



# COLEGIO DE POSTGRADUADOS

---

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN  
EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA  
ESTADÍSTICA

**Una prueba robusta para detectar  
heterogeneidad de varianzas en un diseño  
experimental completamente al azar**

**Annel Hurtado Jaramillo**

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA  
OBTENER EL GRADO DE:

**MAESTRA EN CIENCIAS**

MONTECILLO, TEXCOCO, EDO. DE MÉXICO

2013

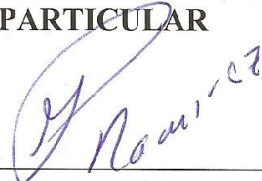
La presente tesis titulada: **Una prueba robusta para detectar heterogeneidad de varianzas en un diseño experimental completamente al azar**, realizada por la alumna: **Annel Hurtado Jaramillo**, bajo la dirección del Consejo Particular indicado ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

## MAESTRA EN CIENCIAS

### SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA ESTADÍSTICA

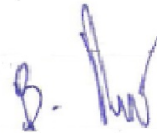
#### CONSEJO PARTICULAR

CONSEJERO



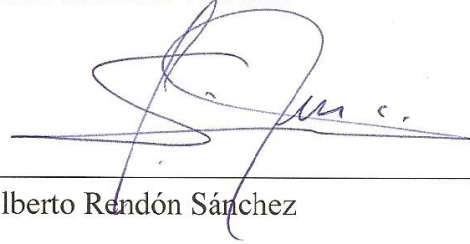
Dr. Gustavo Ramírez Valverde

ASESOR



Dr. Benito Ramírez Valverde

ASESOR



Dr. Gilberto Rendón Sánchez

# Una prueba robusta para detectar heterogeneidad de varianzas en un diseño experimental completamente al azar

Annel Hurtado Jaramillo

Colegio de Postgraduados, 2013

Uno de los supuestos más importantes del ANOVA es que las varianzas de las poblaciones estudiadas son iguales. Una prueba para verificar homogeneidad de varianzas es la prueba de Levene la cual es relativamente sensible a desviaciones de normalidad, se han propuesto algunas modificaciones logrando pruebas más robustas, sin embargo, no han sido del todo satisfactorias. Por lo que, este trabajo tiene como objetivo proponer una prueba robusta para heterogeneidad de varianzas usando el principio de transformación a rangos sobre el valor absoluto de los residuos. Mediante simulación Monte Carlo se comparó la potencia y el tamaño de la prueba propuesta con distintas variantes de la prueba de Levene. Los resultados de la simulación muestran que la prueba de Levene basada en medianas es muy conservadora en muestras pequeñas ( $n=5$ ) y que la prueba de Nordstokke-Zumbo (2010) tiene tamaños de prueba que tienden a incrementar conforme aumenta el tamaño de efecto. La prueba propuesta resultó ser más potente que la prueba de Levene basada en medianas cuando la distribución de los errores corresponde a distribuciones con colas pesadas .

**Palabras clave:** no normalidad, potencia de una prueba, prueba de Levene, robustez, tamaños de prueba, transformación a rangos.

# Robust test to detect heterogeneity of variances in a completely randomized design

Annel Hurtado Jaramillo

Colegio de Postgraduados, 2013

One of the most important assumptions of ANOVA is that the variances of the populations studied are equal. Levene's test for homogeneity of variances is relatively sensitive to deviations from normality, some modifications have been proposed achieving more robust evidence, however, were not entirely satisfactory. This paper suggests a robust test for homogeneity of variances which is based on the principle of rank transformation on the absolute value of the residuals. Monte Carlo simulation was performed to compare the power and size of the proposed test with the different variants of the Levene's test. Simulation results show that the Levene's test based on medians is quite conservative for small sample sizes ( $n = 5$ ), Nordstokke- Zumbo's (2010) test has sizes which tend to increase with increasing effect size. The proposed test was more powerful than the Levene's test based on medians from when the error distribution is double exponential and student's  $t$  distribution.

**Key words:** Levene's test, non-normality, power of test, rank transformation, robustness, test's size.

## AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado para la realización de mis estudios de maestría.

Al Colegio de Postgraduados por haberme brindado la oportunidad de seguir mi formación académica en sus aulas.

Al Dr. Gustavo Ramírez Valverde mi más sincero agradecimiento por su apoyo incondicional, amabilidad y su valiosa ayuda en la realización de este trabajo.

