

## Seminario del IESTA (SIESTA)

# Tablas exactas para la prueba de rangos de Friedman en el caso de empates

Exponente: Esther Hochsztain - Departamento de Métodos Cuantitativos –  
Facultad de Ciencias Económicas y de Administración

Trabajo en conjunto con

- Dr. Ing. Carlos Lopez-Vazquez LatinGEO Lab IGM+ORT, Universidad  
ORT Uruguay, URUGUAY
- M.Sc., Ing. Andromaca Tasistro, PMP

# Plan de la presentación

- **Introducción**
- Test de rangos de Friedman
- Test de rangos de Friedman con empates
- Generación de tablas exactas
- Ejemplo
- Conclusiones y trabajos futuros
- Agradecimiento
- Bibliografía

# Introducción

- Presentación basada en el artículo
- Exact Tables for the Friedman's rank Test for the case with ties
- Aceptado para ser publicado en el Chilean Journal of Statistics

# Introducción

- Las tablas exactas para el estadístico del Test de Rangos de Friedman (1937) estuvieron disponibles desde su propuesta.
- Para otros casos, se dispone de una aproximación asintótica.
- Un estadístico modificado apropiado para el caso de empates recién fue propuesto en 1980.
- Presentamos por primera vez las tablas exactas para un reducido conjunto de casos, llenando así un vacío.
- Ilustramos el efecto de usar las tablas correctas con un ejemplo.

# Plan de la presentación

- Introducción
- **Test de rangos de Friedman**
- Test de rangos de Friedman con empates
- Generación de tablas exactas
- Ejemplo
- Conclusiones y trabajos futuros
- Agradecimiento
- Bibliografía

# Test de Friedman (1937)

- Válido para analizar tablas de rankings: N filas, k opciones
- Ejemplo: N jueces catando k vinos
- $H_0$ : los jueces deciden al azar
- Dada la tabla con rankings  $r_{ij}$ , el estadístico propuesto *asumiendo que no hay empates* es

$$\chi_r^2 = \frac{12}{Nk(k+1)} \sum_{j=1}^k \left( \sum_{i=1}^N r_{ij} \right)^2 - 3N(k+1)$$

- El estadístico se puede aproximar
  - por la distribución  $\chi^2$  para N grande,
  - por la distribución normal para k grande y N pequeño

# Plan de la presentación

- Introducción
- Test de rangos de Friedman
- **Test de rangos de Friedman con empates**
- Generación de tablas exactas
- Ejemplo
- Conclusiones y trabajos futuros
- Agradecimiento
- Bibliografía

## Test de rangos de Friedman con empates

- La corrección del estadístico del Test de rangos de Friedman para el caso de empates, se presentó en varios libros
- Se presentaron distintas expresiones, todas numéricamente equivalentes



# Test de rangos de Friedman con empates

■ Conover, 1980

$$\chi_r = \frac{(k-1) \sum_{j=1}^k [R_j - N(k+1)/2]^2}{\sum_{i=1}^N \sum_{j=1}^k r_{ij}^2 - Nk(k+1)^2/4}$$

■ Siegel and Castellan, 1988

$$\chi_r = \frac{12 \sum_{j=1}^k R_j^2 - 3N^2k(k+1)^2}{Nk(k+1) + \frac{\left( Nk - \sum_{i=1}^N \sum_{j=1}^{g_i} t_{i,j}^3 \right)}{k-1}}$$

■ Corder and Foreman, 2009

$$CF = \frac{Nk}{4} (k+1)^2$$

$$\chi_r^2 = \frac{N(k-1) \left[ \sum_{i=1}^k \frac{R_i^2}{N} - CF \right]}{\sum_{j=1}^k r_{ij}^2 - CF}$$

## Test de rangos de Friedman con empates

- A pesar de presentarse un estadístico apropiado para el caso de empates
- No se disponía con tablas asociadas
  - ⇒ Se continuaron usando las tablas exactas para el caso sin empates, asumiendo implícitamente su validez

