

Prueba de los rangos con signo de Wilcoxon alternativa no paramétrica al t-test de medias dependientes

Joaquín Amat Rodrigo j.amatrodrigo@gmail.com

Enero, 2016

Índice

Ventajas y desventajas de <i>Prueba de los rangos con signo de Wilcoxon</i> frente a <i>t-test</i> de datos pareados.	2
Condiciones para la prueba de los rangos con signo de Wilcoxon	3
Ejemplo.....	3
Solución directa con R	6

Ventajas y desventajas de *Prueba de los rangos con signo de Wilcoxon* frente a *t-test* de datos pareados.

El test no paramétrico *prueba de los rangos con signo de Wilcoxon*, también conocido como *Wilcoxon signed-rank test* permite comparar poblaciones cuando sus distribuciones (normalmente interpretadas a partir de las muestras) no satisfacen las condiciones necesarias para otros test paramétricos. Son una alternativa a los *t-test para muestras dependientes* cuando las muestras no siguen una distribución normal (muestran asimetría o colas) o porque tienen un tamaño demasiado reducido para poder determinar si realmente las poblaciones de las que proceden se distribuyen de forma normal.

A la hora de elegir entre *t-test* o *Wilcoxon signed-rank test*, es importante tener en cuenta que el problema de las muestras pequeñas no se soluciona con ninguno de los dos. Si el tamaño de las muestras es pequeño también lo es la calidad de la inferencia que se puede hacer con ellas. Ahora bien, existen dos situaciones en las que *a priori* se puede recomendar utilizar un *Wilcoxon signed-rank test* antes que un *t-test*:

- Si el tamaño de las muestras es suficientemente grande para determinar (por métodos gráficos o test analíticos) que la distribución de las poblaciones a comparar no es de tipo normal, en tal caso los *t-test* no son adecuados, por lo que mejor emplear los *Wilcoxon signed-rank test* (Bootstrapping también podría ser una alternativa).
- Si el tamaño de la muestra es tal que no se puede determinar con seguridad si las poblaciones de las que proceden se distribuyen de forma normal, y no se dispone de información que pueda orientar sobre la naturaleza de las poblaciones de origen (estudios anteriores, que sea un tipo de variable que se sabe que se distribuye casi siempre de forma normal...) entonces puede ser más apropiado el *Wilcoxon signed-rank test* ya que no requiere asumir la normalidad de las poblaciones.

El test *Wilcoxon signed-rank test* se caracteriza por:

- Compara si las diferencias entre pares de datos siguen una distribución simétrica entorno a un valor. Si las dos muestras proceden de la misma población, es de esperar que las diferencias entre cada par de observaciones se distribuyan de forma simétrica entorno al cero. A pesar de que la definición anterior es la verdadera y correcta, es frecuente encontrar que descrito que el *Wilcoxon signed-rank test* compara la mediana de las diferencias. Esto no es del todo correcto pero es más fácil de entender.

- Trabajan sobre rangos de orden, es decir, utilizan las posiciones que ocupan los datos una vez ordenados. Por lo tanto solo es aplicable a variables cuyos valores se pueden ordenar.
- Son menos potentes (tienen menos probabilidad de rechazar la H_0 cuando realmente es falsa) ya que ignoran valores extremos (en el caso de los t-test al trabajar con medias si los tienen en cuenta). Esto a su vez hace que los Wilcoxon signed-rank test sean una prueba más robusta que los t-test.

Condiciones para la prueba de los rangos con signo de Wilcoxon

- Los datos tienen que ser independientes.
- Los datos tienen que ser ordinales, se tienen que poder ordenar de menor a mayor o viceversa.
- No es necesario asumir que las muestras se distribuyen de forma normal o que proceden de poblaciones normales. Pero sea cual sea el tipo de distribución de las diferencias, tiene que ser simétrica.
- A pesar de considerarse el equivalente no paramétrico del t-test, el Wilcoxon signed-rank test trabaja con medianas, no con medias.
- Preferible al t-test cuando hay valores atípicos, no hay normalidad de los datos o el tamaño de las muestras es pequeño.

Ejemplo

Supóngase que se dispone de dos muestras pareadas, de las que no se conoce el tipo de distribución de las poblaciones de las que proceden y cuyo tamaño es demasiado pequeño para determinar si siguen una distribución normal. ¿Existe una diferencia significativa? *Se emplea un ejemplo con muestras pequeñas para poder ilustrar fácilmente los pasos, no significa que con muestras tan pequeñas el Wilcoxon Signed-rank test sea muy preciso.*

antes	<- c(2, 5, 4, 6, 1, 3)
despues	<- c(5, 6, 2, 7, 1, 6)

1.Hipotesis

- H_0 : La mediana de las diferencias de cada par de datos es cero. $Mediana(diferencias) = 0$.
- H_a : La mediana de las diferencias entre cada par de datos es diferente de cero. $Mediana(diferencias) \neq 0$.
-

