

Práctica 7. Test de comparación de dos medias en R.

Manuela Expósito Ruiz, Pedro J. Femia Marzo, Christian J. Acal González y Miguel Ángel Montero Alonso



**UNIVERSIDAD
DE GRANADA**

Todo el material para el conjunto de actividades de este curso ha sido elaborado y es propiedad intelectual del grupo **BioestadísticaR** formado por:

Antonio Martín Andrés
Juan de Dios Luna del Castillo,
Pedro Femia Marzo,
Miguel Ángel Montero Alonso,
Christian José Acal González,
Pedro María Carmona Sáez,
Juan Manuel Melchor Rodríguez,
José Luis Romero Béjar,
Manuela Expósito Ruíz,
Juan Antonio Villatoro García.

Todos los integrantes del grupo han participado en todas las actividades, en su elección, construcción, correcciones o en su edición final, no obstante, en cada una de ellas, aparecerán uno o más nombres correspondientes a las personas que han tenido la máxima responsabilidad de su elaboración junto al grupo de **BioestadísticaR**.

Todos los materiales están protegidos por la Licencia Creative Commons **CC BY-NC-ND** que permite "descargar las obras y compartirlas con otras personas, siempre que se reconozca su autoría, pero no se pueden cambiar de ninguna manera ni se pueden utilizar comercialmente".

Práctica 7. Test de comparación de dos medias

Manuela Expósito Ruiz, Pedro J Femia Marzo, Christian J. Acal González
y Miguel Ángel Montero Alonso

7.1 Test para dos muestras independientes

En esta práctica, el objetivo que se plantea es comparar las medias de dos poblaciones, para lo que existen diferentes test. La elección del test apropiado va a depender de la naturaleza de los datos, entre otros, si las muestras de las que proceden son independientes o relacionadas, y del supuesto de normalidad de la variable de estudio.

En cualquier caso, el contraste que se plantea para comparar las medias de dos poblaciones es:

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

lo que equivalentemente se puede formular como

$$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$$

siendo $\mu_1 - \mu_2$ el tamaño bruto del efecto.

Para decidir el tipo de test a aplicar en la comparación de dos muestras independientes, en primer lugar hay que comprobar si se cumplen ciertas hipótesis sobre los datos, como la normalidad y la homogeneidad de varianzas.

7.1.1 Hipótesis de normalidad

Una de las condiciones que deben comprobarse sobre los datos, es si estos proceden de una distribución normal, para lo cual debe realizarse el correspondiente contraste de hipótesis:

$$\begin{cases} H_0 : \text{la distribución de la variable es normal} \\ H_1 : \text{la distribución de la variable es distinta a la normal} \end{cases}$$

Nótese que en el caso de rechazar la hipótesis nula, solo se podrá afirmar que la distribución de la que procede la muestra es distinta a la normal, sin permitirnos este contraste afirmar cuál es la naturaleza de la misma. Por otra parte, la normalidad debe comprobarse en cada uno de los grupos en los que posteriormente se realizará el contraste de igualdad e medias.

El test de Shapiro-Wilks permite realizar el contraste de normalidad, y en **R** puede aplicarse con la función **shapiro.test**. Supongamos que se quiere comprobar si existen diferencias en los niveles de hemoglobina glicosilada (**hba1c**) según sexo, a partir de los datos del fichero **osteo**. Para decidir qué test de comparación de medias aplicar, debe comprobarse previamente si la variable **hba1c** se distribuye normalmente. Para ello escribimos en **R**:

```
shapiro.test(hba1c)
```

```
##
## Shapiro-Wilk normality test
##
## data:  hba1c
## W = 0.9833, p-value = 0.2757
```

Esta orden comprueba si la variable `hba1c` en el conjunto de los datos, sigue una distribución normal. Pero como se ha comentado anteriormente, esta hipótesis debe cumplirse en cada uno de los grupos que se quieran comparar, en este caso en hombres y mujeres. En este caso, la función adecuada es:

```
tapply(hba1c,sexo,shapiro.test)
```

```
## $'1'
##
## Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.95798, p-value = 0.1022
##
##
## $'2'
##
## Shapiro-Wilk normality test
##
## data:  X[[i]]
## W = 0.98002, p-value = 0.5661
```

La función **tapply** permite aplicar la función **shapiro.test** a la variable **hba1c** en cada uno de los grupos establecidos por la variable **sexo**. Algunos autores recomiendan rechazar la hipótesis nula con un nivel de significación $\alpha = 0.10$ para este test. En este caso vemos que tanto para los hombres (`sexo=1`) como para las mujeres (`sexo=2`), el p-valor obtenido es superior a 0.10 en ambos grupos, por lo que no podemos rechazar H_0 y aceptamos que la variable *hba1c* sigue una distribución normal en ambos grupos.

