

# Test de Friedman. Alternativa no paramétrica al ANOVA de datos dependientes.

Joaquín Amat Rodrigo [j.amatrodrigo@gmail.com](mailto:j.amatrodrigo@gmail.com)

Enero, 2016

## Índice

Introducción.....	2
Comparaciones <i>Post-Hoc</i> .....	2
Tamaño del efecto.....	3
Ejemplo .....	3
Observaciones.....	5

## Introducción

El test de Friedman es la alternativa no paramétrica a la prueba ANOVA de una vía cuando los datos son dependientes/pareados. Se trata de una extensión de la prueba de los *rangos con signo de Wilcoxon* para más de dos grupos (basada en suma de rangos). Asumiendo ciertas simplificaciones, puede considerarse como una comparación entre las medianas de varios grupos.

El test de Friedman es el test adecuado cuando los datos tienen un orden natural, (cuando para darles sentido tienen que estar ordenados) y además son pareados. Por ejemplo, si se quiere estudiar la diferencia en el rendimiento de un grupo de corredores dependiendo de la estación del año, se hace correr al mismo grupo de personas una vez en cada estación. Como resultado se puede disponer de dos tipos de datos: los tiempos de cada participante (análisis con ANOVA pareado) o las posiciones en las que han terminado la carrera cada participante en cada una de las carreras (análisis con Friedman test).

EL concepto es similar a la suma de rangos de Wilcoxon para dos grupos pero esta vez asignando los rangos dentro de cada fila (ya que hay más de dos medidas para cada individuo) y después haciendo la suma de rangos para cada columna.

El test de Friedman genera un estadístico conocido como *Fr* o *Q* que se distribuye:

- Si el número total de individuos (N) es mayor de 10 la distribución de *Fr* se aproxima a una distribución  $\chi^2$  con  $k-1$  grados de libertad (siendo  $k$  el número de grupos a comparar).
- Si el número de individuos es menor de 10 se recurre a tablas con los valores de significancia para un test de Friedman.

## Comparaciones *Post-Hoc*

Al igual que ocurre con un ANOVA, si el test de Friedman es significativo implica que al menos dos grupos de entre los comparados son significativamente diferentes, pero no indica cuales. Para saberlo es necesario compararlos todos entre ellos. Esto implica realizar una corrección del nivel de significancia para evitar incrementar el error de tipo I. Los dos métodos de comparación post-hoc más empleados para un test de Friedman son:

- Test de *rangos con signo de Wilcoxon* entre cada par de grupos con corrección de significancia `pairwise.wilcox.test( paried = TRUE )`.
- Tukey's range test: en R existe la función `friedmanmc()` del paquete *pgirmess*.

## Tamaño del efecto

No hay una manera general de calcular el tamaño del efecto de una prueba de Friedman. Lo que se suele hacer es calcular el tamaño del efecto sobre las parejas que hayan resultado significativamente distintas en el análisis *post-hoc*. Si las comparaciones *post-hoc* se hacen mediante *prueba de rangos con signo de Wilcoxon*, se tendrá que calcular el efecto asociado a dicho test (ver capítulo dedicado al test).

## Ejemplo

*Un estudio pretende determinar si existe diferencia en como de bueno consideran los consumidores que es un vino dependiendo de la hora del día en la que lo toman. Para ello se selecciona a un grupo de 7 sujetos a los que se les da a probar un vino por la mañana, por la tarde y por la noche. En cada degustación se valora del 1 al 11 el vino (los degustadores no saben que es el mismo vino).*

```
valoracion <- c( 9, 5, 2, 6, 3, 1, 5, 5, 5, 11, 5, 1, 8, 4, 3, 10, 4, 1, 7, 3, 4 )
hora <- factor( rep( c( "mañana", "tarde", "noche" ), 7 ) )
sujeto <- factor( rep( 1:7, each = 3 ) )
datos <- data.frame( valoracion, hora, sujeto )
head(datos)
```

```
##   valoracion   hora sujeto
## 1          9 mañana     1
## 2          5  tarde     1
## 3          2  noche     1
## 4          6 mañana     2
## 5          3  tarde     2
## 6          1  noche     2
```

