

Comportamiento de la función de potencia de la prueba de Wilcoxon para una alternativa de asimetría con muestras seleccionadas de distribuciones de la familia Lambda Generalizada.

Margoth Adriana Valdivieso Miranda*, Víctor Miguel Ángel Burbano**

UNIVERSIDAD PEDAGÓGICA Y TECNOLÓGICA DE COLOMBIA, ESCUELA DE MATEMÁTICAS Y ESTADÍSTICA, TUNJA BOYACÁ

Resumen

Se presenta un análisis del comportamiento de la función de potencia de la prueba de Wilcoxon para dos muestras seleccionadas de distribuciones asimétricas pertenecientes a la Distribución Lambda Generalizada (DLG), considerando pares de muestras para las cuales el tamaño de la muestra combinada " $n + m$ " se ubica entre 20 y 50 para distintos valores en el coeficiente de asimetría de la DLG. Se calculan valores de la potencia de la prueba mediante una aproximación de la función de potencia escrita en términos del parámetro de asimetría de la DLG y del parámetro de curtosis. Los resultados numéricos permiten organizar tablas y trazar curvas de potencia pertinentes para analizar el comportamiento mencionado.

1. Introducción

Las funciones de potencia son de gran importancia en la evaluación de pruebas estadísticas, particularmente en la comparación de regiones críticas que podrían utilizarse para probar una determinada hipótesis nula contra una alternativa dada.

La potencia de las pruebas estadísticas ha sido un tema trabajado minuciosamente por los estadísticos teóricos, tradicionalmente poco tenido en cuenta por los investigadores aplicados y no muy estudiado por los metodólogos, quizá porque para su cálculo se requiere la distribución del estadístico de prueba bajo la hipótesis alternativa, la cual no siempre se puede obtener y cuando es posible hacerlo resultan expresiones matemáticas que solo están al alcance de los teóricos en estadística. Un estudio pionero sobre este tema, Cohen (1962) encontró que en la inmensa mayoría de los trabajos publicados en las revistas científicas de Psicología, la potencia de los contrastes de hipótesis era extraordinariamente baja. Sedlmeier y Gigerenzer (1989) encontraron que, que tras veinte años la situación en lo que respecta a la potencia se encontraba exactamente igual. Clark-Carter (1997) encontraron que después de treinta años el estudio de la potencia de pruebas estadísticas no se había incrementado, aunque ahora hay un pequeño porcentaje de experimentos en los cuales ya se incluye la potencia de la prueba.

En una revisión de investigaciones clínicas, Kazdin y Bass (1989), tras revisar 120 estudios publicados en revistas de investigación en psicoterapia entre 1984 y 1986, encontraron una gran despreocupación con respecto al tema de la potencia. Con

* Docente Escuela de Matemáticas y Estadística UPTC e-mail: mavaldiviesom@unal.edu.co

** Docente Escuela de Matemáticas y Estadística UPTC e-mail: victorburbano@yahoo.es

respecto a los tamaños muestrales, estos mismos autores encontraron que la mediana de los tamaños grupales en las publicaciones estudiadas estaba en 12, con un rango de 3 a 114 y que tres cuartas partes de los estudios incluían menos de 20 sujetos por grupo. En otro estudio sobre este mismo tema, Shapiro y Shapiro (1982) encontraron que los tamaños grupales incluidos en los estudios se consideraron solo en un 10% de los trabajos, se incluyeron 6 o menos sujetos por grupo de tratamiento en un 26% de los estudios y en un 36% se trabajó con 10 o 12 sujetos. Para Keppel (1991) existe un cierto consenso en el sentido de que se deben buscar potencias de alrededor de 0.8 como un criterio razonable para rechazar apropiadamente la hipótesis nula en estudios referidos a las ciencias del comportamiento, de esta manera se manejan de forma equilibrada los errores tipo I y II.

