - y salvo un efecto no controlado que nos separa a los individuos **trandolapril** en dos campanas, una menor y claramente con un efecto bien diferenciado sobre la muestra de placebos - **tiene efecto o no en los pacientes**.

Claramente **hay un factor no controlado en los sujetos trandolapril** que sí provoca un **efecto evidente frente a los sujetos placebo**, por tanto ¿ese efecto lo causa el trandolapril, o es causado por ese factor no controlado y "desconocido"? De otro lado, entre las dos campanas que se "suman" en la "**primera campana**", una centrada en 45 m/s y la otra en 60 m/s ¿existe una diferencia en un contraste de hipótesis de medias de normales con medias y varianzas desconocidas, si nos desvelaran las identidades de qué sujetos han tomado placebo y quiénes **trandolapril** en el doble ciego? De ser así, efectuaríamos una **comparación de medias** de dos normales independientes con media y desviación típica desconocidas, donde tomaríamos la no homocedasticidad en la varianza y el estadístico de contraste sería una **t-Student con f grados de libertad**, viniendo **f** dado por la **fórmula de Welch**.

- d) Con el fin de poder efectuar un estudio poblacional sobre obesidad infantil se efectúa un estudio para evaluar la validez del peso y talla declarados por familiares de menores de 4 años, para clasificar su obesidad. Utilizando los valores medidos y declarados en los niños escolarizados en las escuelas infantiles de la ciudad, se clasificaron a los niños y niñas en dos grupos: con sobrepeso y sin sobrepeso. Se está interesado en contrastar si la clasificación de obesidad efectuada con las mediciones y con los valores declarados es diferente. Indica la prueba estadística a efectuar y justifica su elección.

En este caso, tenemos dos variables categóricas y utilizaremos un estudio mediante el **Coeficiente de Kappa de Cohen**. El Coeficiente kappa de Cohen es una medida estadística que ajusta el efecto del azar en la proporción de la concordancia observada para elementos cualitativos (variables categóricas). Es una medida más robusta que el simple cálculo del porcentaje de concordancia, ya que  $k$  tiene en cuenta el acuerdo que ocurre por azar:

**X:** Clasificación realizada a partir de las declaraciones de los familiares (Con sobrepeso, Sin sobrepeso)

**Y:** Clasificación realizada a partir de las mediciones (Con sobrepeso, Sin sobrepeso)

A partir de ambas clasificaciones podemos construir una tabla de contingencia como la siguiente:

		Criterio (mediciones)		Total
		Con sobrepeso	Sin sobrepeso	
Test (Declaración familiares)	Con sobrepeso	$n_{11}$	$n_{12}$	$n_{1*}$
	Sin sobrepeso	$n_{21}$	$n_{22}$	$n_{2*}$
Total		$n_{*1}$	$n_{*2}$	$n$

**NOTA:** son los mismos sujetos medidos por 2 métodos distintos → **No son Datos pareados o relacionados** pues, en el fondo estamos ante **2 observaciones distintas**, sobre una misma persona, pero en el fondo se miden cosas distintas.

Con los datos de la tabla podemos calcular los siguientes estadísticos:

- Proporción de acuerdo observado

$$p_o = \frac{n_{11} + n_{22}}{n}$$

- Proporción de acuerdo esperado por azar:

$$p_e = \frac{n_{*1} \cdot n_{1*} + n_{*2} \cdot n_{2*}}{n^2}$$

- Coeficiente kappa de cohen, es un estadístico que mide el grado “de acuerdo” en las clasificaciones más allá del esperado por el azar:

$$k = \frac{p_o - p_e}{1 - p_e}$$

Cuánto mayor sea el valor del coeficiente Kappa, mayor es el acuerdo por encima del azar entre las clasificaciones realizadas. Los valores de Kappa, pueden oscilar entre:

$$\frac{-p_e}{1 - p_e} \leq k \leq 1$$

Cohen's Kappa statistic ( $\kappa$ )	Strength of agreement
< 0.00	Poor
0.00–0.20	Slight
0.20–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

Asumiendo normalidad en las mediciones realizadas, tanto declaradas como medidas, podemos calcular un intervalo de confianza para dicho coeficiente, y también, realizar contraste de hipótesis: ( $a$  es un valor de capa poblacional)

- Hipótesis:

$$H_0: k = a$$

$$H_1: k \neq a$$

- Estadístico de contraste

$$Z = \frac{k - a}{\sigma_k} \sim N(0,1) \quad \sigma_k = \sqrt{\frac{p_e}{n \cdot (1 - p_e)}}$$

- Intervalo de confianza

$$k \pm Z_{\alpha/2} \cdot \sigma_k$$

*también se puede calcular el p - valor.*

**NOTA:** podría hacerse el mismo análisis teniendo en cuenta el sexo, siendo una variable categórica y reformulando las frecuencias pertinentes de las estadísticas tomadas en las muestras.

## 2.- Ejercicio práctico

Este ejercicio consta de diversas partes en un intento de simular lo que se lleva a cabo en un estudio real.

Se ha simplificado para hacerlo más practicable por lo que no hace falta que os agobiéis si algo no os cuadra del todo con la realidad. De lo que se trata es que veamos cómo aplicar las distintas técnicas que hemos estudiado, de forma integral, en un problema de análisis de datos.

### Los datos.

Se dispone de datos de 701 pacientes trasplantados y se está interesado principalmente en ver si algunas causas del donante se asocian con un fallo precoz del trasplante(tx) medido como muerte antes de los 90 días (status90) . Se dispone de la edad del donante(edad\_d), edad del receptor (edad\_r), el sexo del donantereceptor(sexodr) donde M es masculino y F es femenino, la causa de muerte del donante(causa) donde además de la anoxia, TCE es traumatismo craneo-encefálico y ACV es accidente cerebrovascular. Además se dispone de algunas características del paciente como el uso de vasopresores ( vasopres), los índices Child- Pugh(pugh) y Quick (quick), la bilirrubina (bilirr) y la creatinina (creati).