7.1.2 Hipótesis de homogeneidad de varianzas

Otra de las hipótesis que deben comprobarse antes de decidir qué test aplicar, es la homocedasticidad o igualdad de varianzas, es decir, las 2 muestras de las que proceden los datos a comparar, deben tener una variabilidad similar, lo que se formula como sigue:

$$\begin{cases} H_0 : \sigma_1 = \sigma_2 \\ H_1 : \sigma_1 \neq \sigma_2 \end{cases}$$

En **R**, se puede resolver el contraste de igualdad de varianzas con el test de Levene (**leveneTest**), Bartlett (**bartlett.test**) o el contraste de razón de varianzas, F test (**var.test**), entre otros. En esta práctica emplearemos este último, que es el más potente. Siguiendo con el ejemplo anterior, se quiere comprobar si la varianza de la variable *hba1c* es similar en hombres y mujeres, con el test F, escribimos:

```
var.test(hba1c ~ sexo)
```

```
##
## F test to compare two variances
##
## data:  hba1c by sexo
## F = 0.87633, num df = 44, denom df = 48, p-value = 0.66
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.4896791 1.5801515
## sample estimates:
## ratio of variances
##          0.8763332
```

En este caso el p-valor obtenido es 0.8763, por lo que no podemos rechazar H_0 , y se asume por tanto que las varianzas son iguales para ambos grupos. Cuando aplicamos este test, por acuerdo consideraremos un nivel de significación $\alpha = 0.20$.

7.1.3 Comparación de dos medias independientes. Test t de Student y Welch

Una vez comprobadas las hipótesis sobre los datos (normalidad y homocedasticidad), si estas se cumplen, se puede aplicar el test t de Student para muestras independientes. Para el ejemplo que nos ocupa, la hipótesis planteada para los niveles de *hba1c*, sería:

$$\begin{cases} H_0 : \mu_{hombres} = \mu_{mujeres} \\ H_1 : \mu_{hombres} \neq \mu_{mujeres} \end{cases}$$

Como en los apartados anteriores se ha verificado que se cumplen las hipótesis de normalidad y homocedasticidad, puede aplicarse el test t de Student, cuyo estadístico de contraste es:

$$t_{exp} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s^2 \frac{n_1 + n_2}{n_1 n_2}}}$$

siendo s^2 el promedio ponderado de las s_i^2 muestrales, ya que estamos suponiendo que las varianzas son desconocidas, y por tanto habrá que estimarlas en la muestra. Por lo que s^2 se calcula mediante la fórmula:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

El estadístico t_{exp} se compara con el valor teórico de la distribución t de Student con $n_1 + n_2 - 2$ grados de libertad, de modo que si se cumple $t_{exp} > t_{\alpha, n_1 + n_2 - 2}$, entonces tendremos suficiente evidencia muestral para rechazar H_0 con un nivel de significación α , y por lo tanto concluir que las medias son distintas en ambos grupos. Veamos como podemos aplicar este test en **R** con la función **t.test**, siguiendo el ejemplo de la variable *hba1c*

```
t.test(hba1c~sexo, var.equal=TRUE)
```

```
##
## Two Sample t-test
##
## data:  hba1c by sexo
```

```
## t = -0.059393, df = 92, p-value = 0.9528
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
## -0.7637643  0.7194106
## sample estimates:
## mean in group 1 mean in group 2
##      8.553333      8.575510
```

Nótese que uno de los parámetros de la función es **var.equal**, al que en este caso asignamos el valor **TRUE**, puesto que previamente se ha comprobado que se cumple la igualdad de varianzas. En caso contrario, habría que indicar **var.equal=FALSE**. El resultado de la función nos muestra que el estadístico experimental es $t_{exp} = -0.05939$, y el p-valor asociado $p=0.9528$, por lo que en este caso, no puede rechazarse H_0 , y asumimos que las medias son iguales en hombres y mujeres. Además, la salida del programa nos muestra la media de la variable *hba1c* en cada uno de los grupos, siendo 8.5533 para los hombres ($\text{sexo}=1$) y 8.5755 para las mujeres ($\text{sexo}=2$). También calcula el intervalo de confianza para la diferencia de medias, en este caso $[-0.7638, 0.7194]$, que al contener el valor 0, nos confirma que la diferencia de medias no es significativamente distinta de 0, y por tanto no puede rechazarse H_0 .