Las consideraciones anteriores han llevado a estudiar el problema de la potencia, analizando tamaños de efecto pequeño en dos y tres muestras, utilizando métodos paramétricos y no paramétricos (prueba t de Student, prueba de Welch y prueba de Wilcoxon en el caso de dos grupos), ello tanto en condiciones de cumplimiento de los supuestos subyacentes a las pruebas como bajo violación de los supuestos de normalidad y homocedasticidad. De otro lado, varios métodos de aproximación de la potencia de la prueba basada en la estadística de Wilcoxon se han propuesto, algunos de ellos se han basado en métodos bootstrap y bajo el supuesto de que la función de distribución de la que se muestrea corresponde a la distribución normal, publicados entre 1999 y 2005 por John Wiley & Sons, Ltda. Usa y otros, también la mencionada potencia ha sido calculada para muestras aleatoria independientes provenientes de distribuciones como la Uniforme (Thompson 1966), la Exponencial (Hayman y Govindarajulu 1966), la Normal (Milton 1970), logística y Cauchy (Randles and Wolfe 1979). Sin embargo, solo se ha considerado tamaños de muestra combinada inferiores a 20.

En este trabajo se amplía el espectro de distribuciones y se analiza el comportamiento de la función de potencia de la prueba de Wilcoxon para una alternativa de asimetría en el problema de dos muestras provenientes de la Distribución Lambda Generalizada (DLG), considerando pares de muestras para las cuales el tamaño de la muestra combinada “ $m + n$ ” corresponda a valores específicos mayores que 20 y menores que 50. Se fijan ciertos valores del coeficiente de asimetría de la DLG y se comparan las funciones de potencia resultantes; para ello, mediante un algoritmo codificado en lenguaje C se calculan los valores para la potencia de la prueba a medida que el parámetro involucrado en las hipótesis se aleja de la hipótesis nula.

2. Potencia de la prueba

Cuando se trabaja con hipótesis compuestas es conveniente considerar las probabilidades $\alpha(\theta)$ de cometer error tipo I para todos los valores de θ dentro del dominio especificado de la hipótesis nula y las probabilidades $\beta(\theta)$ de cometer un error de tipo II para todos los valores de θ dentro del dominio especificado de la hipótesis alternativa. Con base en los dos elementos anteriores, la función de potencia de una prueba de una hipótesis nula contra una hipótesis alternativa se puede definir de la siguiente manera [1]:

$$\pi(\theta) = \begin{cases} \alpha(\theta) & \text{para valores de } \theta \text{ asumidos bajo la hipótesis nula} \\ 1 - \beta(\theta) & \text{para valores de } \theta \text{ asumidos bajo la hipótesis alternativa} \end{cases} \quad (1)$$

Los valores de la función de potencia son las probabilidades de rechazar la hipótesis nula para diversos valores del parámetro θ . Es conveniente destacar que para valores de θ asumidos bajo la hipótesis nula, la función de potencia proporciona las probabilidades de cometer error tipo I, y para valores de θ asumidos bajo la hipótesis alternativa da la probabilidad de no cometer error tipo II (probabilidad de rechazar la hipótesis nula cuando ella es falsa). En síntesis, la potencia de una prueba mide la probabilidad de rechazar acertadamente la hipótesis de nulidad.

En términos generales, para el caso en el cual se tiene: θ un parámetro que toma valores en el conjunto de los números reales no negativos y U una estadística de prueba basada en una muestra de tamaño N de una distribución F , para probar la hipótesis:

$$H_0 : \theta = 0 \quad \text{frente a la alternativa} \quad H_1 : \theta > 0$$

La estadística U rechaza la hipótesis H_0 a favor de la hipótesis H_1 cuando $U \geq c$ donde c es el valor crítico de la prueba determinado por $P_{H_0}(U \geq c) \leq \alpha$ para un nivel de significancia α .

La función de potencia de la prueba basada en la estadística U de acuerdo con (1) se puede expresar de la siguiente forma:

$$\pi_U(\theta) = P_\theta(U \geq c) \tag{2}$$

para todo θ en el espacio de parámetros.

2.1 Modelo de muestreo para este trabajo.

Sean X_1, \dots, X_m y Y_1, \dots, Y_n dos muestras aleatorias independientes entre sí, provenientes de distribuciones $F(x)$ y $G(y) = F(y - \theta)$ continuas con f.d.p $f(x)$, la distribución F se considera asimétrica y proveniente de la familia Lambda Generalizada.