Los datos los tenéis disponibles en Stata (dades\_tx.dta), en Excel (dades\_tx.xlsx) y en formato csv(dades\_tx.csv). La estructura está en la siguiente tabla

N de variables: 11

N de observaciones: 701

Variable	Type	Format	Description	Label	Valid cases Missing
ident	float	%9.0g	Número identificación		701 ( 100.00%) 0 ( 0.00%)
status90	double	%10.0g	Estatus a los 90 días	0: Vivo 1: Muerto	701 ( 100.00%) 0 ( 0.00%)
pugh	double	%10.0g	Puntuación Child-Pugh Score		633 ( 90.30%) 68 ( 9.70%)
quick	double	%10.0g	Índice de Quick		694 ( 99.00%) 7 ( 1.00%)
bilirr	double	%10.0g	Bilirrubina		694 ( 99.00%) 7 ( 1.00%)
creati	double	%10.0g	Creatinina		691 ( 98.57%) 10 ( 1.43%)
vasopres	double	%10.0g	vasopresores	0: no 1: si	701 ( 100.00%) 0 ( 0.00%)
sexodr	double	%10.0g	Sexo donante-receptor	0: MM 1: MF 2: FM 3: FF	701 ( 100.00%) 0 ( 0.00%)
edad_d	double	%10.0g	Edad Donante		701 ( 100.00%) 0 ( 0.00%)
edad_r	double	%10.0g	Edad Receptor		701 ( 100.00%) 0 ( 0.00%)
causa	double	%10.0g	Causa de la muerte del donante	0: TCE 1: ACV 2: anoxia	691 ( 98.57%) 10 ( 1.43%)

Selecciona 600 casos al azar del fichero para generar la base datos del trabajo personalizada

```
 #(set.seed(#####) Sustituye ##### por el número de vuestro código UOC
 library(foreign)
 dades<-as.data.frame(read.dta("dades_tx.dta"))
 dades600<-dades[sample(1:nrow(dades),600,replace=FALSE),]
```

### Las preguntas.

En las siguientes preguntas además de escoger las hipótesis adecuadas, justifica el uso de la prueba o pruebas utilizadas e interpreta los resultados.

- a) ¿Existe asociación entre el sexo donante-receptor( sexodr) y el estatus a los 90 días (status90)? Justifica la/s prueba/s elegida/s.

```
 #Cargamos los datos del fichero.
 datos <- read.csv("dades_tx.csv", sep = ",", header = TRUE, dec = ".")
 #Visualizamos los primeros datos.
 head(datos)
 #Las clases o tipos de los datos.
 str(datos)
 #Miramos si hay valores nulos en los datos cargados del fichero.
 sapply(datos, function(x) sum(is.na(x)))
```

```
 ident status90  pugh  quick  bilirr  creati vasopres  sexodr  edad_d  edad_r
 causa
```

```
 0 0 68 7 7 10 0 0 0 0 0
```

#Eliminamos del dataset las filas que contienen datos o columnas vacías.

```
 datos <- na.omit(datos)
 #Comprobamos que ahora no hay datos nulos.
 sapply(datos, function(x) sum(is.na(x)))
```

```
 ident status90  pugh  quick  bilirr  creati vasopres  sexodr  edad_d  edad_r
 causa
```

```
 0 0 0 0 0 0 0 0 0 0 0
```

#Vemos que ahora ya no tenemos datos nulos y, al tomar una muestra aleatoria de 600 sujetos, seguro que todos tienen valores en sus campos pertinentes.

#Fijamos la semilla para el GNA y cargamos librerías pertinentes de funciones que utilizaremos.

```
 set.seed(1234)
 library(foreign)
```

#Tomamos una muestra aleatoria sin remplazo y lo cargamos en un dataset.

```
 dades600<-dades[sample(1:nrow(dades),600,replace=FALSE),]
```

#Miramos las dimensiones del dataset y la clase de sus datos.

```
 dim(dades600)
```

```
 600 11
```

```
 str(dades600)
```

#ejercicio 2a

#Hacemos un test Ji-Cuadrado de Pearsons porque son dos variables cualitativas, el sexo. Si hubiesen sido cuantitativas podríamos hacer una regresión, pero aquí sólo podemos comparar las proporciones y ver si son o no dependientes.

```
#Contrastamos H0: son independientes y H1: son dependientes.  
#install.packages('vcd')  
#Prueba ji-cuadrado de independencia al ser cualitativas las dos variables y buscar  
#su posible asociación
```

```
library(vcd)  
(tabla=table(dades600$status90,dades600$sexodr))  
(test.1=chisq.test(tabla))  
test.1$expected
```

```
FF FM MF MM  
Muerto 18 18 16 31  
Vivo 71 135 76 235
```

Pearson's Chi-squared test

```
data: tabla  
X-squared = 5.6362, df = 3, p-value = 0.1307  
FF FM MF MM  
Muerto 12.31167 21.165 12.72667 36.79667  
Vivo 76.68833 131.835 79.27333 229.20333
```

#El p-valor es superior a 0.05 en el Test de Pearson Ji-Cuadrado, lo que indica que no podemos rechazar H0 y decir que son dependientes.

```
#Ahora calculamos la intensidad de esta asociación entre las variables, utilizando  
assocstats:  
assocstats(x = tabla)
```

