

Cheat sheet Práctica 7: Test de comparación de dos medias

Juan Antonio Villatoro García – Grupo BioestadísticaR



**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



UNIVERSIDAD
DE GRANADA

Comprobación de los supuestos

Hipótesis de la normalidad: Test de Shapiro-Wilks

Función shapiro.test() por defecto en R

Caso comprobación normalidad 1 variable

shapiro.test(variable)

```
> shapiro.test(osteo$hba1c)
```

Shapiro-Wilk normality test

```
data:  osteo$hba1c  
W = 0.9833, p-value = 0.2757
```

Caso comprobación normalidad por grupos

tapply(variable, factor, shapiro.test)

```
> tapply(osteo$hba1c, osteo$sexo, shapiro.test)
```

\$Hombre

Shapiro-Wilk normality test

```
data:  X[[i]]  
W = 0.95798, p-value = 0.1022
```

\$Mujer

Shapiro-Wilk normality test

```
data:  X[[i]]  
W = 0.98002, p-value = 0.5661
```

Homogenidad de varianzas: F Test

Función var.test() por defecto en R

var.test(variable ~ factor)

```
> var.test(osteo$hba1c ~ osteo$sexo)
```

F test to compare two variances

```
data:  osteo$hba1c by osteo$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
```

Comparación de dos muestras independientes. Test *t* de Student y Welch

Función t.test() por defecto en R

Test t de student

t.test(variable~factor, var.equal = TRUE)

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

Two Sample t-test

```
data:  osteo$hba1c by osteo$sexo  
t = -0.059393, df = 92, p-value = 0.9528  
alternative hypothesis: true difference in means between group  
Hombre and group Mujer is not equal to 0  
95 percent confidence interval:  
 -0.7637643 0.7194106  
sample estimates:  
mean in group Hombre mean in group Mujer  
 8.553333 8.575510
```

Test de Welch

t.test(variable~factor, var.equal = FALSE)

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

welch Two Sample t-test

```
data:  osteo$hba1c by osteo$sexo  
t = -0.059562, df = 91.963, p-value = 0.9526  
alternative hypothesis: true difference in means between group  
Hombre and group Mujer is not equal to 0  
95 percent confidence interval:  
 -0.7616689 0.7173151  
sample estimates:  
mean in group Hombre mean in group Mujer  
 8.553333 8.575510
```

Práctica 7: : Test de comparación de dos medias



UNIVERSIDAD
DE GRANADA

Comparación de medias de muestras independientes con el paquete BioestadisticaR: Función test.t()

Función aplicada a variables

`test.t(m = variable, grupos = factor)` $\longrightarrow$ `> test.t(m=osteo$peso, grupos=osteo$sexo)`

Información muestral

```
# t-test para 2 Muestras Independientes
#-----
# Información muestral
# Niveles de agrupacion: Hombre , Mujer

[1] Para grupo = Hombre
     n = 45
     m = 66.9311
     s = 10.8468
     sem= 1.6169
     95 %-IC(m)= ( 63.6724 , 70.1898 )

[2] Para grupo = Mujer
     n = 49
     m = 61
     s = 12.0388
     sem= 1.7198
     95 %-IC(m)= ( 57.5421 , 64.4579 )
```

Pruebas de normalidad

```
# Pruebas de normalidad (test de Shapiro-Wilk)
[1] Para grupo = Hombre , W =0.928, g.l. =45, p =0.008
[2] Para grupo = Mujer , W =0.924, g.l. =45, p =0.0037
```

Test de homogeneidad de varianzas

```
# Test de homogeneidad de varianzas. Fexp = (var[2]/var[1])
Fexp =1.2319, gl1 = 48, gl2 = 44, p = 0.486
```

Test de Student y Welch

```
# 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)
```

Función para información muestral resumida

`test.t(n1, m1, s1, n2, m2, s2)`

n1: tamaño muestral de la primera muestra

m1: media de la primera muestra

s1: desviación típica de la primera muestra

n2: tamaño muestral de la segunda muestra

m2: media de la segunda muestra

s2: desviación típica de la segunda muestra

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

```
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)
```

Fijar la potencia y detectar el tamaño de efecto (delta, δ)

Función `test.t()` con mismo parámetros que en el caso anterior, pero añadiendo los parámetros potencia y delta. Se obtienen los mismo resultados más el estudio de la potencia

`test.t(..., potencia, delta)` $\longrightarrow$ **potencia:** potencia del test
delta: tamaño del efecto

```
> test.t(n1=50, m1=63, s1=10, n2=67,
+        m2=58, s2=20, delta=5, potencia=0.85)
```

```
# Estudio de la potencia: delta= 5 -> [-5, 5], potencia = 85%
70%-IC(mu1-mu2) = (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]
```

© Juan Antonio Villatoro-García – Grupo BioestadísticaR

Práctica 7: : Test de comparación de dos medias



UNIVERSIDAD
DE GRANADA

Contrastes no paramétricos (test de Wilcoxon)

Contraste no paramétricos para 2 muestras independientes

Método asintótico (muestras grandes)

Función `wilcox.test` por defecto en R

`Wilcox.test(variable~factor)`

```
> wilcox.test(osteo$imc~osteo$sexo)
```

Wilcoxon rank sum test with continuity correction

data: osteo\$imc by osteo\$sexo

W = 1060, p-value = 0.7506

alternative hypothesis: true location shift is not equal to 0

Método exacto (muestras pequeñas)

Función `wilcox_test` del paquete `coin`

`Wilcox_test(variable~factor, distribution = "exact")`

```
> library(coin)
> m1<-c(1,1,1,2,3,5,5,7)
> m2<-c(1,1,2,2,3,4)
> grupo=as.factor(rep(c("m1","m2"),c(8,6)))
> valores = c(m1,m2)
> wilcox_test(valores~grupo, 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

Contraste no paramétricos para 2 muestras relacionadas

Método asintótico (muestras grandes)

Función `wilcox.test()` por defecto en R añadiendo `paired = TRUE`

`Wilcox.test(variable~factor, paired = TRUE)`

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

Wilcoxon signed rank test with continuity correction

data: osteo\$szcue and osteo\$sztri

V = 1809.5, p-value = 0.1111

alternative hypothesis: true location shift is not equal to 0

Método exacto (muestras pequeñas)

Función `wilcoxsign_test()` del paquete `coin`

`Wilcoxsign_test(variable~factor, distribution = "exact")`

```
> p1<-c(2,2,3,3,4,5,7,7)
> p2<-c(1,1,2,2,2,2,3,3)
> wilcoxsign_test(p1~p2, ditribution = "exact")
```

Asymptotic Wilcoxon-Pratt Signed-Rank Test

data: y by x (pos, neg)

stratified by block

Z = 2.5552, p-value = 0.01061

alternative hypothesis: true mu is not equal to 0