El test t de Student para muestras independientes, se aplica cuando se cumple tanto la hipótesis de normalidad como la de homocedasticidad. En caso de cumplirse la normalidad, pero no la igualdad de varianzas, lo adecuado es aplicar el test de Welch, cuyo estadístico de contraste es

$$t_{exp} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{A + B}}$$

Estando A y B definidas de la siguiente forma:

$$A = \frac{s_1^2}{n_1} \text{ y } B = \frac{s_2^2}{n_2}$$

El resultado del estadístico t_{exp} se compara con el valor $t_\alpha(f)$ de la distribución t de Student, siendo f

$$f = \frac{(A + B)^2}{\frac{A^2}{n_1 - 1} + \frac{B^2}{n_2 - 1}}$$

Seguendo el ejemplo anterior, basta con modificar la función **t.test**, indicando **var.equal=FALSE**, para que **R** aplique el test de Welch a los datos

```
t.test(hba1c~sexo, var.equal=FALSE)
```

```
##
## Welch Two Sample t-test
##
## data: hba1c by sexo
## t = -0.059562, df = 91.963, p-value = 0.9526
## alternative hypothesis: true difference in means between group 1 and group 2 is not equal to 0
## 95 percent confidence interval:
## -0.7616689  0.7173151
## sample estimates:
## mean in group 1 mean in group 2
##      8.553333      8.575510
```

en este caso se obtiene un p-valor de 0.9526, similar al del contraste anterior, por lo que en este caso tampoco puede rechazarse H_0 .

7.1.3 Comparación de dos medias independientes con el paquete BioestadísticaR

El paquete **BioestadísticaR** contiene una función que nos permite calcular todo lo visto anteriormente de una sola vez. La función `test.t`, comprueba tanto la hipótesis de normalidad como la de igualdad de varianzas, además de calcular simultáneamente los estadísticos de contraste por el test de la T de Student y el de Welch, de modo que podremos realizar el contraste de igualdad de medias de forma más rápida y cómoda. Por ejemplo, si se quiere comprobar si la distribución del peso de los pacientes del fichero `osteo` es igual entre hombres y mujeres, puede resolverse ejecutando la siguiente orden, donde `m` representa la variable numérica a comparar (peso), y `grupos` la variable de agrupación (sexo):

```
test.t(m=peso, grupos=sexo)

##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
## Tamaño muestral: n = 45
## Media: m = 66.9311
## Desviación típica: s = 10.8468
## Error estándar de la media: sem = 1.6169
##
## Estimación:
## 95 %-IC( $\mu$ ): ( 63.6724 , 70.1898 )
## Precisión obtenida: 3.2587
##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
## Tamaño muestral: n = 49
## Media: m = 61
## Desviación típica: s = 12.0388
## Error estándar de la media: sem = 1.7198
##
## Estimación:
## 95 %-IC( $\mu$ ): ( 57.5421 , 64.4579 )
## Precisión obtenida: 3.4579
##
## # t-test para 2 Muestras Independientes
## #-----
## # Información muestral
## Niveles de agrupacion: 1 , 2
##
## [1] Para grupo = 1
## n = 45
## m = 66.9311
## s = 10.8468
## sem= 1.6169
## 95 %-IC(m)= ( 63.6724 , 70.1898 )
##
## [2] Para grupo = 2
## n = 49
## m = 61
```

```
##      s = 12.0388
##      sem= 1.7198
##      95 %-IC(m)= ( 57.5421 , 64.4579 )
##
## # Pruebas de normalidad (test de Shapiro-Wilk)
## [1] Para grupo = 1 , W =0.928, g.l. =45, p =0.008
## [2] Para grupo = 2 , W =0.924, g.l. =45, p =0.0037
##
## # Test de homogeneidad de varianzas. Fexp = (var[2]/var[1])
##      Fexp =1.2319, gl1 = 48, gl2 = 44, p = 0.486
##
## # Diferencia de medias (grupo[1] - grupo[2])
##      Diferencia a contrastar: m0 = 0
##
## a) Test de Student (varianzas homogéneas)
##      texp = 2.5014 , g.l. = 92 , p = 0.0141
##      95%-IC(m1-m2) = (1.2218, 10.6404)
##
## b) Test de Welch (varianzas no homogéneas)
##      texp = 2.5126 , g.l. = 91.97 , p = 0.0137
##      95%-IC(mu1-mu2) = (1.2428, 10.6194)
```