A los integrantes de mi Consejo Particular, Dr. Benito Ramírez Valverde y Dr. Gilberto Rendón Sánchez, por su colaboración en la revisión de este escrito.

A mis profesores del CP.

A los integrantes del personal administrativo del ISEI por su amabilidad y apoyo que siempre me brindaron, en especial a Isabel por la ayuda, muchas gracias Isa.

A mis compañeros y amigos del CP, por todos los momentos compartidos.

## DEDICATORIA

A mis **padres**:  
Magdalena y Agustín

A mis **hermanos**:  
Hoguer, Edith e Ibeth.

A mi cuñada Maribel.

A dos personitas muy especiales en mi vida:  
Eder e Isabella.

A Erick por acompañarme y siempre brindarme su apoyo.

# Índice

<b>1. Introducción</b>	<b>1</b>
<b>2. Pruebas de homogeneidad de varianzas</b>	<b>3</b>
2.1. Prueba de Bartlett (B) . . . . .	4
2.2. Prueba F-Max de Hartley . . . . .	4
2.3. Prueba de Levene (L) . . . . .	5
2.4. Modificaciones a la Prueba de Levene . . . . .	6
2.4.1. Prueba Levene modificada por Brown y Forshyte . . . . .	6
2.4.2. Modificaciones para tamaños de muestras desbalanceados . . . . .	7
2.4.3. Versión no paramétrica (NP) de la prueba de Levene . . . . .	7
<b>3. Método propuesto</b>	<b>9</b>
3.1. Prueba Levene Transformada . . . . .	9
3.2. Ejemplo de aplicación . . . . .	10
<b>4. Evaluación del método propuesto</b>	<b>15</b>

# Índice

---

4.1. Factores estudiados en la simulación . . . . .	15
4.1.1. Pruebas de homogeneidad de varianzas . . . . .	16
4.1.2. Función de distribución de los errores . . . . .	16
4.1.3. Tamaño de muestra . . . . .	16
4.1.4. Tamaño de efecto . . . . .	16
4.2. Potencia y Tamaño de la prueba . . . . .	17
4.2.1. Potencia de la prueba . . . . .	17
4.2.2. Tamaño de la prueba . . . . .	18
4.3. Determinación de la robustez . . . . .	18
<b>5. Resultados y discusión</b>	<b>20</b>
5.1. Tamaño de la prueba . . . . .	20
5.2. Potencia de la prueba . . . . .	26
<b>6. Conclusiones</b>	<b>31</b>
<b>Anexos</b>	<b>33</b>
<b>A. Tamaños de prueba</b>	<b>34</b>
<b>B. Potencia de prueba</b>	<b>38</b>
<b>Referencias</b>	<b>38</b>



# Índice de tablas

3.1. Datos del experimento de aprendizaje de pares asociados. . . . .	10
3.2. Valores de $z_{ij} =  y_{ij} - \bar{y}_i $ . . . . .	11
3.3. Valores de la transformación a rangos ( $r_{ij}$ ) de los valores $z_{ij}$ . . . . .	11
3.4. Valor absoluto de las desviaciones de las observaciones respecto a la mediana de cada tratamiento . . . . .	13
3.5. Transformación a rangos de las observaciones originales ( $y_{ij}$ ) . . . . .	14
3.6. Resumen de resultados . . . . .	14
A.1. Tamaño estimado de prueba, n=4 . . . . .	34
A.2. Tamaño estimado de prueba, n=5 . . . . .	35
A.3. Tamaño estimado de prueba, n=10 . . . . .	36
A.4. Tamaño estimado de prueba, n=20 . . . . .	37
B.1. Potencia estimada, n=4 . . . . .	38
B.2. Potencia estimada, n=5 . . . . .	39
B.3. Potencia estimada, n=10 . . . . .	40

B.4. Potencia estimada,  $n=20$  . . . . . 41

# Índice de figuras

5.1. Tamaño de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de muestra $n=4$ . . . . .	21
5.2. Tamaño de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de muestra $n=5$ . . . . .	23
5.3. Tamaño de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de muestra $n=10$ . . . . .	24
5.4. Tamaño de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de muestra $n=20$ . . . . .	25
5.5. Potencia de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de prueba $n=4$ . . . . .	27
5.6. Potencia de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de prueba $n=5$ . . . . .	28
5.7. Potencia de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de prueba $n=10$ . . . . .	29