# Plan de la presentación

- Introducción
- Test de rangos de Friedman
- Test de rangos de Friedman con empates
- **Generación de tablas exactas**
- Ejemplo
- Conclusiones y trabajos futuros
- Agradecimiento
- Bibliografía

# Generación de tablas exactas

## Etapas

1. Generación de todas las posibles alternativas considerando  $k$  rangos (con y sin empates)
2. Generación de todas las posibilidades para  $N$  casos y  $k$  rangos (con y sin empates)
3. Cálculo de valor del estadístico propuesto por Corder and Foreman, 2009 para todas las tablas con  $N$  filas y  $k$  columnas posibles anteriormente determinadas
4. Construcción de la tabla exacta para cada pareja  $(N,k)$ 
  - Lenguajes de programación: **R** para 1, 2 y 3; **Octave** para 4

# Tiempo de procesamiento.

- El tiempo de cálculo en la segunda etapa es manejable para  $k$  y  $N$  pequeños
- Crece rápidamente a medida que  $k$  y  $N$  aumentan.
- Las combinaciones válidas se agruparon en conjuntos y se computaron de forma independiente para realizar el procesamiento en paralelo

# Tiempo de procesamiento

- Por ejemplo: para el caso  $k=3$   $N=7$  se requirieron más de 90 días
    - Usando hasta 100 nodos en paralelo en un cluster LINUX (Intel Xeon-Gold 6138 2.00GHz procesadores.)
  - Por consiguiente, debido a la naturaleza combinatoria del problema, nuestro enfoque es capaz de abordar casos modestos en un tiempo razonable.
- ⇒ Problema abierto: mejorar los tiempos de procesamiento!!

# Plan de la presentación

- Introducción
- Test de rangos de Friedman
- Test de rangos de Friedman con empates
- Generación de tablas exactas
- **Ejemplo**
- Conclusiones y trabajos futuros
- Agradecimiento
- Bibliografía

# Ejemplo

- Ejemplo propuesto por Corder y Foreman (2009)
- Presenta los rangos para siete empleados ( $N = 7$ ) bajo tres alternativas ( $k = 3$ ) para lidiar con su ‘tardanza’:
  - a) no hacer nada (denotado como *baseline*),
  - b) un mes con un incentivo monetario de \$ 10,
  - c) otro mes con doble incentivo.
- Objetivo: determinar si alguna de las estrategias modificó la ‘tardanza’ de los empleados.



# Ejemplo

## ■ Tabla de datos

Employee	Rank of Monthly Tardiness		
	Baseline	Month 1	Month 2
1	2	3	1
2	3	2	1
3	2	3	1
4	3	2	1
5	3	1.5	1.5
6	2.5	2.5	1
7	3	2	1

Valor del estadístico:  $\chi_r = 10.23$

*(con la fórmula correcta, considerando empates)*

# Ejemplo

- Para  $\alpha=0.05$  Corder and Foreman (2009) indican que el valor crítico es 7.14
  - Este valor corresponde al caso exacto de tablas *sin* empates
  - Dado que el valor del estadístico es 10.23  $\Rightarrow$  la decisión es **Rechazar la Hipótesis Nula.**
  - Utilizando las tablas que proponemos para el caso con empates el valor crítico es 5.769  $\Rightarrow$  la decisión coincide: **Rechazar la Hipótesis Nula.**
- Si se considera  $\alpha=0.005$  **las decisiones no coinciden**
  - El valor crítico usando la tabla exacta sin empates es 10.286  $\Rightarrow$  la decisión es **No rechazar la Hipótesis Nula**
  - El valor crítico utilizando las tablas que proponemos para el caso con empates es 9.083  $\Rightarrow$  la decisión es **Rechazar la Hipótesis Nula**

# Plan de la presentación

- Introducción
- Test de rangos de Friedman
- Test de rangos de Friedman con empates
- Generación de tablas exactas
- Ejemplo
- **Conclusiones y trabajos futuros**
- Agradecimiento
- Bibliografía

# Conclusiones

- Se cumplió el objetivo de presentar por primera vez tablas exactas para el test de rangos de Friedman en el caso de empates
- Se generaron las tablas para
  - $k=3$  y  $N$  hasta 11
  - $k=4$  y  $N$  hasta 6
  - $k=5$  y  $N$  hasta 4
  - $k=6$  y  $N$  hasta 3
  - $k=7$  y  $N=2$