La muestra combinada de X_1, \dots, X_m y Y_1, \dots, Y_n está definida de la siguiente forma: $X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_{m+n}$ con $X_{m+j} = Y_j$ para $j = 1, 2, \dots, n$

2.2 Estadística de Wilcoxon.

Para probar hipótesis como las dadas en (2), Wilcoxon en 1945 propuso el siguiente procedimiento:

- Ordenar de menor a mayor la muestra combinada.
- Determinar los rangos correspondientes.
- Calcular la suma U de los rangos R_j de los Y_j para $j = 1, 2, \dots, n$, es decir

$$U = \sum_{j=1}^n R_j$$

A la anterior estadística se le denomina estadística de Wilcoxon.

En 1947 Mann y Whitney propusieron la siguiente expresión para la estadística U :

$$U = V + \frac{n(n+1)}{2}$$

donde

$$V = \sum_{i=1}^m \sum_{j=1}^n T_{ij} \text{ con } T_{ij} = \begin{cases} 1 & \text{si } Y_j > X_i \\ 0 & \text{en otro caso} \end{cases}$$

La estadística V se denomina estadística de Mann-Whitney, la cual se puede escribir de la siguiente forma:

$$V = U - \frac{n(n+1)}{2}$$

Bajo la hipótesis nula $H_0 : \theta = 0$, la esperanza y la varianza para la estadística U son:

$$E(U) = \frac{1}{2}n[(m+n)+1] \quad \text{Var}(U) = \frac{1}{12}mn[(m+n)+1]$$

Bajo la hipótesis alternativa $H_1 : \theta > 0$, la esperanza y la varianza para la estadística U son:

$$\begin{aligned} E(U) &= mnp_1 + \frac{1}{2}n(n+1) \\ \text{Var}(U) &= mn(p_1 - p_1^2) + mn(n-1)(p_2 - p_1^2) + mn(m-1)(p_3 - p_1^2) \\ p_1 &= P(Y > X) = \int_{-\infty}^{\infty} F(y)g(y)dy = \int_{-\infty}^{\infty} (1-G(x))f(x)dx \\ p_2 &= P(Y_1 > X_1, Y_2 > X_1) = \int_{-\infty}^{\infty} (1-G(x))^2 f(x)dx \\ p_3 &= P(Y_1 > X_1, Y_1 > X_2) = \int_{-\infty}^{\infty} (F(y))^2 g(y)dy \end{aligned}$$

La estadística siguiente tiene una distribución normal estándar [2]:

$$\frac{V - E(V)}{\sqrt{\text{Var}(V)}} \Rightarrow \frac{U - E(U)}{\sqrt{\text{Var}(U)}} \rightarrow N(0,1)$$

Con los elementos anteriores, la función de potencia de la prueba basada en la estadística U se puede aproximar por la distribución normal [3] y la expresión (2) puede escribirse de la siguiente forma:

$$\pi_U(\theta) = P_\theta(U \geq c) \cong \Phi \left(\sqrt{\frac{12mn}{N+1}} f^*(0)\theta - u_\alpha \right) \quad (3)$$

donde α es el nivel de significación, c es tal que $P_{H_0}(U \geq c) \leq \alpha$, $\Phi(u_\alpha) = 1 - \alpha$ y

$$f^*(0) = \int_{-\infty}^{\infty} f^2(x)dx.$$

3. Expresión para la función de potencia en términos de la DLG

La Distribución Lambda Generalizada se define de la siguiente manera [4]:

$$F^{-1}(y) = F^{-1}(y, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2} \quad (4)$$

con $\lambda_2 \neq 0$, $0 \leq y \leq 1$. En esta definición λ_1 es el parámetro de localización, λ_2 es el parámetro de escala, λ_3 determina el sesgo (coeficiente de asimetría) y λ_4 determina la curtosis. Utilizando la expresión (4) se puede obtener la función de densidad de la distribución Lambda Generalizada:

$$f(x) = \frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}} \text{ en } x = F^{-1}(y).$$