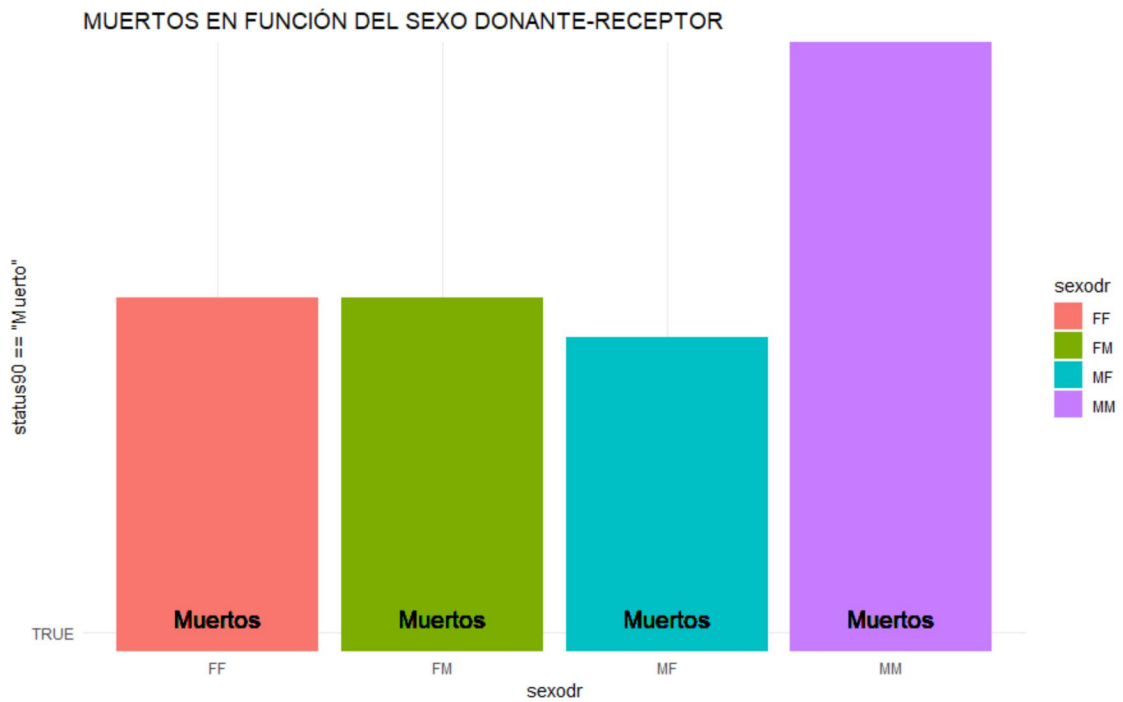
```
X^2 df P(> X^2)  
Likelihood Ratio 5.3295 3 0.14920  
Pearson 5.6362 3 0.13072
```

```
Phi-Coefficient : NA  
Contingency Coeff.: 0.096  
Cramer's V : 0.097
```

```
#Coeficientes de Pearson, la razón de verosimilitud (es un test de hipótesis  
equivalente al ji-Cuadrado)  
#No rechazamos la hipótesis nula, podemos confiar significativamente en que son  
independientes.
```

```
#Ver enlace para interpretar los otros coeficientes que miden la intensidad,  
#Coeficiente de contingencia y V de Cramer, que vemos que son bajos.  
#https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize
```

```
library(ggplot2)  
#Graficamos los Muertos en función del sexo donante-receptor.  
df <- dades600[dades600$status90=='Muerto',]  
ggplot(data=df, aes(x=sexodr, y=status90=='Muerto', fill=sexodr)) +  
geom_bar(stat="identity")+  
geom_text(aes(label='Muertos'), vjust=-0.3, size=5)+  
theme_minimal()+  
ggtitle("MUERTOS EN FUNCIÓN DEL SEXO DONANTE-RECEPTOR")
```



b) ¿ Hay normalidad en las variables de los parametros Puntuación Child-Pug(pugh), índice quick( quick), bilirrubina(bilirr) y creatinina(creati)?

Aplicaremos el test de Shapiro

##ejercicio 2b

#Son 4 variables cuantitativas.

```
shapiro.test(dades600$pugh)
```

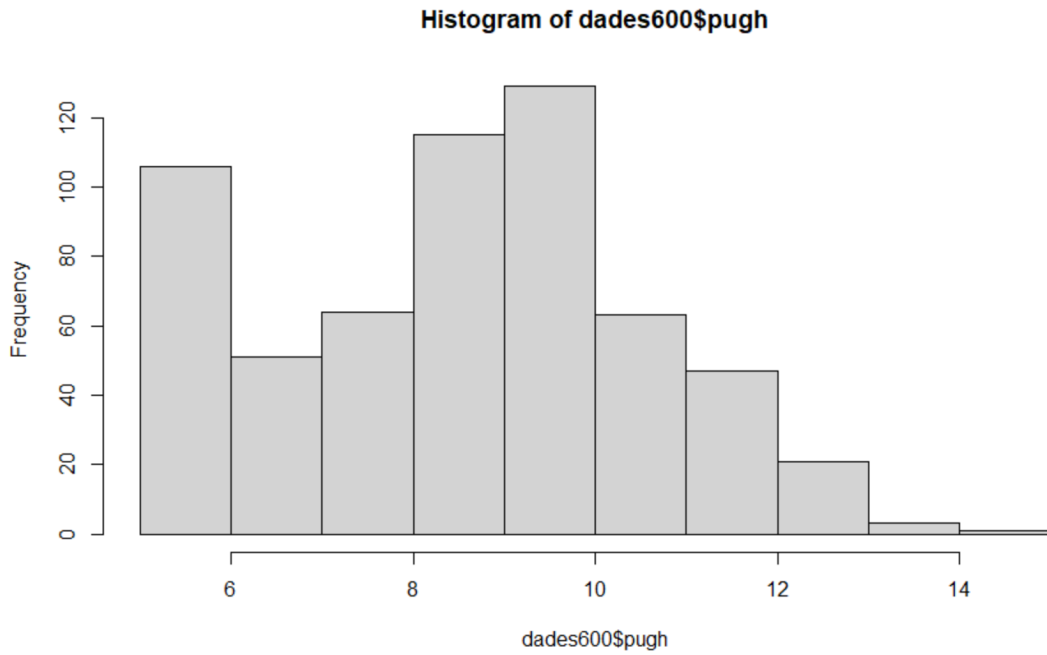
Shapiro-Wilk normality test

*data: dades600\$pugh*

*W = 0.95929, p-value = 7.873e-12*

```
hist(dades600$pugh)
```





*Ilustración 1. Vemos que no sigue una distribución Normal*

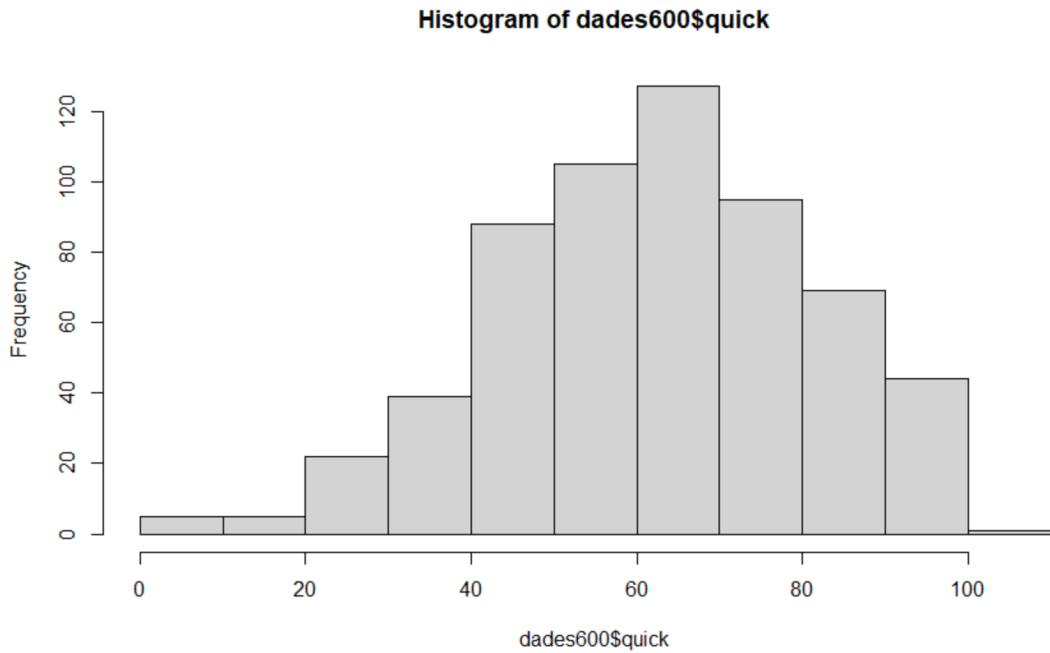
**#No es normal visto el p-valor y su histograma.**

```
shapiro.test(dades600$quick)
```