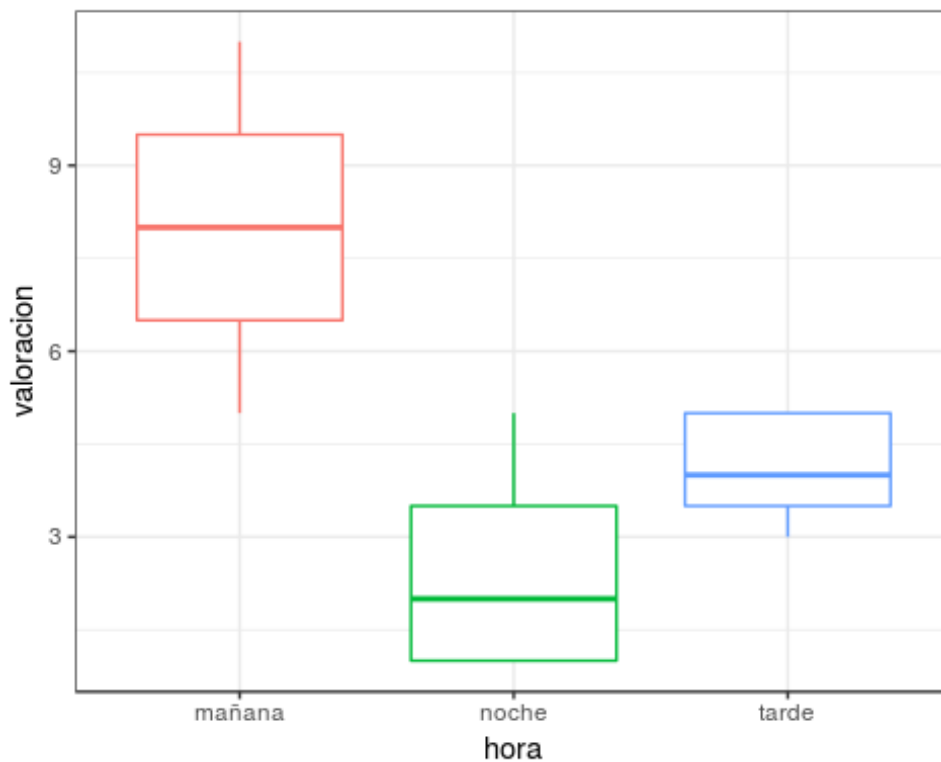
```
by(data = datos$valoracion, INDICES = datos$hora, FUN = median)
```

```
## datos$hora: mañana
## [1] 8
## -----
```

```
## datos$hora: noche
## [1] 2
## -----
## datos$hora: tarde
## [1] 4
```

## Representación gráfica de los datos

```
library(ggplot2)
ggplot(data = datos, mapping = aes(x = hora, y = valoracion, colour = hora)) +
  geom_boxplot() +
  theme_bw() +
  theme(legend.position = "none")
```



El *box-plot* parece indicar que sí hay diferencia dependiendo de la hora del día. Al tratarse de escalas de valoración (valores ordinales) no se puede emplear un ANOVA, se emplea el test de Friedman.

## Test de Friedman

```
friedman.test(valoracion, hora, sujeto)

##
##  Friedman rank sum test
##
## data:  valoracion, hora and sujeto
## Friedman chi-squared = 10.333, df = 2, p-value = 0.005704
```

El test encuentra diferencias significativas entre al menos dos grupos. Para saber cuáles son, se realizan comparaciones dos a dos empleando el test de Wilcoxon rangos con signo.

## Comparaciones múltiples

```
pairwise.wilcox.test(datos$valoracion, datos$hora, paired = TRUE, p.adjust.method =
"holm")

##
##  Pairwise comparisons using Wilcoxon signed rank test
##
## data:  datos$valoracion and datos$hora
##
##      mañana noche
## noche 0.1      -
## tarde 0.1      0.1
##
## P value adjustment method: holm
```

A pesar de que el test de Friedman resulta significativo, las comparaciones por pares mediante *wilcox* test con corrección de *holm* no encuentran ninguna diferencia significativa.

## Observaciones

El test de Friedman parece tener ciertos inconvenientes. Para ver los detalles consultar <https://www.r-bloggers.com/beware-the-friedman-test/>