La salida que nos ofrece el programa nos muestra en primer lugar los estadísticos muestrales media (m), desviación típica (s) error estandar (sem) el numero de sujetos (n) y el intervalo de confianza al 95% para la media muestral, de cada uno de los grupos. A contiunción realiza el test de normalidad de Shapiro-Wilk, tanto para hombres como para mujeres. En este caso ambos p-valores son inferiores a 0.10, por lo que no se cumple la hipótesis de normalidad, aunque como es un ejemplo a modo ilustrativo, continuamos comentando las salidas. El test de homocedasticidad implementado en el paquete, es el de razón de varianzas, que en este caso nos lleva a aceptar la hipotesis nula. Por último el programa nos muestra la salida de los test de comparación de medias, tanto de la t de Student como el de Welch. Lo único que el usuario debe hacer es seleccionar el test adecuado, en función del resultado obtenido en el test de razón de varianzas. En este caso lo adecuado sería el test t de Student, por lo que interpretaríamos el valor $p=0.0141$, que nos lleva a rechazar la H_0 , concluyendo que el peso medio es distinto según el sexo.

Esta función también permite realizar el test de comparación de medias a partir de datos resumidos, si no disponemos del vector de datos, como era el caso anterior. Para ello, debe indicarse al programa cuáles son los valores de tamaño muestral, media y desviación típica para cada uno de los grupos, mediante los parámetros **n1**, **m1**, **s1** para la primera muestra, **n2**, **m2** y **s2** para la segunda. También se puede modificar el nivel de confianza (por defecto 95%) con el parámetro **alfa**, así como la potencia deseada para detectar un efecto δ (parámetros **potencia** y **delta** respectivamente) Veamos un ejemplo en el que se indica la información resumida para cada grupo, y se desea declarar significativa una diferencia de $\delta = 5$ unidades (efecto bruto) con una potencia de, al menos, el 85%.

```
test.t(n1=50, m1=63, s1=10, n2=67, m2=58, s2=20, delta=5, potencia=.85)
```

```
##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
##      Tamaño muestral: n = 50
##      Media: m = 63
##      Desviación típica: s = 10
##      Error estándar de la media: sem = 1.4142
##
```

```
## Estimación:
## 95 %-IC( $\mu$ ): ( 60.158 , 65.842 )
## Precisión obtenida: 2.842
##
##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
## Tamaño muestral: n = 67
## Media: m = 58
## Desviación típica: s = 20
## Error estándar de la media: sem = 2.4434
##
## Estimación:
## 95 %-IC( $\mu$ ): ( 53.1216 , 62.8784 )
## Precisión obtenida: 4.8784
##
##
## # t-test para 2 Muestras Independientes
## #-----
## # Información muestral
## Niveles de agrupacion: 1 , 2
##
## [1] Para grupo = 1
## n = 50
## m = 63
## s = 10
## sem= 1.4142
## 95 %-IC(m)= ( 60.158 , 65.842 )
##
## [2] Para grupo = 2
## n = 67
## m = 58
## s = 20
## sem= 2.4434
## 95 %-IC(m)= ( 53.1216 , 62.8784 )
##
## # Test de homogeneidad de varianzas. Fexp = (var[2]/var[1])
## Fexp =4, gl1 = 66, gl2 = 49, p < 1e-04
##
## # Diferencia de medias (grupo[1] - grupo[2])
## Diferencia a contrastar: m0 = 0
##
## a) Test de Student (varianzas homogéneas)
## texp = 1.6217 , g.l. = 115 , p = 0.1076
## 95%-IC(m1-m2) = (-1.1071, 11.1071)
##
## b) Test de Welch (varianzas no homogéneas)
## texp = 1.7711 , g.l. = 102.18 , p = 0.0795
## 95%-IC( $\mu_1-\mu_2$ ) = (-0.5996, 10.5996)
##
##
## # Estudio de la potencia: delta= 5 -> [-5, 5], potencia = 85%
## 70%-IC( $\mu_1-\mu_2$ ) = (2.0591, 7.9409)
```



```
##      ---[---|(-)-]--      potencia < 85%
##
##      Leyenda:  --(---)--      --[---|---]--
##                IC- IC+      m0-d      m0      m0+d
##
## # Estimación del tamaño muestral para detectar una diferencia delta=5 con potencia=85%
## (1) Considerando las varianzas homogéneas:
##      (n1 = n2) >= 199 casos en cada grupo
##
## (2) Considerando las varianzas heterogéneas: r= s[2]/s[1] = 2 , (g.l.'= 111.09 )
##      n1 >= 110 casos en el grupo [1]
##      n2 >= 220 casos en el grupo [2]
```