2.Cálculo de las diferencias y ordenación de las observaciones

Se calcula la diferencia de cada par de observaciones.

```
diferencias <- c(antes - despues)
rbind(antes, despues, diferencias)
```

```
##           [,1] [,2] [,3] [,4] [,5] [,6]
## antes      2   5   4   6   1   3
## despues    5   6   2   7   1   6
## diferencias -3  -1   2  -1   0  -3
```

Se asigna al valor absoluto de cada diferencia un número de posición, siendo la posición 1 para el valor más pequeño. En caso de haber valores repetidos, fenómeno conocido como ligadura o *ties*, se les asigna como valor de posición la media de las posiciones que ocupan. En la asignación de posiciones se ignoran las diferencias que sean 0.

```
rangosdiferencias <- rank(abs(diferencias[diferencias != 0]))
# La función rank() de R calcula las posiciones automáticamente,
# solucionando las ligaduras en caso de que las haya.
rangosdiferencias <- c(rangosdiferencias[1:4], 0, rangosdiferencias[5])
# Se incorpora un 0 en la posición que ocupan las diferencias nulas para
# cuadrar las posiciones
```

Se compila toda la información:

```
tabla <- data.frame(antes = antes, despues = despues, signo = sign(diferencias),
  diferencia = abs(diferencias), rangos = rangosdiferencias)
tabla
```

```
##      antes  despues  signo  diferencia  rangos
## 1      2      5     -1         3      4.5
## 2      5      6     -1         1      1.5
## 3      4      2      1         2      3.0
## 4      6      7     -1         1      1.5
## 5      1      1      0         0      0.0
## 6      3      6     -1         3      4.5
```

3.Cálculo del estadístico W

$$W = \min(W+, W-)$$

$W +$ = suma de los rangos con signo positivo

$W -$ = suma de los rangos con signo negativo

```
sumaPositivos <- sum(tabla[tabla$signo == 1,"rangos"])
sumaNegativos <- sum(tabla[ tabla$signo == -1,"rangos"] )
W <- min(c( sumaPositivos, sumaNegativos ))
W
```

```
## [1] 3
```

4.Cálculo de *p-value*

Una vez ha obtenido el valor del estadístico W se puede calcular cual es la probabilidad de que adquiera valores igual o más extremos que el observado.

- Si el tamaño $n < 25$, se compara el valor obtenido de W con los valores de una tabla Wilcoxon. Si W cae dentro del intervalo correspondiente en la tabla para ese n , la diferencia NO es significativa.
- Si $n > 25$ se puede asumir que W se distribuye de forma aproximadamente Normal, rechazando H_0 si Z calculado es menor que el valor de Z para el α (esta es la que utiliza la función de R).

$$Z = \frac{W - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{12}}} \sim N(0,1)$$

6. Cálculo del tamaño del efecto (side effect)

Parece indicarse que la forma de calcular el tamaño del efecto de un test de Wilcoxon es mediante la correlación de rangos "*Rank correlation*".

Solución directa con R

R contiene una función llamada `wilcox.test()` que realiza el test de Wilcoxon entre dos muestras cuando se indica que *paired*= *TRUE*.

```
wilcox.test(x = antes, y = despues, alternative = "two.sided", mu = 0, paired = TRUE)
```

```
## Warning in wilcox.test.default(x = antes, y = despues, alternative =  
## "two.sided", : cannot compute exact p-value with ties  
  
## Warning in wilcox.test.default(x = antes, y = despues, alternative =  
## "two.sided", : cannot compute exact p-value with zeroes  
  
##  
## Wilcoxon signed rank test with continuity correction  
##  
## data:  antes and despues  
## V = 3, p-value = 0.2763  
## alternative hypothesis: true location shift is not equal to 0
```

En la salida devuelta por R el estadístico se denomina V en lugar de W.

Cuando hay ligaduras o *ties*, `wilcox.test()` no es capaz de calcular el *p-value* exacto por lo que devuelve un *p-value* aproximado asumiendo que W se distribuye de forma ~ Normal. En estos casos, o cuando los tamaños muestrales son mayores de 25 y se quiere la aproximación por la Normal, es más recomendable emplear la función `wilcoxsigned_test()` del paquete *coin* que devuelve el valor exacto de p-value en lugar de una aproximación.

```
require(coin)
# La función wilcoxsigned_test() del paquete coin requiere pasarle los
# argumentos en forma de función (~), por lo que los datos tienen que estar
# almacenados en forma de data frame.

datos <- data.frame(antes = antes, despues = despues)
wilcoxsign_test(antes ~ despues, data = datos, distribution = "exact")
```

```
##
## Exact Wilcoxon-Pratt Signed-Rank Test
##
## data: y by x (pos, neg)
## stratified by block
## Z = -1.272, p-value = 0.25
## alternative hypothesis: true mu is not equal to 0
```