Asignando valores adecuados a los parámetros de la DLG, se obtienen las densidades de las distribuciones de probabilidad más usuales tales como la distribución normal ($\lambda_1 = 0$, $\lambda_2 = 0.1975$, $\lambda_3 = 0.1349$, $\lambda_4 = 0.1349$), la distribución uniforme en $[-1,1]$ ($\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 2$, $\lambda_4 = 2$), con los parámetros que se indican a continuación: $\lambda_1 = 0$, $\lambda_2 = -0.0003637$, $\lambda_3 = -0.000363$, $\lambda_4 = -0.000363$, se tiene la distribución logística, etc.

Para expresar (3) en términos de los parámetros de la distribución Lambda Generalizada, se calcula:

$$f^*(0) = \int_{-\infty}^{\infty} f^2(x) dx = \int_{-\infty}^{\infty} f(x) f(x) d(x) = \int_0^1 \frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}} dy$$

reemplazando en (3) resulta:

$$\pi_U(\theta) = P_\theta(U \geq c) \cong \Phi \left(\frac{\sqrt{12mn}}{\sqrt{(m+n)+1}} \left(\int_0^1 \frac{\lambda_2}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}} dy \right) \theta - u_\alpha \right) \quad (5)$$

La expresión (5) es una aproximación de la potencia de la prueba basada en la estadística U para muestras provenientes de distribuciones seleccionadas de la familia Lambda Generalizada.

4. Función de potencia en términos de λ_3 y λ_4

Si F es la distribución Lambda Generalizada, la mediana corresponde al valor θ que hace que se cumpla las siguientes igualdades:

$$F(\theta) = \frac{1}{2} \quad \Rightarrow \quad F^{-1}\left(\frac{1}{2}\right) = \lambda_1 + \frac{\left(\frac{1}{2}\right)^{\lambda_3} - \left(\frac{1}{2}\right)^{\lambda_4}}{\lambda_2} = \theta \quad (6)$$

La hipótesis

$$H_0 : \theta = 0 \quad \text{versus} \quad H_1 : \theta > 0$$

Se puede expresar de la siguiente forma:

$$H_0 : \lambda_1 + \frac{(\frac{1}{2})^{\lambda_3} - (\frac{1}{2})^{\lambda_4}}{\lambda_2} = 0 \quad \text{versus} \quad H_1 : \lambda_1 + \frac{(\frac{1}{2})^{\lambda_3} - (\frac{1}{2})^{\lambda_4}}{\lambda_2} > 0$$

Si $\lambda_1 = 0$ las hipótesis anteriores pueden ser escritas así:

$$i) \quad H_0 : \lambda_3 = \lambda_4 \quad \text{versus} \quad H_1 : (\lambda_3 < \lambda_4) \cap (\lambda_2 > 0)$$

$$ii) \quad H_0 : \lambda_3 = \lambda_4 \quad \text{versus} \quad H_1 : (\lambda_3 > \lambda_4) \cap (\lambda_2 < 0)$$

A la hipótesis $H_0 : \lambda_3 = \lambda_4$ se denominará hipótesis de simetría y a la hipótesis H_1 se le llamará hipótesis de asimetría tanto en *i*) como en *ii*).

Si $\lambda_1 = 0$ la expresión (6) hace posible que (5) tome la forma:

$$\pi_U(\theta) = P_\theta(U \geq c) \cong \Phi \left(\sqrt{\frac{12mn}{(m+n)+1}} \left(\int_0^1 \frac{(\frac{1}{2})^{\lambda_3} - (\frac{1}{2})^{\lambda_4}}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}} dy \right) - u_\alpha \right) \quad (7)$$

La expresión (7) corresponde a una de las expresiones de la función de potencia de la prueba basada en la estadística de Wilcoxon para la alternativa de asimetría en términos de los parámetros λ_3 y λ_4 , con $(\lambda_3 \leq \lambda_4)$ y $(\lambda_2 > 0)$ ó $(\lambda_3 \geq \lambda_4)$ y $(\lambda_2 < 0)$.