Las salidas ahora incluyen el tamaño muestral necesario para detectar un efecto de 5 unidades con una potencia mínima del 85%, que sería en este caso de n=110 hombres y n=220 mujeres, ya que las varianzas no son homogéneas.

7.2 Test para dos muestras relacionadas

En determinados estudios, se dispone de datos donde cada individuo proporciona dos observaciones (antes/después, brazo derecho/izquierdo...); se habla entonces de muestras *relacionadas* o *apareadas*. Existe una versión del test t de Student para aplicar en estos casos, que puede realizarse con la función `test.t` del paquete **BioestadísticaR**, indicando los parámetros `m1` y `m2`, que indican cuales son las variables que se quieren comparar.

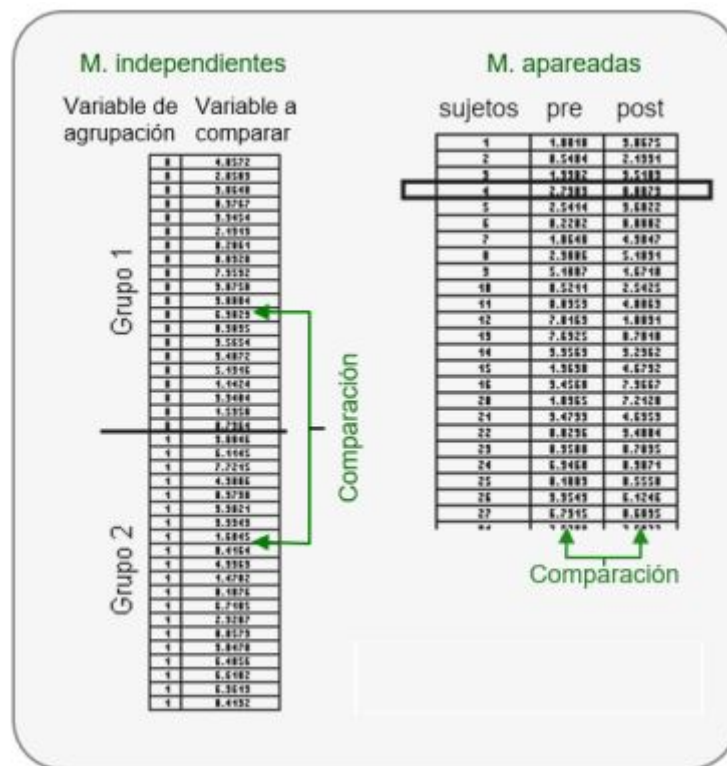


Figure 1: Formato computacional de los datos de un *data.frame* para realizar contrastes de comparación de medias.

Una diferencia que se encuentra al realizar los contrastes para muestras relacionadas, es la forma en que deben estar organizados los datos en el *data.frame*, que debe ser en dos variables distintas, indicando cada una de ellas, el grupo de datos a comparar (que generalizando, señalaremos como *pre* y *post*), como puede verse a modo ilustrativo en la figura 7.1

En el caso de muestras independientes, los datos de la variable a analizar quedan divididos por la variable factor (en el ejemplo visto anteriormente, el sexo), que actúa como variable de agrupación (parte izquierda de la figura 7.1). En el caso de *muestras relacionadas* o *apareadas*, la variable a analizar se organiza en dos columnas distintas del *data.frame* (parte derecha de la figura 7.1).