5.8. Potencia de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de prueba  $n=20$ . . . . . 30

# Capítulo 1

## Introducción

Uno de los supuestos más importantes del ANOVA es que las varianzas de las poblaciones estudiadas son iguales. Al no cumplirse este supuesto se altera seriamente el tamaño y la potencia de la prueba F asociada al análisis de varianza cuando se compara la media de dos o más poblaciones. Para la verificación del supuesto de homogeneidad de varianzas existen varias pruebas, [Conover \*et al.\* \(1981\)](#) menciona la existencia de 56 pruebas, la mayoría de las cuales son variaciones de aquellas más usadas. Entre estas pruebas destacan: la prueba de [Bartlett \(1937\)](#) la cual depende crucialmente del supuesto de normalidad en los datos, la F de [Hartley \(1950\)](#), la cual es muy sensible a alejamientos del supuesto de normalidad y requiere que los tamaños de prueba sean iguales y la prueba de [Levene \(1960\)](#) una prueba relativamente menos sensible a la no normalidad de los datos.

La prueba de Levene se basa en un análisis de varianza de una vía sobre  $z_{ij} = |y_{ij} - \bar{y}_i|$ , donde  $z_{ij}$  representa el valor absoluto de las desviaciones de los valores muestrales respecto a la media. Esta prueba, a lo largo del tiempo, ha tenido diversas modificaciones que buscan mejorar su robustez, entre estas se pueden encontrar investigaciones de [Brown y Forsythe \(1974\)](#), [O'Brien \(1978\)](#), [Carroll y Schneider \(1985\)](#), [Yitnosumarto y O'Neill \(1986\)](#), [Loh \(1987\)](#), y de manera más reciente, [Nordstokke y Zumbo \(2010\)](#) entre otros. Las modificaciones a la prueba Levene se han vuelto muy populares y se utilizan en una amplia variedad de aplicaciones, por ejemplo, en datos clínicos, clima, ciencia animal, psicología, educación ([Gastwirth \*et al.\* \(2009\)](#)).

Una de las primeras modificaciones que se le hizo a la prueba de Levene fue la de

## 1. Introducción

---

Brown y Forsythe (1974) quienes proponen usar la mediana y la media truncada al 10 % como estadístico de centralidad en lugar de la media en la prueba original. Loh (1987), por su parte, examina la eficacia de la aplicación del método de corrección de grados de libertad (Satterthwaite) y la potencia de la prueba basada en transformaciones del test de Levene usando la mediana de los grupos en lugar de la media. Loh considera que el método de Satterthwaite puede mejorar la robustez de la prueba de Levene para muestras pequeñas, pero las transformaciones de la prueba incrementan el nivel de significancia. Yitnosumarto y O'Neill (1986) dan otro método para la modificación de los grados de libertad. Por su parte, Nordstokke y Zumbo (2010), mencionan que la prueba de Levene, basada en medias, es sensible a distribuciones sesgadas por lo que proponen una versión no paramétrica de la prueba original la cual busca ser más robusta para poblaciones sesgadas, sin embargo, en su estudio de simulación no toman en cuenta el tamaño del efecto en el modelo y solo manejan la distribución chi-cuadrada.

A pesar de que las modificaciones a la prueba Levene buscan ser más robustas no han sido totalmente satisfactorias, por lo que, en este trabajo se tiene como objetivo general proponer una prueba de homogeneidad de varianzas que sea robusta a la falta de normalidad basada en la transformación a rangos (Conover y Iman (1981)) sobre el valor absoluto de los residuos y como objetivos específicos: 1) desarrollar una prueba robusta a la falta de normalidad y 2) comparar la potencia y el tamaño de la prueba propuesta contra las pruebas sugeridas en la literatura.

El documento se encuentra dividido en 5 capítulos, siendo el primero esta introducción. En el capítulo 2, se describen algunas de las pruebas de homogeneidad de varianzas más utilizadas así como algunas de las modificaciones que ha sufrido la prueba de Levene; en el capítulo 3 se define la prueba propuesta y se da un ejemplo de aplicación, mientras que en el 4 se describe el método de simulación usado para evaluar el comportamiento de la prueba propuesta en cuanto a su potencia y tamaño de prueba; el capítulo 5 muestra los resultados y la discusión de comparar la prueba propuesta con la prueba Bartlett, la prueba Levene original y dos de sus variantes (Brown y Forsythe (1974) y Nordstokke y Zumbo (2010)), por último, en el capítulo 6 se tienen las conclusiones del trabajo.