También (7) permite verificar que bajo la hipótesis de simetría $H_0 : \lambda_3 = \lambda_4$ el valor de la función de potencia corresponde al nivel de significancia α .

Con el objeto de comparar valores de la función de potencia para distintas distribuciones, es conveniente expresar (7) en términos de $\frac{\theta}{\sigma}$, esto se puede lograr haciendo uso del segundo momento para una variable aleatoria X con distribución Lambda Generalizada [4], dados por:

$$\alpha_2 = \sigma^2 = E[(X - \mu)^2] = \frac{B - A^2}{\lambda_2^2}$$

$$A = \frac{1}{1 + \lambda_3} - \frac{1}{1 + \lambda_4}$$

$$B = \frac{1}{1 + 2\lambda_3} + \frac{1}{1 + 2\lambda_4} - 2\text{Beta}(1 + \lambda_3, 1 + \lambda_4)$$

Y después de realizar algunas sustituciones, cálculos algebraicos y usar el resultado siguiente:

$$\text{Beta}(1 + \lambda_3, 1 + \lambda_4) = \int_0^1 t^{\lambda_3} (1-t)^{\lambda_4} dt$$

Si $\lambda_1 = 0$, bajo la hipótesis alternativa se tiene que,

$$\pi_U(\theta) = P_\theta(U \geq c') \cong \Phi \left(\sqrt{\frac{12mn}{(m+n)+1}} \left(\int_0^1 \frac{E}{\lambda_3 y^{\lambda_3-1} + \lambda_4 (1-y)^{\lambda_4-1}} dy \right) \frac{\theta}{\sigma} - u_\alpha \right) \quad (8)$$

Donde

$$\lambda_2 = \frac{E}{\sigma} \quad \text{con} \quad E = \sqrt{\frac{1}{1+2\lambda_3} + \frac{1}{1+2\lambda_4} - 2 \int_0^1 t^{\lambda_3} (1-t)^{\lambda_4} dt - \left(\frac{1}{1+\lambda_3} - \frac{1}{1+\lambda_4} \right)^2}$$

La expresión (8) corresponde a la función de potencia de la prueba basada en la estadística de Wilcoxon para la alternativa de asimetría en términos de los parámetros λ_3 y λ_4 .

5. Resultados numéricos

La integral involucrada en (8) no se puede resolver usando métodos analíticos y fue necesario usar métodos numéricos. Se desarrolló un programa en Lenguaje C que contiene básicamente la regla de Simpson para calcular integrales. Con la ayuda del programa se calcularon los valores de la función de potencia para valores de $\frac{\theta}{\sigma}$ correspondientes a 0; 0.3; 0.6; 0.9; 1.2 y 1.5. El tamaño de la muestra combinada “ $m+n$ ” se fijó para valores entre 20 y 50, cuando “ $m+n$ ” corresponde a un número par se tomó $m=n$ y cuando “ $m+n$ ” corresponde a un número impar se trabajó con $n=m+1$. Como se indica en la tabla 8. Así mismo, se consideraron valores específicos de los parámetros λ_3 y λ_4 que definen distribuciones asimétricas provenientes de la familia Lambda Generalizada. Se utilizó un nivel de significancia $\alpha=0.05$. Para la interpretación y análisis de resultados numéricos se han organizado tablas resumen cuyos valores se tomaron de tablas completas similares a la tabla 8 (Anexo) que se obtuvieron en la investigación, y con ellas se han graficado algunas curvas de la función de potencia de la prueba de Wilcoxon para muestras de las distribuciones asimétricas.