*Shapiro-Wilk normality test*

*data: dades600\$quick  
W = 0.98906, p-value = 0.0001876*

```
hist(dades600$quick)
```



*Ilustración 2. Vemos que no sigue una distribución Normal*

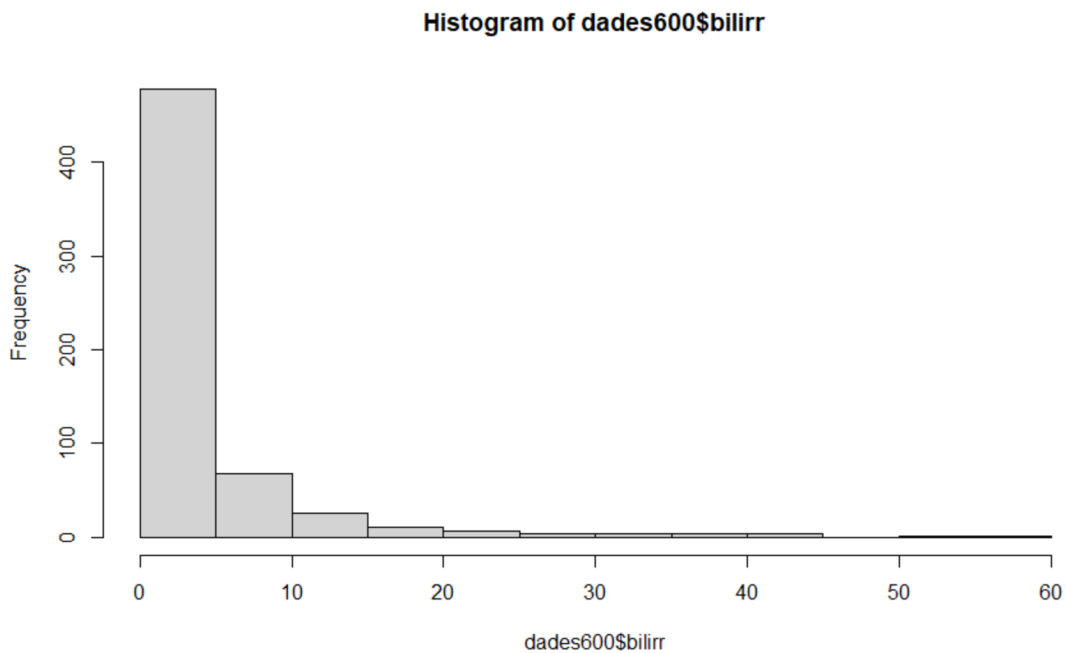
**#No es normal dado el p-valor y su histograma.**

```
shapiro.test(dades600$bilirr)
```

*Shapiro-Wilk normality test*

*data: dades600\$bilirr  
W = 0.53293, p-value < 2.2e-16*

```
hist(dades600$bilirr)
```



*Ilustración 3. Vemos que no sigue una distribución Normal*

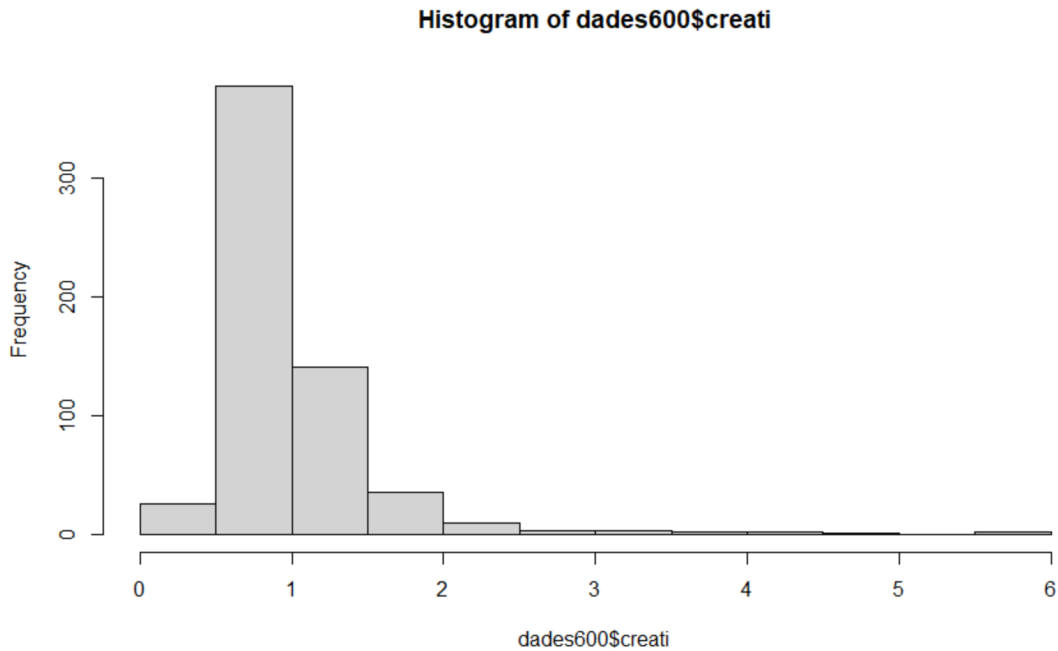
**#No es normal dado el p-valor y el histograma.**

```
shapiro.test(dades600$creati)
```

*Shapiro-Wilk normality test*

*data: dades600\$creati*  
*W = 0.65562, p-value < 2.2e-16*

```
hist(dades600$creati)
```



*Ilustración 4. Vemos que no sigue una distribución Normal*

**#No es normal dado el p-valor y el histograma.**

- c) **¿Existen diferencias entre la edad del receptor y la edad del donante. Justifica la prueba elegida.**

```
#ejercicio 2c
```

```
#Son variables cuantitativas. Es un diseño de medidas relacionadas, podemos o no aplicar una prueba paramétrica o no, dependiendo de si se cumplen las condiciones de normalidad. Nos centraremos en la diferencia entre medidas.
```

```
d=dades600$edad_d-dades600$edad_r
```

```
#Miramos la normalidad mediante un test de Shapiro.
```

```
shapiro.test(d)
```

