



UNIVERSIDAD NACIONAL DE ROSARIO
FACULTAD DE CIENCIAS ECONÓMICAS Y ESTADÍSTICA
SECRETARIA DE CIENCIA Y TECNOLOGIA E INSTITUTOS DE INVESTIGACIONES

Resumen Ampliado

Jornadas Anuales

“Investigaciones en la Facultad”

Ciencias Económicas y Estadística



Bussi, Javier; Prunello, Marcos; Toledo, Paulina

Instituto de Investigaciones Teóricas y Aplicadas, Escuela de Estadística.

UN ESTUDIO COMPARATIVO DE LA POTENCIA DE LOS TESTS DE KOLMOGOROV-SMIRNOV Y WALD-WOLFOWITZ PARA DOS MUESTRAS INDEPENDIENTES EN EL CASO DE LA DISTRIBUCIÓN NORMAL¹

Resumen

El Test de Kolmogorov-Smirnov (K-S) y el Test de Rachas de Wald-Walfowitz (W-W) para dos muestras independientes sirven para ensayar la misma prueba de hipótesis que evalúa si ambas poblaciones cuentan con igual función de distribución de probabilidad. Si bien son pruebas muy difundidas, no se han encontrado antecedentes en la revisión bibliográfica que determinen cuál de las dos pruebas es más potente. En este trabajo se realizó un estudio de simulación en donde se seleccionaron distintos valores para los parámetros de las dos poblaciones normales a comparar. Se determinaron diversos tamaños muestrales, tanto iguales como diferentes. En total se plantearon 1.032 escenarios y se realizaron 1.000 repeticiones de cada uno de ellos. En cada simulación se computó una estimación empírica de la potencia para ambas pruebas. El test de K-S resultó ser más potente cuando las distribuciones normales difieren solo en posición. En el caso de que las muestras provengan de distribuciones normales que difieran solo en dispersión, el test de W-W resultó ser más potente. Cuando las poblaciones difieren en localización y en dispersión, K-S presenta mayores potencias en la mayoría de los casos. En términos generales, en este estudio, se puede concluir que K-S es superior en cuanto a potencia que el Test de Rachas, en el caso que las muestras provengan de distribuciones normales (71% de los escenarios simulados).

Palabras clave: Kolmogorov-Smirnov, Rachas de Wald-Walfowitz, Potencia

Abstract

The Kolmogorov-Smirnov Test (K-S) and the Wald-Walfowitz Runs Test (W-W) for two independent samples consider the same hypothesis that evaluates whether both populations have equal probability distribution function. Although these tests are very well known, no precedents have been found in the bibliographic review that determine which of the two tests is more powerful. In this work, a simulation study was carried out where different values were selected for the parameters of the two normal populations to be compared and several sample sizes (equal and different) were determined. In total, 1,032 scenarios were proposed and 1,000 repetitions of each of them were carried out. In each simulation, an empirical estimate of power was computed for both tests. The K-S test turned out to be more powerful when the normal distributions differed only in location. In the case that the samples come from normal distributions that differ only in dispersion, the W-W test turned out to be more powerful. When populations differ in location and dispersion, K-S presents higher powers in most cases. In general terms, in this study, it can be concluded that K-S is superior in terms of power than the Run Test, in the case that the samples come from normal distributions (71% of the simulated scenarios).

Keywords: Kolmogorov-Smirnov, Wald-Walfowitz Runs Test, Power

¹ Trabajo elaborado en el marco del Proyecto (80120210200167UR), titulado: "Métodos muestrales y series de tiempo en las estadísticas oficiales", dirigido por Mg. Gonzalo P.D. Marí.



Introducción

El Test de Kolmogorov-Smirnov y el Test de Rachas de Wald-Walfowitz para dos muestras independientes sirven para ensayar las mismas hipótesis. La hipótesis nula especifica que ambas poblaciones cuentan con la igual función de distribución de probabilidad. La hipótesis alternativa sostiene que esas distribuciones son distintas Sin embargo no se han encontrado antecedentes en la revisión bibliográfica que determinen cuál de las dos pruebas es más potente. Se desea estudiar si una de ellas posee una potencia superior brindando así una prueba que cuente con mejores cualidades para evaluar el test de hipótesis planteado. Obtener las funciones de potencia de manera explícita y calcular eficiencias relativas asintóticas representa un desafío muy complejo que resulta inalcanzable. En este trabajo se presenta como alternativa la realización de un estudio de simulación. A partir de distribuciones normales con distintos parámetros, se simulan pares de muestras con diversos tamaños, tanto iguales como diferentes, y en cada escenario planteado, se computa una estimación empírica de la potencia para ambas pruebas.

Objetivo

Comparar la potencia entre las pruebas de hipótesis no paramétricas de Kolmogorov-Smirnov (K-S) y Wald-Wolfowitz (W-W) para dos muestras independientes provenientes de poblaciones con distribución normal.

Metodología

Las dos pruebas plantean las mismas hipótesis utilizadas para distribuciones continuas:

$$H_0) F_1(X) = F_2(X)$$

$$H_1) F_1(X) \neq F_2(X)$$

donde $F_i(X)$ es la función de distribución teórica de la población i .

Ambas pruebas están muy difundidas, K-S computa la diferencia máxima absoluta entre las funciones de distribución empíricas de las dos muestras. W-W combina las observaciones de ambas muestras y computa el número de rachas.

La estadística utilizada en K-S es:

$$D_{n_1, n_2} = |S_{n_1}(X) - S_{n_2}(X)|$$

donde $S_{n_i}(X)$ es la función de distribución empírica calculada con las observaciones de la muestra n_i de la población i .