# Capítulo 2

## Pruebas de homogeneidad de varianzas

Si  $y_{ij}$  denota el  $j$ -ésimo sujeto del  $i$ -ésimo tratamiento (población), entonces el modelo ANOVA de una vía puede escribirse como

$$y_{ij} = \mu + \tau_i + e_{ij}, \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, n_i \quad (2.1)$$

donde  $\mu$  es la media total para los  $k$  tratamientos,  $\tau_i$  es el efecto del  $i$ -ésimo tratamiento y  $e_{ij}$  representa el término de error el cual se asume independiente y normalmente distribuido con

$$E(e_{ij}) = 0, \quad Var(e_{ij}) = \sigma_i^2. \quad (2.2)$$

Estamos interesados en probar la hipótesis nula de homogeneidad de varianzas. Esto es,

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 \quad \text{vs} \quad H_1 : \sigma_i^2 \neq \sigma_j^2 \quad \text{para algún } i, j = 1, 2, \dots, k \text{ e } i \neq j. \quad (2.3)$$

donde  $\sigma_i^2$  es la varianza para el  $i$ -ésimo tratamiento.

## 2.1. Prueba de Bartlett (B)

---

Para probar (2.3) existen varias pruebas en la literatura, tan solo [Conover et al. \(1981\)](#) menciona la existencia de 56 pruebas para probar homogeneidad de varianzas, la mayoría de las cuales son variaciones de las pruebas más usadas. Algunas de estas pruebas se describen en las siguientes secciones.

### 2.1. Prueba de Bartlett (B)

Propuesta por [Bartlett \(1937\)](#), es una modificación de la prueba de Neyman y Pearson para “corregir el sesgo”, tiene la ventaja de no requerir que  $n_i$  sea igual, esta prueba es probablemente la más usada ([Conover y Iman \(1981\)](#)). El estadístico de la prueba Bartlett es ([Milliken y Johnson \(2002\)](#))

$$U = \frac{1}{C} \left[ \nu \log_e (\hat{\sigma}^2) - \sum_{i=1}^t \log_e (\hat{\sigma}_i^2) \right] \quad (2.4)$$

donde  $\nu = n_i - 1$ ,  $\nu = \sum_{i=1}^k \nu_i$ ,  $\hat{\sigma}^2 = \sum_{i=1}^k \hat{\nu}_i^2 / \nu$

$$\text{y } C = 1 + \frac{1}{3(k-1)} \left[ \sum_{i=1}^k \frac{1}{\nu_i} - \frac{1}{\nu} \right]$$

La hipótesis de igualdad de varianzas se rechaza si  $U > \chi_{\alpha, k-1}^2$ . Una de las desventajas de esta prueba es que es muy sensible a desviaciones de normalidad.

### 2.2. Prueba F-Max de Hartley

Esta prueba requiere que todas las muestras sean del mismo tamaño. La prueba se basa en el siguiente estadístico

$$F_{max} = \frac{\max_i \{\hat{\sigma}_i^2\}}{\min_i \{\hat{\sigma}_i^2\}} \quad (2.5)$$

donde  $i = 1, \dots, k$ , con  $k$  igual al número de muestras.



### 2.3. Prueba de Levene (L)

---

La hipótesis nula,  $H_0$  se rechaza si  $F_{max} > F_{\alpha, \nu, k}$ , donde  $\nu = n - 1$ . Si los tamaños de muestra no son iguales entonces hay dos versiones sobre el cálculo de  $\nu$ :

1. Se puede obtener una prueba “liberal” haciendo  $\nu = \max(n_i) - 1$
2. Si los tamaños de muestra no son tan desiguales, se puede obtener una prueba conservadora usando  $\nu = \min(n_i) - 1$

Esta prueba es muy sensible a alejamientos del supuesto de normalidad y requiere que los tamaños de muestra sean iguales; cuando los tamaños de muestra no son iguales, la prueba no tiene soporte teórico fuerte.