*Shapiro-Wilk normality test*

*data: d*  
*W = 0.98377, p-value = 3.208e-06*  

```
hist(d)
```

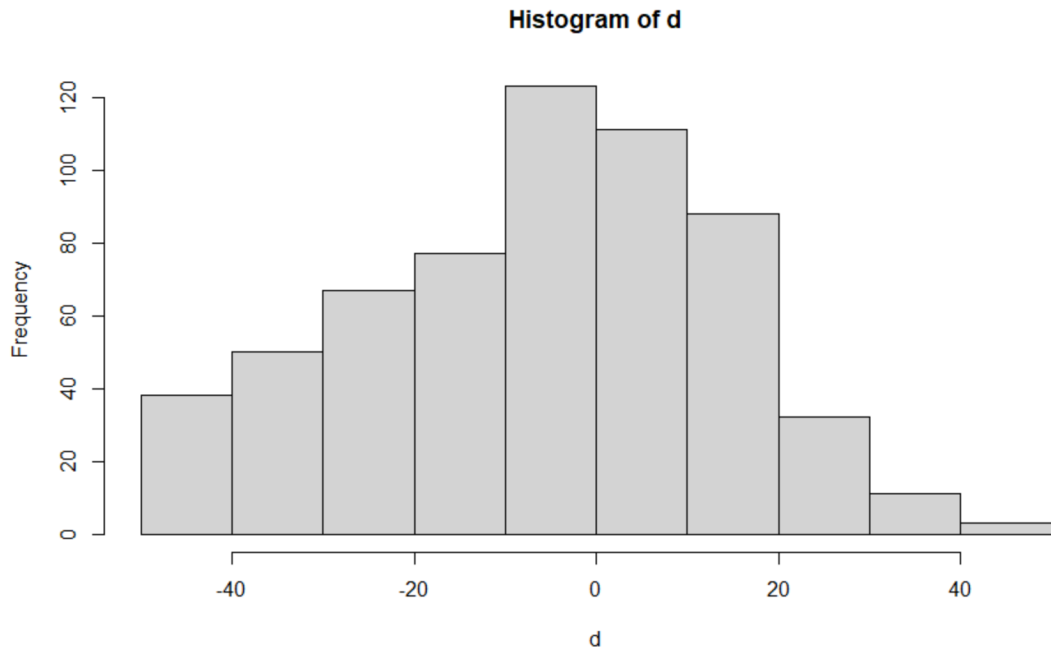


Ilustración 5. Vemos que no sigue una distribución Normal

**#No son normales, aplico un test no paramétrico.**

```
edad_d=dades600$edad_d
edad_r=dades600$edad_r
```

#Puedo aplicar otros: Kruskal Wallis, pero me quedo con wilcoxon porque se aplica para diferencias.

#Ahora un test no paramétrico (test Wilcoxon para muestras relacionadas).  
`wilcox.test(edad_d, edad_r, paired = TRUE )`

*Wilcoxon signed rank test with continuity correction*

*data: edad\_d and edad\_r*  
*V = 60351, p-value = 8.068e-10*  
*alternative hypothesis: true location shift is not equal to 0*

```
median(edad_d)
51
median(edad_r)
56
```

**#Si hay diferencias visto el test, se rechaza la H0**

**#Mirando las medianas y no las medias, que no son representativas, utilizando un estadístico más robusto como la mediana.**

**#En este caso podemos decir que la mediana de edad de los donantes es menor que la mediana de edad de los receptores.**

```
#Bootstrap
install.packages('MKinfer')
library(MKinfer)
```

```
#Hacemos un test bootstrap utilizando el estadístico t-student (por defecto hace
10 mil réplicas o muestreos de los 600)
boot.t.test(edad_d,edad_r,paired = T)
```

*Bootstrapped Paired t-test*

```
data: edad_d and edad_r
bootstrapped p-value < 2.2e-16
95 percent bootstrap percentile confidence interval:
-7.440083 -4.316583
```

*Results without bootstrap:*

```
t = -7.2461, df = 599, p-value = 1.327e-12
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-7.441908 -4.268092
sample estimates:
mean of the differences
-5.855
```

**#Rechazamos la hipótesis nula. Podemos ver el resultado del bootstrap y su intervalo de confianza, y también del intervalo de confianza del test paramétrico de la t-student.**

```
#Hacemos un test de permutación.
MKinfer::perm.t.test(edad_d,edad_r, paired=T)
```

*Permutation Paired t-test*

```
data: edad_d and edad_r
(Monte-Carlo) permutation p-value < 2.2e-16
95 percent (Monte-Carlo) permutation percentile confidence interval:
-7.452346 -4.254698
```

*Results without permutation:*

```
t = -7.2461, df = 599, p-value = 1.327e-12
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-7.441908 -4.268092
sample estimates:
mean of the differences
-5.855
```

**#Salen resultados con y sin permutaciones. la conclusión es que la edad de los donantes es menor que las no donantes.**

d) ¿Es el nivel de la creatinina(creati) diferente según el sexo del donante-receptor(sexodr).

```
#ejercicio 2d
install.packages('RVAideMemoire')
library(RVAideMemoire)
byf.shapiro(creati ~ factor(sexodr),dades600 )
```

*Shapiro-Wilk normality tests*

*data: creati by factor(sexodr)*

```
      W      p-value
FF 0.7557 7.373e-11 ***
FM 0.6757 < 2.2e-16 ***
MF 0.5648 4.394e-15 ***
MM 0.6495 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#Hay 4 niveles o grupos y evaluamos la normalidad en cada uno de ellos, observando con este test que en cada grupo la variable no se distribuye de forma normal.

```
install.packages('car')
library(car)
```

#Para igualdad de varianzas.

```
leveneTest(creati ~ factor(sexodr),dades600)
```

```
Levene's Test for Homogeneity of Variance (center = median)
  Df F value Pr(>F)
group 3  0.8591 0.4621
  596
```

#Si hay homogeneidad de varianzas, pero no normalidad  
#Utilizamos un Test no paramétrico, la prueba de Kruskal Wallis, que compara rangos promedios de los 4 grupos.