Veamos un ejemplo en el que aplicamos el test para muestras relacionadas en **R**. Se quiere comparar la densidad de masa ósea medida en el *triángulo* (variable **sztri**) y en el *cueillo* (variable **szcue**) de los pacientes del fichero **oste**. Al tratarse de la misma variable (densidad de masa ósea) medida en dos partes distintas del mismo individuo (triángulo y cuello), nos encontramos ante un caso de muestras *apareadas* o *relacionadas*. Como puede comprobarse en el fichero de datos, las dos variables se recogen en dos columnas distintas, por lo que podemos aplicar el test en **R** como sigue:

```
test.t(m1=szcue, m2=sztri)

##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
##   Tamaño muestral: n = 94
##   Media: m = -1.0502
##   Desviación típica: s = 1.2444
##   Error estándar de la media: sem = 0.1284
##
## Estimación:
##   95 %-IC( $\mu$ ): ( -1.3051 , -0.7953 )
##   Precisión obtenida: 0.2549
##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
##   Tamaño muestral: n = 94
##   Media: m = -0.8956
##   Desviación típica: s = 1.2111
##   Error estándar de la media: sem = 0.1249
##
## Estimación:
##   95 %-IC( $\mu$ ): ( -1.1437 , -0.6476 )
##   Precisión obtenida: 0.2481
##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
##   Tamaño muestral: n = 94
##   Media: m = -0.1546
##   Desviación típica: s = 1.1197
##   Error estándar de la media: sem = 0.1155
##
## Estimación:
```

```

## 95 %-IC( $\mu$ ): ( -0.3839 , 0.0748 )
## Precisión obtenida: 0.2293
##
##
## # t-test para dos muestras relacionadas
## # -----
##
## # Información muestral
## [1] Variable pre:
## n = 94
## m = -1.0502
## s = 1.2444
## sem= 0.1284
## 95 %-IC( $\mu$ )= ( -1.3051 , -0.7953 )
##
## [2] Variable post:
## n = 94
## m = -0.8956
## s = 1.2111
## sem= 0.1249
## 95 %-IC( $\mu$ )= ( -1.1437 , -0.6476 )
##
## # Correlación pre-post
## r = 0.5845
##
## # Normalidad de la diferencia (Test de Shapiro-Wilk)
## W =0.9812, g.l. =94, p =0.1945
##
## # Diferencia pre-post
## n = 94
## m = -0.1546
## s = 1.1197
## sem= 0.1155
## 95 %-IC(m)= ( -0.3839 , 0.0748 )
##
## # Valor de la diferencia media a contrastar: m0 = 0
## texp = 1.3385 , g.l. = 93 , p = 0.184
## 95%-IC(m1-m2) = (-0.3839, 0.0748)

```

La salida que nos muestra el programa incluye información muestral tanto del pre como del post, así como la diferencia media entre ambas variables, en este caso -0.1546 , y su intervalo de confianza al 95%. El p-valor asociado al contraste es $p=0.184$, que nos indica que no puede rechazarse la hipótesis nula, lo que también puede comprobarse con el intervalo de confianza $(-0.3839, 0.0748)$ que contiene al valor 0, indicando que la diferencia de medias no es significativamente distinta de 0, y por lo tanto las medias de la densidad de masa ósea medida en triángulo y cuello pueden considerarse iguales.

El test t de Student para muestras relacionadas se puede realizar también con la información resumida, en el caso en el que no se disponga de los datos en un *data.frame*. Los parámetros que hay que indicar a la función en este caso, son el tamaño muestral (**n**), la media de la diferencia entre las variables pre y post (**m**) y su desviación típica (**s**). La situación entonces, es la de un test con una sola muestra. Por ejemplo, conocidos los valores descriptivos de la variable diferencia pre-post del análisis anterior, podría escribirse de la siguiente forma:

```
test.t(m=-0.1546, s=1.1197, n=94)
```

```
##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
##   Tamaño muestral: n = 94
##   Media: m = -0.1546
##   Desviación típica: s = 1.1197
##   Error estándar de la media: sem = 0.1155
##
## Estimación:
##   95 %-IC( $\mu$ ): ( -0.3839 , 0.0747 )
##   Precisión obtenida: 0.2293
##
## # t-Test con una muestra
## # -----
## # Información muestral
##   n = 94
##   m = -0.1546
##   s = 1.1197
##   sem= 0.1155
##   95 %-IC(m)= ( -0.3839 , 0.0747 )
##
## # Valor de la media a contrastar: m0 = 0
##   texp = 1.3387 , g.l. = 93 , p = 0.1839
##   95%-IC( $\mu$ ) = (-0.3839, 0.0747)
```