Se rechaza la hipótesis nula si:

$$D_{n_1, n_2} \geq D_{n_1, n_2, 1-\alpha}$$

siendo $D_{n_1, n_2, 1-\alpha}$ el valor crítico de la distribución de las diferencias bajo la hipótesis nula para un nivel de significación α .

La estadística utilizada en W-W es el número de rachas R . Se ordenan las observaciones de la muestra combinada de menor a mayor indicando a que muestra pertenecen. Se computa el número de rachas y se rechaza la hipótesis nula si:



$$R \leq R_{\alpha}$$

donde R_{α} es el valor crítico de la distribución del número de rachas bajo la hipótesis nula para un nivel de significación α .

Se realizó un estudio de simulación en donde a partir de la distribución normal se seleccionaron distintos valores para los parámetros de las dos poblaciones a comparar. Se determinaron diversos tamaños muestrales, tanto iguales como diferentes. En total se plantearon 1.032 escenarios y se realizaron 1.000 repeticiones de cada uno de ellos. En cada simulación se computó una estimación empírica de la potencia para ambas pruebas. Se utilizó el paquete estadístico R para realizar las simulaciones.

Resultados

Se plantearon 516 escenarios donde las distribuciones diferían solo en medias, pero no en variancias:

$$\begin{aligned}\mu_1 &= 1, \mu_2 = 1.5, 2, 2.5, 3, 3.5, 4 \\ \sigma_1 &= \sigma_2 = 1, 2\end{aligned}$$

Se consideraron tamaños muestrales iguales:

$$n_1 = n_2 = 4, 5, \dots, 12, 100$$

y distintos:

$$n_1 = 4, 5, \dots, 9 \quad n_2 = 5, 6, \dots, 12$$

Se observó que en el 89% de los 516 escenarios K-S resultó más potente que W-W, en el 4% de los casos fueron igualmente potentes y solo en el 7% de los casos W-W presentó una potencia mayor.

De manera similar, se plantearon 258 escenarios donde las distribuciones diferían solo en variancias:

$$\begin{aligned}\sigma_1 &= 1, \sigma_2 = 1.5, 2, 2.5, 3, 3.5, 4 \\ \mu_1 &= \mu_2 = 1, 2\end{aligned}$$

Nuevamente se consideraron tamaños muestrales iguales y distintos de manera similar al caso anterior. Se observó que en el 78% de los 258 escenarios W-W resultó más potente que K-S, en el 4% de los casos fueron igualmente potentes y que en el 18% de los casos K-S presentó una potencia mayor.

Por último, se plantearon 258 escenarios donde las distribuciones diferían en medias y en variancias:

$$\begin{aligned}\mu_1 &= 1, \mu_2 = 1.5, 2, 2.5, 3, 3.5, 4 \\ \sigma_1 &= 1, \sigma_2 = 1.5, 2, 2.5, 3, 3.5, 4\end{aligned}$$

También en estos escenarios se consideraron tamaños muestrales iguales y distintos. Se observó que en el 87% de los 258 escenarios K-S resultó más potente que W-W, en el 3% de los casos fueron igualmente potentes y que en el 10% de los casos W-W presentó una potencia mayor.



Conclusiones

En este estudio, el test de K-S resultó ser más potente cuando las distribuciones normales difieren solo en posición. En el caso de que las muestras provengan de distribuciones normales que difieran solo en dispersión, el test de W-W resultó ser más potente. Cuando las poblaciones difieren en localización y en dispersión, K-S presenta mayores potencias en la mayoría de los casos. En términos generales, en este estudio en particular, se puede concluir que K-S es superior en cuanto a potencia que el Test de Rachas, en el caso que las muestras provengan de distribuciones normales (71% de los escenarios simulados).

Discusión

El test de K-S parece ser más sensible a diferencias de posición que el test de rachas mientras que este parece ser más sensible cuando hay diferencias de variabilidad, pero no en medias. Cuando las poblaciones difieren en localización y en dispersión, parece que tiene más peso la sensibilidad del test de K-S a la diferencia de localización que la del test de rachas a la diferencia en variabilidad. Si la distribución de las poblaciones a comparar fuese normal, no se aplicarían estas pruebas, ya que se utilizarían las pruebas clásicas. Sin embargo, cuando se estudia la eficiencia relativa de las pruebas no paramétricas, se las compara con las pruebas clásicas bajo teoría normal. En consecuencia, este trabajo representa un primer paso para el estudio de la comparación de la potencia entre ambas pruebas. Es de interés en el futuro seguir estudiando la potencia de estos tests cuando las poblaciones no son normales, es decir, considerando poblaciones con leves o fuertes asimetrías, tales como la distribución Lognormal o Exponencial.

REFERENCIAS BIBLIOGRÁFICAS

- ARÓNICA, E. "Inferencia Estadística. Métodos de Distribución Libre" (2011)
- HOLLANDER, M., WOLFE, D. A. "Nonparametric Statistical Methods" (2013)
- KENDALL, M. C., STUART, A. "The Advanced Theory of Statistics", Vol. I, II y III (1963)
- MARASCUILO, L. A., MC SWEENEY, M. "Nonparametric and Distribution Free Methods for the Social Sciences" (1977)
- SIEGEL, S. "Estadística No Paramétrica" (1995)
- SHESKIN, D.J. "Handbook of parametric and nonparametric statistical procedures." Sheskin, D.J..Fifth ed. (2020)