```
kruskal.test(creati ~ factor(sexodr),dades600)
```

*Kruskal-Wallis rank sum test*

```
data: creati by factor(sexodr)
Kruskal-Wallis chi-squared = 28.478, df = 3, p-value = 2.883e-06
```

#p-valor bajo, lo que significa que hay diferencias en el nivel de creatinina según el sexo donante/receptor.

#No son distribuciones semejantes, no vienen de la misma población o se distribuyen de distinta manera.

#Las muestras son distintas pero no sabemos cuáles son distintas a cuáles. Hacemos un test post-hoc, el Dunn Test, que es un test no paramétrico, pues nuestros datos no son normales y necesitamos una estadística no paramétrica con datos pareados para hacer múltiples comparaciones entre las muestras que vamos generando.

```
library(PMCMRplus)
#Test no paramétrico para datos pareados con múltiples comparaciones de las sumas
de los rangos. Compara por pares de grupos para ver si hay diferencia.
kwAllPairsDunnTest(dades600$creati, factor(dades600$sexodr), method =
"Chisq",p.adjust.method="bonferroni")
```

*Pairwise comparisons using Dunn's all-pairs test*

*data: dades600\$creati and factor(dades600\$sexodr)*

```
      FF      FM      MF
FM 1e-05 -      -
MF 0.63683 0.01562 -
MM 0.00014 1.00000 0.12943
```

*P value adjustment method: bonferroni*

*alternative hypothesis: two.sided*

*Warning message:*

*In kwAllPairsDunnTest.default(dades600\$creati, factor(dades600\$sexodr), :  
Ties are present. z-quantiles were corrected for ties.*

#p-valores ajustados por el método de Bonferroni. MM-MF,MM-FM,MF-FF no hay diferencia en la creatinina, en el resto sí.

#APLICAR WILCOXON

```
install.packages("WRS2")
```

```
library(WRS2)
```

#Tenemos una variable independiente con 4 grupos de sexo donante/receptor, siendo la creatinina la variable dependiente

#Si se cumplen las condiciones aplicaríamos un ANOVA, pero al no cumplirse podemos aplicar Kruskal, ya visto, o un test de ANOVA robusto con bootstrap

```
t1waybt(creati ~ factor(sexodr),dades600,nboot = 1000) #mil bootstrapping
```

*Call:*

```
t1waybt(formula = creati ~ factor(sexodr), data = dades600, nboot = 1000)
```

*Effective number of bootstrap samples was 1000.*

*Test statistic: 8.9757*

*p-value: 0*

*Variance explained: 0.083*

*Effect size: 0.287*

**#Si hay diferencias entre grupos, cosa que ya decía la anterior prueba de Kruskal Wallis anterior.**

#Ahora miramos las comparaciones por pares entre grupos, una vez sabiendo que hay un efecto significativo de la variable donante/receptor sobre la creatinina

#se hacen comparaciones múltiples a posteriori para ver entre qué grupos se encuentran diferencias.

```
lincon(creati ~ factor(sexodr),dades600)
```

Call:

```
lincon(formula = creati ~ factor(sexodr), data = dades600)
```

```
          psihat ci.lower ci.upper p.value
FF vs. FM -0.18698 -0.28938 -0.08459 0.00002
FF vs. MF -0.07334 -0.20040  0.05372 0.23998
FF vs. MM -0.14628 -0.24183 -0.05073 0.00043
FM vs. MF  0.11364 -0.01031  0.23760 0.06491
FM vs. MM  0.04070 -0.05014  0.13155 0.23998
MF vs. MM -0.07294 -0.19153  0.04565 0.23998
```

#Sale el estadístico, un intervalo y un p-valor. Tomando el criterio del p-valor de 0.05, sólo hay diferencia en la 1ª y en la 3ª, o podemos ver el intervalo de confianza, si no contiene el valor 0 significa que hay diferencias.

**e) ¿Existe relación entre la edad del donante(edad\_d) y el status a los 9 días (status90)?**

**Justifica la/s prueba/s elegida/s.**

#ejercicio 2e

```
library(RVAideMemoire)
```

```
byf.shapiro(edad_d ~ factor(status90),dades600 )
```

*Shapiro-Wilk normality tests*

```
data: edad_d by factor(status90)
```

```
      W      p-value
Muerto 0.9465 0.001742 **
Vivo    0.9559 2.603e-11 ***
```

---

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#No es normal

```
library(car)
```

```
leveneTest(edad_d ~ factor(status90),dades600 )
```

*Levene's Test for Homogeneity of Variance (center = median)*

```
      Df F value Pr(>F)
group 1  2.1688 0.1414
      598
```

**#Las varianzas si son homogéneas.**

#Hacemos un test no paramétrico, al ser muestras independientes que no cumplen normalidad, y comparamos los rangos promedios mediante el test de Wilconxon.

```
wilcox.test(edad_d ~ factor(status90),dades600 )
```

*Wilcoxon rank sum test with continuity correction*



```
data: edad_d by factor(status90)
W = 23337, p-value = 0.1995
alternative hypothesis: true location shift is not equal to 0
```

#Dependiendo del nivel de significación aceptaríamos la hipótesis nula bajo este test y no habría diferencia en la edad.

```
boot.t.test(edad_d ~ factor(status90),dades600,var.equal=T)
```

*Bootstrapped Two Sample t-test*

```
data: edad_d by factor(status90)
bootstrapped p-value = 0.2154
95 percent bootstrap percentile confidence interval:
-1.494463 6.932393
```

*Results without bootstrap:*

```
t = 1.2405, df = 598, p-value = 0.2153
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-1.564023 6.927404
sample estimates:
mean in group Muerto  mean in group Vivo
      50.18072         47.49903
```

#Con el bootstrap si miramos el p-valor hay diferencias significativas (0.049) pero observando el intervalo de confianza calculado, no, porque incluye al 0.

```
MKinfer::perm.t.test(edad_d ~ factor(status90),dades600,var.equal=T)
```

*Permutation Two Sample t-test*