### 2.3. Prueba de Levene (L)

Levene (1960) propone contrastar si  $k$  tratamientos tienen la misma varianza haciendo un análisis de varianza de una vía sobre el valor absoluto de las desviaciones de los valores muestrales respecto a la media. El valor absoluto de los residuales está dado por  $z_{ij} = |y_{ij} - \bar{y}_i|$ ,  $i = 1, 2, \dots, k$ ;  $j = 1, 2, \dots, n_i$ . El estadístico a calcular es

$$L = \frac{\sum_{i=1}^k n_i (\bar{z}_{i\cdot} - \bar{z}_{\cdot\cdot})^2 / (k - 1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i\cdot})^2 / \sum_{i=1}^k n_i - k} \quad (2.6)$$

Donde:

$$\bar{z}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} z_{ij} \text{ y } \bar{z}_{\cdot\cdot} = \sum_{i=1}^k \sum_{j=1}^{n_i} z_{ij} / \sum_{i=1}^k n_i$$

La regla de decisión es rechazar  $H_0$  si  $L > F_{(N-k)}^{(k-1)}$ .

La prueba de Levene es menos sensible a la falta de normalidad que la prueba Bartlett y es una herramienta muy popular para probar homogeneidad de varianzas, por lo que ha tenido diversas modificaciones las cuales buscan mejorar su potencia, algunas de estas modificaciones se mencionan a continuación.

## 2.4. Modificaciones a la Prueba de Levene

### 2.4.1. Prueba Levene modificada por Brown y Forsythe

Una de las primeras modificaciones que se le hizo a la prueba de Levene fue la propuesta por [Brown y Forsythe \(1974\)](#), quienes extienden la prueba original al utilizar la mediana ( $\tilde{y}_i$ ) y la media truncada al 10% ( $\bar{y}_i^{(10)}$ ) en lugar de la media ( $\bar{y}_i$ ), (la media truncada al 10% se calcula obteniendo la media de las observaciones de la muestra eliminando el 10% de los datos del extremo inferior y superior).

Por lo tanto, el nuevo estadístico resulta de usar  $w_{1ij} = |y_{ij} - \tilde{y}_i|$  y  $w_{2ij} = |y_{ij} - \bar{y}_i^{(10)}|$  en lugar de  $z_{ij}$  en la ecuación (2.6). Por lo que, el estadístico propuesto por [Brown y Forsythe \(1974\)](#) basado en la mediana (Lmed) se calcularía como:

$$Lmed = \frac{\sum_{i=1}^k n_i (\bar{w}_{1i\cdot} - \bar{w}_{1\cdot\cdot})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (w_{1ij} - \bar{w}_{1i\cdot})^2 / \sum_{i=1}^k n_i - k} \quad (2.7)$$

Donde:

$$\bar{w}_{1i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} w_{1ij} \text{ y } \bar{w}_{1\cdot\cdot} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} w_{1ij}}{\sum_{i=1}^k n_i}$$

La regla de decisión es rechazar  $H_0$  si  $Lmed > F_{(N-k)}^{(k-1)}$ .

Los estudios de Monte Carlo de [Brown y Forsythe \(1974\)](#) muestran que al utilizar la media truncada se mejora la prueba original cuando los datos siguen una distribución de Cauchy (colas grandes) y la mediana conseguía mejorarlo cuando los datos seguían una  $\chi_4^2$  (distribución asimétrica).

## 2.4. Modificaciones a la Prueba de Levene

---

### 2.4.2. Modificaciones para tamaños de muestras desbalanceados

Para el caso de diseños no balanceados existen diversas contribuciones, tal es el caso de [Keyes y Levy \(1997\)](#) quienes resuelven el problema de desbalance sugiriendo usar  $u_{ij} = |y_{ij} - \bar{y}_i| / (1 - 1/n_i)^{1/2}$  en lugar de  $z_{ij}$  en (2.6). Por su parte [O'Brien \(1978\)](#) sugiere una prueba similar usando  $v_{ij} = u_{ij}^2$ .

[O'Brien \(1981\)](#) también sugiere reemplazar las observaciones originales  $y_{ij}$  con nuevas observaciones dadas por

$$r_{ij} = \frac{(w - n_i - 2)n_i(y_{ij} - \bar{y}_i)^2 - ws_i^2(n_i - 1)}{(n_i - 1)(n_i - 2)} \quad (2.8)$$

donde  $n_i$ ,  $\bar{y}_i$  y  $s_i$  son el tamaño muestral, la media y la desviación estándar del  $i$ -ésimo grupo y  $w$  ( $0 \leq w \leq 1$ ) es un parámetro.