# Trabajos futuros

- Generar tablas exactas para el test de Friedman con empates para valores mayores de  $k$  y  $N$ , basados en algoritmos del estilo de Van de Wiel (2000) y Van de Wiel *et al.*, (1999).
- Identificar en la literatura casos en que se incurre en decisiones equivocadas al utilizar tablas exactas del caso sin empates, existiendo empates

# Plan de la presentación

- Introducción
- Test de rangos de Friedman
- Test de rangos de Friedman con empates
- Generación de tablas exactas
- Ejemplo
- Conclusiones y trabajos futuros
- **Agradecimiento**
- Bibliografía

# Agradecimiento

- El trabajo se desarrolló utilizando el ClusterUY <https://cluster.uy>
- Agradecemos al Dr. Sergio Nesmachnow y al Dr. Gabriel Usera por brindarnos el acceso al mismo, sin el cual este trabajo no se hubiera podido concretar

# Bibliografía

- Conover, W. J. (1980) Practical Nonparametric Statistics 2nd. Edition, Wiley, 462 pp.
- Corder, G. W. and Foreman, D. I. (2009) Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach. John Wiley & Sons, 247 pp.
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. Journal of the American Statistical Association, 32, 200, 675-701. <http://dx.doi.org/10.2307/2279372>
- Hollander, M. & Wolfe, D.A. (1973). Nonparametric statistical methods. Wiley, New York, 503 pp.
- Kendall, M.G. & Babington-Smith, B. (1939). The problem of m rankings. Annals of Mathematical Statistics, 10, 3, 275-287. <http://dx.doi.org/10.1214/aoms/1177732186>
- López-Vázquez, C. and Hochsztain, E. (2019) Extended and updated tables for the Friedman rank test. Communications in Statistics - Theory and Methods, 48, 2, 268-281. <http://dx.doi.org/10.1080/03610926.2017.1408829>
- Martin, L., Leblanc, R. and Toan, N. (1993). Tables for the Friedman rank test. Canadian Journal of Statistics, 21, 1, 39-43. <http://dx.doi.org/10.2307/3315656>
- Odeh, R. E. (1977). Extended tables of the distribution of Friedman's S-statistic in the two way layout. Communications in Statistics-Simulation and Computation, 6 (1): 29-48. <https://doi.org/10.1080/03610917708812025>



# Bibliografía

- Owen, D. B. (1962) Handbook of statistical tables. Addison-Wesley Publishing Company, Reading, Massachusetts. 580 pp.
- Schneider, G., Chicken, E. and Becvarik, R. (2016). NSM3 Package: Functions and Datasets to Accompany Hollander, Wolfe, and Chicken - Nonparametric Statistical Methods, Third Edition. <http://CRAN.R-project.org/web/packages/NSM3/NSM3.pdf> (Accessed September 2019).
- Sen, P. K.; Salama, I. A.; and Quade, D. 2011. Spearman's footrule: Asymptotics in applications. Chilean Journal of Statistics 2(1), 3–20
- Siegel, S., & Castellan, N. J., Jr. (1988). Nonparametric statistics for the behavioral sciences (2nd ed.). New York, NY, England: McGraw-Hill Book Company.
- Van de Wiel, M. A. (2000). Exact distributions of distribution-free test statistics. Ph.D. Thesis, Eindhoven University of Technology, The Netherlands, 165 pp.
- Van de Wiel, M. A.; Di Bucchianico, A. and Van Der Laan, P. (1999) Symbolic Computation and Exact Distributions of Nonparametric Test Statistics. Journal of the Royal Statistical Society: Series D (The Statistician), 48, 4, 507-516. <http://dx.doi.org/10.1111/1467-9884.00208>

## Seminario del IESTA (SIESTA)

# Muchas gracias por su atención

¿Comentarios?  
¿Preguntas?

**Tablas exactas para la prueba de rangos de Friedman  
en el caso de empates**