```
data: edad_d by factor(status90)
(Monte-Carlo) permutation p-value = 0.2123
95 percent (Monte-Carlo) permutation percentile confidence interval:
-1.544639 6.775605
```

*Results without permutation:*

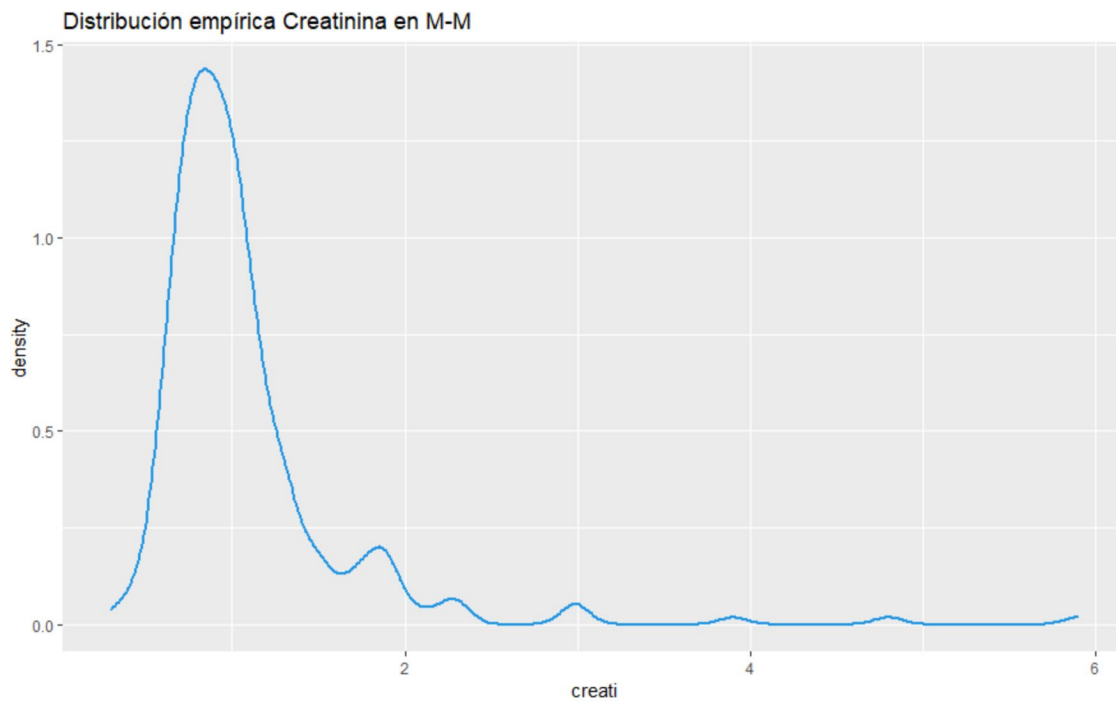
```
t = 1.2405, df = 598, p-value = 0.2153
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-1.564023 6.927404
sample estimates:
mean in group Muerto  mean in group Vivo
      50.18072         47.49903
```

#No ocurre la paradoja, el p-valor es >0.05 y se acepta H0 y su intervalo incluye el 0 (que no hay diferencia entre los 2 grupos).

#Gráficos de densidad de distribución empírica de la creatinina según el sexo donante-receptor para comprobar los resultados numéricos obtenidos por los métodos.

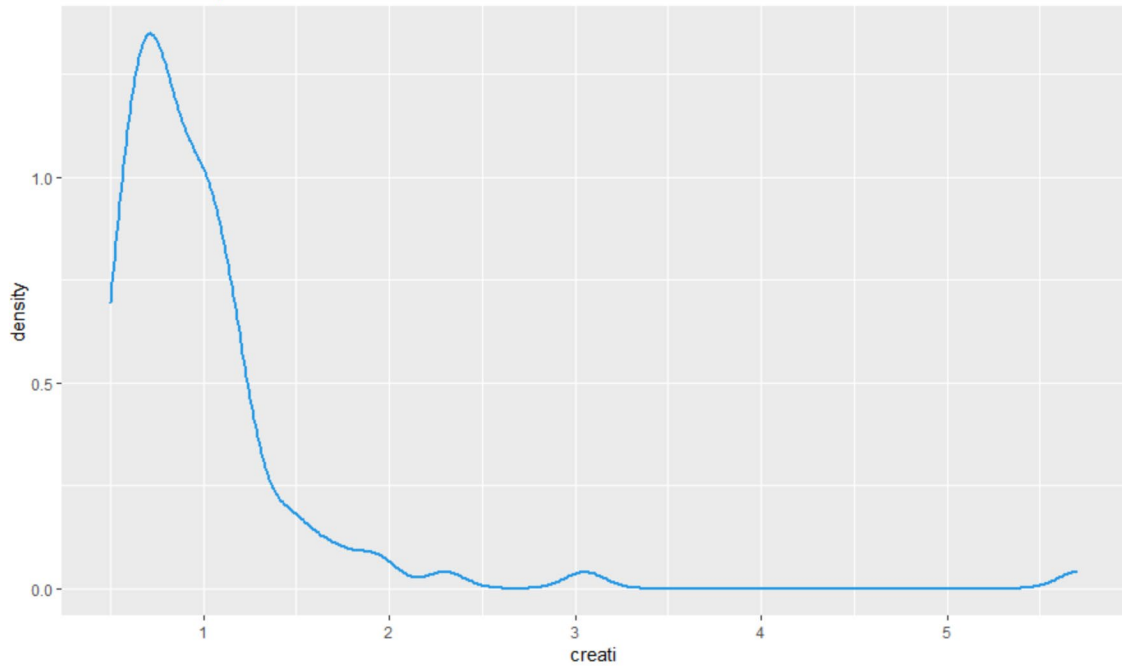
```
df1 <- dades600[dades600$sexodr=='MM',]
df2 <- dades600[dades600$sexodr=='MF',]
df3 <- dades600[dades600$sexodr=='FM',]
df4 <- dades600[dades600$sexodr=='FF',]
```