5.1 Distribución asimétrica y platicúrtica obtenida de la Distribución Lambda Generalizada (DLG). Coeficiente de asimetría $\alpha_3=0.3$. Coeficiente de curtosis $\alpha_4=2$. Parámetros de la DLG: $\lambda_1=0$; $\lambda_2=0.2660$; $\lambda_3=-0.00001$; $\lambda_4=0.7019$

$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
20	0,05	0,162983	0,374389	0,63383	0,842258	0,952213
30	0,05	0,204264	0,496299	0,790441	0,947897	0,992776
40	0,05	0,243093	0,599148	0,884783	0,984101	0,998964
50	0,05	0,280159	0,6844	0,938422	0,995327	0,999812

Tabla 1. $m=n$; $\alpha=0,05$

5.2 Distribución asimétrica y mesocúrtica obtenida de la Distribución Lambda Generalizada (DLG). Coeficiente de asimetría $\alpha_3=0.3$. Coeficiente de curtosis $\alpha_4=3$. Parámetros de la DLG: $\lambda_1=0$; $\lambda_2=0.1991$; $\lambda_3=0.09255$; $\lambda_4=0.1859$

$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
20	0,05	0,158993	0,361937	0,614953	0,825791	0,94336
30	0,05	0,198415	0,479841	0,772383	0,938619	0,990256
40	0,05	0,235635	0,580603	0,870211	0,979982	0,998599
50	0,05	0,271294	0,665184	0,928129	0,993656	0,999878

Tabla 2. $m = n$; $\alpha = 0,05$

5.3 Distribución asimétrica y leptocúrtica obtenida de la Distribución Lambda Generalizada (DLG). Coeficiente de asimetría $\alpha_3 = 0.3$. Coeficiente de curtosis $\alpha_4 = 4$. Parámetros de la DLG: $\lambda_1 = 0$; $\lambda_2 = 0.03891$; $\lambda_3 = 0.01902$; $\lambda_4 = 0.02549$

$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
20	0,05	0,16898	0,392521	0,660111	0,864141	0,962858
30	0,05	0,212534	0,519532	0,814816	0,959296	0,995238
40	0,05	0,253836	0,625119	0,90323	0,988737	0,9995
50	0,05	0,293049	0,710895	0,951198	0,997068	0,999977

Tabla 3. $m = n$; $\alpha = 0,05$

5.4 Distribución asimétrica y mesocúrtica obtenida de la Distribución Lambda Generalizada (DLG). Coeficiente de asimetría $\alpha_3 = 0.8$. Coeficiente de curtosis $\alpha_4 = 3$. Parámetros de la DLG: $\lambda_1 = 0$; $\lambda_2 = 0.1995$; $\lambda_3 = 0.006849$; $\lambda_4 = 0.3356$

$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
20	0,05	0,180979	0,429317	0,710589	0,901142	0,978309
30	0,05	0,229812	0,565811	0,857822	0,975873	0,998
40	0,05	0,275828	0,675171	0,933601	0,994616	0,999829
50	0,05	0,319813	0,760408	0,970121	0,998867	0,999949

Tabla 4. $m = n$; $\alpha = 0,05$

5.5 Distribución asimétrica y leptocúrtica obtenida de la Distribución Lambda Generalizada (DLG). Coeficiente de asimetría $\alpha_3 = 0.8$. Coeficiente de curtosis $\alpha_4 = 4$. Parámetros de la DLG: $\lambda_1 = 0$; $\lambda_2 = 0.1042$; $\lambda_3 = 0.03035$; $\lambda_4 = 0.09887$

$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
20	0,05	0,17211	0,402042	0,673772	0,874524	0,967618
30	0,05	0,21702	0,53181	0,826949	0,984327	0,996199
40	0,05	0,2319	0,571123	0,862405	0,977447	0,998234
50	0,05	0,30004	0,724429	0,956967	0,997743	0,999949

Tabla 5. $m = n$; $\alpha = 0,05$

5.6 Distribución asimétrica y leptocúrtica obtenida de la Distribución Lambda Generalizada (DLG). Coeficiente de asimetría $\alpha_3 = 1.10$. Coeficiente de curtosis $\alpha_4 = 4$. Parámetros de la DLG: $\lambda_1 = 0$; $\lambda_2 = 0.1458$; $\lambda_3 = 0.00604$; $\lambda_4 = 0.2013$

$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
20	0,05	0,198512	0,48011	0,772552	0,938841	0,990397
30	0,05	0,254559	0,627113	0,904676	0,989197	0,999543
40	0,05	0,307336	0,738192	0,962318	0,998214	0,999951
50	0,05	0,357369	0,819456	0,985828	0,999756	0,999971