Cuando se tienen muestras pequeñas (no más de 10 observaciones en cada grupo), la prueba de Levene puede ser poco conservadora al trabajar con medianas. [Hines y O'Hara~Hines \(2000\)](#) proponen una modificación, si  $n_i$  en el  $i$ -ésimo grupo es impar, se debe usar  $z_{im}$ ,  $m = [n_i/2] + 1$ , ( $[y]$  es la función que redondea a  $y$ ), si  $n_i$  es par las dos desviaciones iguales y más pequeñas  $z_{i[n_i/2]}$  y  $z_{i[n_i/2+1]}$  son reemplazadas por un solo valor  $\sqrt{2}z_{i[n_i/2]}$ .

Este procedimiento incrementa la variabilidad de  $z_{ij}$ , y proporciona un tamaño de prueba cercano al nominal, además ofrece una prueba tipo Levene con incremento de potencia.

### 2.4.3. Versión no paramétrica (NP) de la prueba de Levene

Una de las modificaciones más recientes a la prueba de Levene es la que proponen [Nordstokke y Zumbo \(2010\)](#) quienes hacen uso del principio de transformación a rangos ([Conover y Iman \(1981\)](#)) sobre las observaciones originales. Ellos sustituyen en (2.6)  $z_{ij}$  por  $w_{ij}^* = |R_{ij} - \bar{y}_j^R|$ , donde  $R_{ij}$  es el rango de cada una de las observaciones originales y  $\bar{y}_j^R$  es la media de los rangos por grupo.

## 2.4. Modificaciones a la Prueba de Levene

---

$$NP = \frac{\sum_{i=1}^k n_i (\bar{w}_{i.}^* - \bar{w}^{*..})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (w_{ij}^* - \bar{w}_{i.}^*)^2 / \sum_{i=1}^k n_i - k} \quad (2.9)$$

Donde:

$$\bar{w}_{i.}^* = \frac{1}{n_i} \sum_{j=1}^{n_i} w_{ij}^* \text{ y } \bar{w}^{*..} = \sum_{i=1}^k \sum_{j=1}^{n_i} w_{ij}^* / \sum_{i=1}^k n_i$$

La regla de decisión es rechazar  $H_0$  si  $NP > F_{(N-k)}^{(k-1)}$ .

# Capítulo 3

## Método propuesto

### 3.1. Prueba Levene Transformada

La prueba propuesta, que se nombrará Levene Transformada (LT), hace uso del principio de transformación a rangos sobre el valor absoluto de los residuales del modelo, para después aplicar un análisis de varianza a los datos obtenidos, es decir, LT se define de la siguiente manera:

Sea  $z_{ij} = |y_{ij} - \bar{y}_i|$ ,  $i = 1, 2, \dots, k$ ;  $j = 1, 2, \dots, n_i$ , el valor absoluto de los residuos y sea  $r_{ij}$  la transformación a rangos de  $z_{ij}$ , entonces

$$LT = \frac{\sum_{i=1}^k n_i (\bar{r}_{i\cdot} - \bar{r}_{\cdot\cdot})^2 / (k - 1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (r_{ij} - \bar{r}_{i\cdot})^2 / \sum_{i=1}^k n_i - k} \quad (3.1)$$

Con  $\bar{r}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} r_{ij}$  y  $\bar{r}_{\cdot\cdot} = \sum_{i=1}^k \sum_{j=1}^{n_i} r_{ij} / \sum_{i=1}^k n_i$

Donde  $r_{ij}$  representa la transformación a rangos del residual en valor absoluto del j-ésimo sujeto en el i-ésimo tratamiento.

### 3.2. Ejemplo de aplicación

## 3.2. Ejemplo de aplicación

En esta sección se da un ejemplo de aplicación de la forma en la que se calcula la prueba LT y las demás que se usaron en el estudio de simulación (B, L, Lmed y NP). Los datos fueron tomados del Ejemplo 2.4-Drogas y Errores del libro de [Milliken y Johnson \(2002\)](#) y corresponden al efecto de dos posibles drogas. Estos resultados se encuentran divididos en cuatro grupos, el grupo 1 es el grupo control (no se les aplicó la droga), al grupo 2 se le dio la droga 1, al grupo 3 se le dio la droga 2 y al grupo 4 se le dieron ambas drogas. Los datos se presentan en la tabla 3.1.

**Tabla 3.1:** Datos del experimento de aprendizaje de pares asociados.

	Sin droga	Droga 1	Droga 2	Drogas 1 y 2
	1	12	12	13
	8	10	4	14
	9	13	11	14
	9	13	7	17
	4	12	8	11
	0	10	10	14
	1	-	12	13
	-	-	5	14
$n$	7	7	8	8
Suma	32	70	69	110
Mediana	4	12	9	14
Media	4.5714	11.6667	8.6250	13.750
Varianza	16.2857	1.8667	9.6964	2.786

Los tamaños de muestra, así como las sumas, las medianas, las medias y las varianzas de cada grupo se encuentran en la tabla 3.1. El primer estadístico que se calcula es el de la prueba LT.