```
ggplot(df1, aes(x = creati)) +  
  geom_density(color = 4, # Color  
              lwd = 1,   # Ancho  
              linetype = 1) + # Tipo  
  ggtitle("Distribución empírica Creatinina en M-M")
```



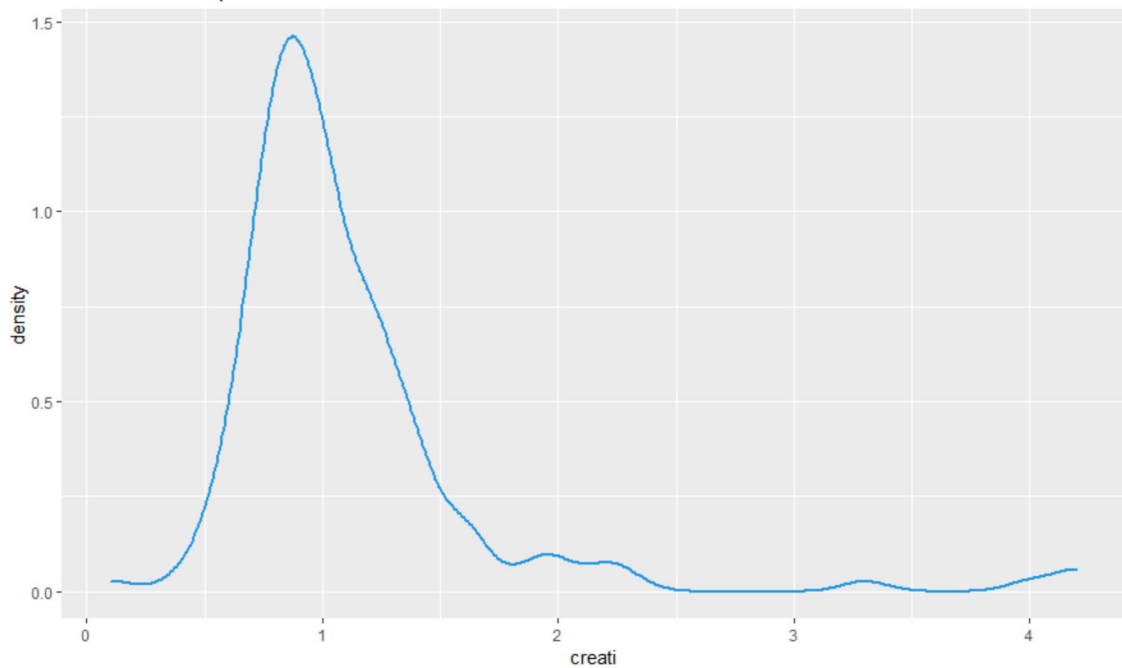
```
ggplot(df2, aes(x = creati)) +  
  geom_density(color = 4, # Color  
              lwd = 1,   # Ancho  
              linetype = 1) + # Tipo  
  ggtitle("Distribución empírica Creatinina en M-F")
```

Distribución empírica Creatinina en M-F

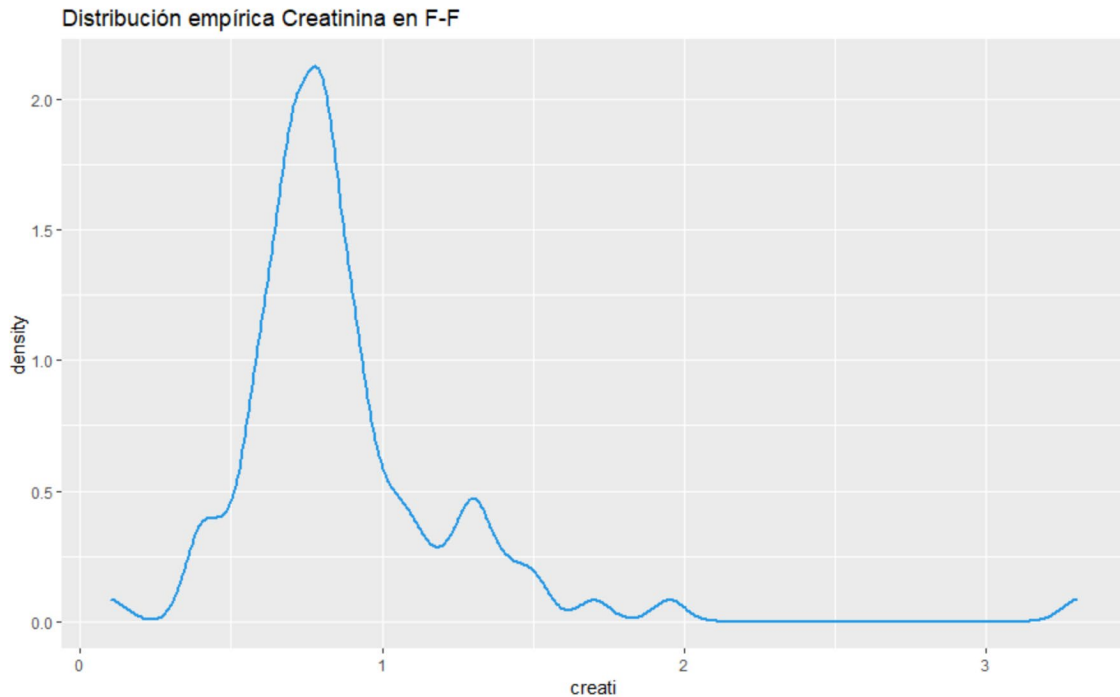


```
ggplot(df3, aes(x = creati)) +  
  geom_density(color = 4, # Color  
              lwd = 1,   # Ancho  
              linetype = 1) + # Tipo  
  ggtitle("Distribución empírica Creatinina en F-M")
```

Distribución empírica Creatinina en F-M



```
ggplot(df4, aes(x = creati)) +  
  geom_density(color = 4, # Color  
              lwd = 1,   # Ancho  
              linetype = 1) + # Tipo  
  ggtitle("Distribución empírica Creatinina en F-F")
```



### 3.- Algunas cosas más:

a) Indica porqué hay que utilizar técnicas de comparación múltiple como la corrección de Bonferroni.

Cuando se realizan **comparaciones múltiples**, el nivel de **significación global se incrementa** a medida que **aumenta el número de comparaciones** y, por tanto, hay más probabilidad de **cometer al menos un error tipo I en las comparaciones**. Esto **se puede corregir** aplicando algún **procedimiento de ajuste del p-valor** de cada una de las comparaciones, como por ejemplo, el de **Bonferroni**. El ajuste de **Bonferroni** consiste en multiplicar cada **p-valor** por el **número de comparaciones a realizar**.

b) Indica para que te podrían servir las técnicas de bootstrap en los análisis estadísticos que has realizado. Calcula el intervalo utilizando Bootstrap de la media del valor de la creatinina. Compáralo con el intervalo de confianza asintótico habitual.

Las técnicas de **Bootstrap** se pueden utilizar cuando **no se cumplan los supuestos asociados a las pruebas paramétricas** (Normalidad y homogeneidad de varianzas). Estas técnicas nos permiten **realizar las comparaciones aunque no se cumplan los supuestos**. Se hace un **remuestreo del muestreo para simular lo acontecido a nivel poblacional**.

```
#ejercicio 3b
boot.t.test(dades600$creati)
```

*Bootstrapped One Sample t-test*

```
data: dades600$creati
bootstrapped p-value < 2.2e-16
```

95 percent bootstrap percentile confidence interval:  
0.9986833 1.0917342

Results without bootstrap:

$t = 44.406$ ,  $df = 599$ ,  $p\text{-value} < 2.2e-16$

alternative hypothesis: true mean is not equal to 0

95 percent confidence interval:

0.9966799 1.0889201

sample estimates:

mean of  $x$

1.0428

**#Tenemos el intervalo bootstrapping y el asintótico (el calculado a partir de la distribución normal, clásico). Vemos los 2 intervalos, ninguno contiene al cero, son muy similares y los p-valores son bajos.**