Tabla 6. $m = n$; $\alpha = 0,05$

5.7 Distribución asimétrica y leptocúrtica obtenida de la Distribución Lambda Generalizada (DLG). Coeficiente de asimetría $\alpha_3 = 1.80$. Coeficiente de curtosis $\alpha_4 = 8$. Parámetros DLG: $\lambda_1 = 0$; $\lambda_2 = 0.00291$; $\lambda_3 = 0.0001925$; $\lambda_4 = 0.002798$

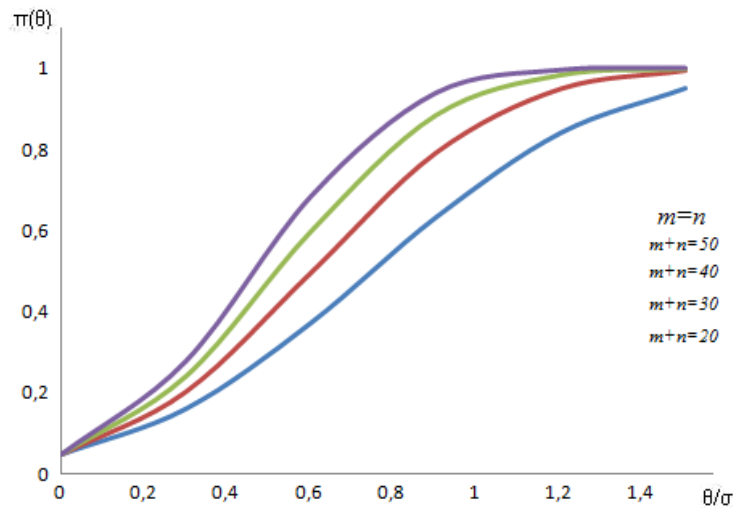
$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
20	0,05	0,236678	0,583191	0,872308	0,980582	0,998592
30	0,05	0,308732	0,740552	0,963159	0,998382	0,999971
40	0,05	0,485844	0,942103	0,999211	0,99995	0,99998
50	0,05	0,437627	0,908212	0,997664	0,999961	0,999999

Tabla 7. $m = n$; $\alpha = 0,05$

6. Análisis de resultados

En la tabla 1 y la gráfica 1 se observa que a medida que se incrementa el tamaño de la muestra, la función de potencia crece más rápidamente; lo cual indica que la potencia de la prueba aumenta a medida que se incrementa el tamaño de la muestra. Así mismo, se puede notar que la potencia de la prueba aumenta a medida que θ/σ se aleja de cero (de la hipótesis nula). Una situación similar se presenta para los casos que se indican en las tablas: 2, 3, 4, 5, 6 y 7.

Al comparar los valores de las tablas 1, 2 y 3 manteniendo fijos el tamaño de la muestra combinada y el coeficiente de asimetría en $\alpha_3 = 0.3$ y haciendo variar el coeficiente de curtosis (apuntamiento $\alpha_4 = 2, 3, 4$) se puede interpretar que entre más apuntamiento tenga la distribución asimétrica muestreada de una DLG, la prueba basada en Wilcoxon resulta más potente.



Si se mantienen fijos el tamaño de la muestra combinada y el coeficiente de curtosis (apuntamiento $\alpha_4 = 3$) y se hace variar el coeficiente de asimetría ($\alpha_3 = 0.3, 0.8$. Ver tablas 2 y 4) se puede concluir que entre más asimétrica sea la distribución de proveniente de la familia DLG, la potencia de la prueba basada en Wilcoxon es mayor. Similar análisis se puede hacer si se mantiene fijo $\alpha_4 = 4$ y se hace variar el coeficiente de asimetría pasando de $\alpha_3 = 0.8$. a 1.10. Ver tablas 5 y 6.

En la tabla 7 se presentan los valores de la función de potencia consideran un coeficiente de asimetría en $\alpha_3 = 1.80$ y un coeficiente de curtosis de $\alpha_4 = 8$, se puede afirmar que entre más asimétrica y mayor apuntamiento tenga la distribución seleccionada de la familia Lambda Generalizada, más potente es la prueba basada en la estadística de Wilcoxon. La anterior al comparar los valores de la tabla 7 con los valores de potencia que se presentan correspondientemente en las tablas: 1, 2, 3, 4, 5, y 6.