Para calcular la prueba LT se comienza con el cálculo de las desviaciones de los valores muestrales respecto a la media de cada tratamiento (residuales del modelo), para después aplicar el valor absoluto a estos valores, los resultados se muestran en la tabla 3.2.

Como paso siguiente al cálculo del valor absoluto de los residuales esta la transformación a rangos,  $r_{ij}$ , de los valores  $z_{ij}$ , los resultados de aplicar este paso se muestran en la tabla 3.3.

### 3.2. Ejemplo de aplicación

**Tabla 3.2:** Valores de  $z_{ij} = |y_{ij} - \bar{y}_i|$

Sin droga	Droga 1	Droga 2	Drogas 1 y 2
3.571	0.333	3.375	0.750
3.429	1.667	4.625	0.250
4.429	1.333	2.375	0.250
4.429	1.333	1.625	3.250
0.571	0.333	0.625	2.750
4.571	1.667	1.375	0.250
3.571	-	3.375	0.750
-	-	3.625	0.250

**Tabla 3.3:** Valores de la transformación a rangos ( $r_{ij}$ ) de los valores  $z_{ij}$

Sin droga	Droga 1	Droga 2	Drogas 1 y 2
23.5	6.5	20.5	3
22	15.5	29	3
26.5	11.5	17	3
26.5	11.5	14	19
8	6.5	9	18
28	15.5	13	3
23.5	-	20.5	10
-	-	25	3
$\bar{r}_i.$	22.571	11.167	18.5

Finalmente, se aplica un análisis de varianza a los datos de la tabla 3.3, usando la ecuación (3.1).

Algunos de los cálculos se muestran a continuación:

$$\sum_{i=1}^4 n_i = 29, \quad \bar{r}_{..} = \frac{\sum_{i=1}^4 \sum_{j=1}^{n_i} r_{ij}}{\sum_{i=1}^4 n_i} = \frac{23.5+22+\dots+10+3}{29} = 15$$

$$\frac{\sum_{i=1}^4 n_i (\bar{r}_i. - \bar{r}_{..})^2}{(4-1)} = \frac{7(22.571 - 15)^2 + \dots + 8(7.75 - 15)^2}{3} = \frac{1007.952}{3} = 335.984$$

### 3.2. Ejemplo de aplicación

---

$$\frac{\sum_{i=1}^4 \sum_{j=1}^{n_i} (r_{ij} - \bar{r}_{i.})^2}{\sum_{i=1}^4 n_i - 4} = \frac{(23.5 - 22.571)^2 + (22 - 22.571)^2 + \dots + (10 - 7.75)^2 + (3 - 7.75)^2}{29 - 4}$$
$$= \frac{1009.048}{25} = 40.362$$

Por lo que  $LT = \frac{335.984}{40.362} = 8.324$ , como LT sigue la misma distribución que el estadístico original de Levene, entonces se pueden usar los valores de la tabla F, con lo cual se obtiene que el nivel de significancia observado para un valor de tablas F de 8.324 con 3 y 25 grados de libertad es 0.0005.

Los cálculos para la prueba de Bartlett son:

$$C = 1 + \frac{1}{3 \times 3} \left( \frac{1}{6} + \frac{1}{5} + \frac{1}{7} + \frac{1}{7} - \frac{1}{25} \right)$$

y

$$\hat{\sigma}^2 = \frac{6(16.2857) + 5(1.8667) + 7(9.6964) + 7(2.7860)}{25} = 7.7769$$

Entonces, de la ecuación (2.4),

$$U = \frac{1}{C} \left( \nu \log_e \hat{\sigma}^2 - \sum_{i=1}^4 \nu_i \log_e \hat{\sigma}_i^2 \right)$$
$$= \frac{1}{1.068} [25 \log_e(7.7769) - 6 \log_e(16.2857) - 5 \log_e(1.8667)$$
$$- 7 \log_e(9.6964) - 7 \log_e(2.7860)]$$
$$= 7.8111$$

La distribución muestral asintótica asociada a U es la distribución chi-cuadrada basada en tres grados de libertad. El nivel de significancia de la prueba es 0.0501.