De la misma forma que se ha visto en el contraste para muestras independientes, se pueden añadir a la función `test.t` los parámetros **delta** y **potencia**, para realizar el estudio de la fiabilidad de la aceptación de la hipótesis nula del contraste, así como obtener una estimación del tamaño de muestra necesario para declarar significativo un efecto de tamaño mayor o igual a **delta**, con la potencia indicada. En el ejemplo anterior, si quisieramos calcular el tamaño muestral necesario para un tamaño del efecto de 0.20, con una potencia mínima del 85%:

```
test.t(m=-0.1546, s=1.1197, n=94, delta=0.2, potencia=0.85)
```

```
##
## Intervalo de confianza bilateral para la media de una VA normal
## -----
## Información muestral:
##   Tamaño muestral: n = 94
##   Media: m = -0.1546
##   Desviación típica: s = 1.1197
##   Error estándar de la media: sem = 0.1155
##
## Estimación:
##   95 %-IC( $\mu$ ): ( -0.3839 , 0.0747 )
##   Precisión obtenida: 0.2293
##
##
```

```
## # t-Test con una muestra
## # -----
## # Información muestral
##   n = 94
##   m = -0.1546
##   s = 1.1197
##   sem= 0.1155
##   95 %-IC(m)= ( -0.3839 , 0.0747 )
##
## # Valor de la media a contrastar: m0 = 0
##   texp = 1.3387 , g.l. = 93 , p = 0.1839
##   95%-IC(mu) = (-0.3839, 0.0747)
##
## # Estudio de la potencia: delta= 0.2 -> [-0.2, 0.2], potencia = 85%
##   70%-IC(mu) = (-0.275, -0.0342)
##
##   --(-)-|---]---   potencia < 85%
##
##   Leyenda:  --(---)--   --[---|---]--
##             IC- IC+   m0-d   m0   m0+d
##
## # Estimación del tamaño muestral para detectar delta=0.2 con potencia=85%
##   n >= 288 casos
```

7.3 Contrastes no paramétricos

Los métodos paramétricos vistos en 7.1 y 7.2, son válidos cuando las variables a contrastar siguen una distribución normal, aunque también pueden aplicarse cuando se dispone de un tamaño muestral suficientemente grande ($n > 60$). En los casos en los que los datos se desvían mucho de la normalidad, o bien el tamaño muestral es pequeño, se dispone de los métodos no paramétricos, que siempre son válidos, sea cual sea la distribución de la que proceden los datos.

7.3.1 Contraste no paramétrico para 2 muestras independientes

En el caso que se quieran comparar 2 muestras independientes, la opción no paramétrica es el **test de Mann-Whitney** o **test de Wilcoxon** para muestras independientes. Cuando el tamaño muestral es grande ($n_1 + n_2 > 30$) se aplica el método asintótico, que utiliza la aproximación a la normal. En otro caso, lo correcto es aplicar el método exacto.

7.3.1.1 Método asintótico (muestras grandes)

Veamos un ejemplo en el que el tamaño de muestra nos permite aplicar este método. Se quiere comparar si el *índice de masa corporal (imc)* de los pacientes del fichero **osteo**, difiere entre hombres y mujeres. Se ha realizado el test de Shapiro Wilks previamente, y se comprueba que los datos no siguen una distribución normal. Como $n_1 = 50$ y $n_2 = 57$, en este caso resulta adecuado usar el método asintótico del test de Wilcoxon. En **R**, puede usarse la función **wilcox.test** para este fin, indicando la siguiente orden al programa:

```
wilcox.test(imc~sexo)
```

Nótese que esta orden puede producir error si la variable **sexo** está definida como tipo numérico (lo que suele pasar si el archivo se abre importando el fichero desde SPSS). En ese caso habría que indicarle previamente al programa, que considere la variable como tipo factor:

```
fsexo <- factor(sexo, labels=c("Hombre", "Mujer"))
wilcox.test(imc~fsexo)
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: imc by fsexo
## W = 1060, p-value = 0.7506
## alternative hypothesis: true location shift is not equal to 0
```

Se obtiene un p-valor $p=0.7506$, por lo que no podemos rechazar H_0 , y se concluye que el *imc* está igualmente distribuido entre hombres y mujeres.