A manera de conclusión se puede decir que la prueba basada en la estadística de Wilcoxon es más potente cuando se muestrea de distribuciones asimétricas y muy apuntadas, las cuales se han seleccionado de la familia Lambda Generalizada. Además, la potencia de la prueba aumenta a medida que el tamaño de la muestra combinada $m + n$ aumenta, esto hace pensar que la prueba es consistente.

7. Referencias

- [1] FREUND, J. E. (2000). Estadística Matemática con aplicaciones. Pearson Educación. México.
- [2] HETTMANSPERGER, T.P. (1984). Statistical Inference based on Ranks. John.Wiley & Sons. USA.
- [3] LEHMAN, E. L. (1975). Nonparametric Statistical Methods Based on Ranks. McGraw-Hill International Book Company. New York.
- [4] KARIAN, Z.A., DUDEWICZ, E.J. (2000). Fitting Statistical Distributions: The Generalized Lambda Distribution and Generalized Bootstrap Methods. CRC Press. Boca Ratón, FL.

Anexo

$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
20	0,05	0,162983	0,374389	0,63383	0,842258	0,952213
22	0,05	0,171451	0,400416	0,671421	0,872792	0,966714
24	0,05	0,179805	0,425475	0,705525	0,897768	0,977133
26	0,05	0,187913	0,449882	0,736711	0,918067	0,984281
28	0,05	0,196181	0,473523	0,764976	0,934647	0,989324
30	0,05	0,204264	0,496299	0,790441	0,947897	0,992776
32	0,05	0,212056	0,518289	0,813651	0,958775	0,995115
34	0,05	0,219896	0,539537	0,834412	0,967354	0,996698
36	0,05	0,227709	0,560188	0,853005	0,97418	0,997716
38	0,05	0,235353	0,580071	0,869855	0,979716	0,998432
40	0,05	0,243093	0,599148	0,884783	0,984101	0,998964
42	0,05	0,250592	0,617657	0,898185	0,987484	0,999313
44	0,05	0,258012	0,635323	0,910126	0,990224	0,999556
46	0,05	0,265522	0,652368	0,920622	0,992365	0,999779
48	0,05	0,273042	0,668717	0,930044	0,994207	0,999847
50	0,05	0,280159	0,6844	0,938422	0,995327	0,999812

Tabla 8. $m = n$; $\alpha = 0,05$

$m+n$	$\theta/\sigma=0$	$\theta/\sigma=0.3$	$\theta/\sigma=0.6$	$\theta/\sigma=0.9$	$\theta/\sigma=1.2$	$\theta/\sigma=1.5$
21	0,05	0,167088	0,386855	0,652254	0,857634	0,959875
23	0,05	0,175567	0,412482	0,688125	0,885316	0,972285
25	0,05	0,183772	0,437232	0,720879	0,908142	0,980826
29	0,05	0,200096	0,484566	0,777664	0,94135	0,991112
31	0,05	0,207936	0,507019	0,801969	0,953386	0,993842
33	0,05	0,215872	0,528698	0,823999	0,963198	0,995967
35	0,05	0,223711	0,549678	0,843651	0,971022	0,99736
37	0,05	0,231431	0,570028	0,861408	0,977133	0,998164
39	0,05	0,239232	0,589477	0,877268	0,982073	0,998754
41	0,05	0,246778	0,608362	0,891411	0,985948	0,999091
43	0,05	0,254172	0,626309	0,904111	0,988875	0,999404
45	0,05	0,261738	0,6437	0,915412	0,991246	0,999619
47	0,05	0,269165	0,660412	0,925436	0,993187	0,999681
49	0,05	0,276559	0,676553	0,934392	0,994781	0,999858
50	0,05	0,280159	0,6844	0,938422	0,995327	0,999812

Tabla 8. (Continuación) $n = m + 1$; $\alpha = 0,05$