7.3.1.2 Método exacto (muestras pequeñas)

Para ilustrar este ejemplo, consideramos los datos de la estancia hospitalaria en días, de dos grupos de pacientes (**m1** y **m2**), de tamaño 8 y 6 respectivamente, con los que crearemos un data.frame:

```
m1<-c(1,1,1,2,3,5,5,7)
m2<-c(1,1,2,2,3,4)
datos<-data.frame(
  grupo=as.factor(rep(c("m1","m2"),c(8,6))),
  valores=c(m1,m2))
```

El tamaño muestral es pequeño, y suponemos que no se cumple la hipótesis de normalidad, por lo que se dan las condiciones apropiadas para aplicar el test exacto de Wilcoxon. En este caso, puede usarse la función **wilcox_test** del paquete **coin**, indicando el parámetro **distribution = "exact"**:

```
wilcox_test(valores~grupo, data=datos, distribution = "exact")
```

```
##
## Exact Wilcoxon-Mann-Whitney Test
##
## data: valores by grupo (m1, m2)
## Z = 0.59829, p-value = 0.5671
## alternative hypothesis: true mu is not equal to 0
```

El p-valor obtenido $p=0.5671$ nos conduce a aceptar la hipótesis nula, y concluir que la duración de la estancia hospitalaria está igualmente distribuida en ambos grupos.

Nota: la función **wilcox_test** del paquete **coin** también permite calcular el test asintótico, suprimiendo el parámetro **distribution = "exact"**. Sin embargo no se ha utilizado en esta práctica por no incluir la corrección por continuidad, vista en teoría.

7.3.2 Contraste no paramétrico para 2 muestras relacionadas

Al igual que ocurría en los test paramétricos, puede darse el caso de que los datos a comparar procedan de muestras relacionadas. En este caso también existe una versión del test no paramétrico de **Wilcoxon para muestras relacionadas o apareadas**. Bajo los mismos supuestos de tamaño muestral vistos en 7.3.1, existe un test asintótico y otro exacto.

7.3.2.1 Método asintótico (muestras grandes)

En el ejemplo visto en la sección 7.2, en el que se comparaba la *densidad de masa ósea* medida en el triángulo y cuello del mismo paciente, estas variables no siguen una distribución normal, lo que puede comprobarse con la función `shapiro.test`.

```
shapiro.test(osteo$szcue)
```

```
##
## Shapiro-Wilk normality test
##
## data:  osteo$szcue
## W = 0.96758, p-value = 0.01958
```

```
shapiro.test(osteo$sztri)
```

```
##
## Shapiro-Wilk normality test
##
## data:  osteo$sztri
## W = 0.97049, p-value = 0.03179
```

Para aplicar el test de Wilcoxon en este caso, escribiremos:

```
wilcox.test(szcue, sztri, paired=TRUE)
```

```
##
## Wilcoxon signed rank test with continuity correction
##
## data:  szcue and sztri
## V = 1809.5, p-value = 0.1111
## alternative hypothesis: true location shift is not equal to 0
```

Aceptándose en este caso la hipótesis nula con $p=0.1111$, por lo que se asume que ambas mediciones están igualmente distribuidas en triángulo y cuello.

7.3.2.2 Método exacto (muestras pequeñas)

También es posible aplicar el método exacto cuando el tamaño de muestra es pequeño o existen empates, que impiden la aplicación del método asintótico. Supongamos que tenemos los datos de estancia hospitalaria de un grupo de 8 mujeres en su primer y segundo parto (**p1** y **p2** respectivamente), con los que creamos un *data.frame*.

```
p1<-c(2,2,3,3,4,5,7,7)
p2<-c(1,1,2,2,2,2,3,3)
datos<-data.frame(
  grupo=as.factor(rep(c("p1","p2"),c(8,8))),
  valores=c(p1,p2))
```

En este caso utilizaremos la función `wilcoxsign_test` del paquete `coin`:

```
wilcoxsign_test(p1 ~ p2, distribution = "exact")
```

```
##  
## Exact Wilcoxon-Pratt Signed-Rank Test  
##  
## data: y by x (pos, neg)  
## stratified by block  
## Z = 2.5552, p-value = 0.007812  
## alternative hypothesis: true mu is not equal to 0
```

La salida del programa nos muestra un p-valor de 0.0078, que nos lleva a rechazar H_0 con un nivel de significación $\alpha = 0.05$, por lo que se concluye que la duración de la estancia hospitalaria del primer parto es distinta de la del segundo parto.

Nota: igual que se ha comentado en 7.3.2.1, la función **wilcoxsign_test** del paquete **coin**, también permite calcular el test exacto de Wilcoxon para muestras relacionadas, suprimiendo el término **distribution="exact"**, aunque no tiene en cuenta la corrección por continuidad vista en este curso.