Los primeros cálculos para el test de Levene (L) son iguales a la prueba LT, se comienza con el cálculo de los residuos de los tratamientos. Seguido de los valores



### 3.2. Ejemplo de aplicación

absolutos de los residuos, los cuales se calculan como se ilustra en la Tabla 3.2. Finalmente, se aplica un análisis de varianza de una vía a estos datos, usando la ecuación (2.6).

$$\bar{z}_{..} = \frac{\sum_{i=1}^4 \sum_{j=1}^{n_i} z_{ij}}{\sum_{i=1}^4 n_i} = \frac{3.571 + 3.429 + \dots + 0.750 + 0.250}{29} = 2.094$$

$$\begin{aligned} L &= \frac{\sum_{i=1}^4 n_i (\bar{z}_{i.} - \bar{z}_{..})^2 / (4 - 1)}{\sum_{i=1}^4 \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_{i.})^2 / (29 - 4)} \\ &= \frac{7(3.510 - 2.094)^2 + \dots + 8(1.063 - 2.094)^2 / 3}{(3.571 - 3.510)^2 + \dots + (0.250 - 1.063)^2 / 25} \\ &= 6.97 \end{aligned}$$

El valor resultante del estadístico L es 6.97, que se basa en los 3 y 25 grados de libertad. El nivel de significancia observado de la prueba de Levene es 0.0015.

Para la prueba Lmed, se calcula el valor absoluto de las desviaciones de las observaciones respecto a la mediana de cada tratamiento (las medianas se muestran en la Tabla 3.2). La Tabla 3.4 contiene los valores absolutos de las desviaciones de las observaciones respecto a las medianas.

**Tabla 3.4:** Valor absoluto de las desviaciones de las observaciones respecto a la mediana de cada tratamiento

Sin droga	Droga 1	Droga 2	Drogas 1 y 2
3	0	3	1
4	2	5	0
5	1	2	0
5	1	2	3
0	0	1	3
4	2	1	0
3	-	3	1
-	-	4	0

### 3.2. Ejemplo de aplicación

---

El paso siguiente es hacer el análisis de varianza de una sola vía usando (2.7), realizando los cálculos correspondientes se obtiene que el valor de Lmed es 5.49 y el nivel de significancia observada es 0.0049 (también basado en 3 y 25 grados de libertad).

Por último, para calcular el estadístico NP, se requiere de los rangos de las observaciones originales ( $y_{ij}$ ), los cuales se muestran en la tabla 3.5, también se incluye la media de los rangos por grupo.

**Tabla 3.5:** Transformación a rangos de las observaciones originales ( $y_{ij}$ )

Sin droga	Droga 1	Droga 2	Drogas 1 y 2	
	2.5	3.5	7.5	2.5
	5	1.5	1	5.5
	6.5	5.5	6	5.5
	6.5	5.5	3	8
	4	3.5	4	1
	1	1.5	5	5.5
	2.5	-	7.5	2.5
	-	-	2	5.5
media	4	3.5	4.5	4.5

El siguiente paso es calcular el estadístico de la prueba NP, usando (2.9), al realizar los cálculos correspondientes se obtiene que el resultado para NP es 2.676, con 3 y 25 grados de libertad, con un nivel de significancia observado de 0.069.

En la tabla 3.6 se muestra un resumen de los resultados obtenidos en cada una de las pruebas.

**Tabla 3.6:** Resumen de resultados

Estadístico de Prueba	Valor	Significancia
LT	8.324	0.0005
B	7.8111	0.0501
L	6.97	0.0015
Lmed	5.49	0.0049
NP	2.678	0.069

# Capítulo 4

## Evaluación del método propuesto

Para evaluar el comportamiento de las pruebas mencionadas y de la prueba LT se usó el método Monte Carlo y se analiza la potencia y el tamaño de las pruebas en diversas situaciones. El nivel de significancia usado es de 5% y los resultados de simulación se basan en 10,000 réplicas.

### 4.1. Factores estudiados en la simulación

Las variables que se manipularon en la simulación son las siguientes:

- Pruebas de homogeneidad de varianzas
- Función de distribución de los errores
- Homogeneidad y heterogeneidad de varianzas
- Tamaño de muestra
- Tamaño de efecto

## 4.1. Factores estudiados en la simulación

---

### 4.1.1. Pruebas de homogeneidad de varianzas

Las pruebas de homogeneidad de varianzas que se utilizaron en el estudio de simulación fueron 5, la prueba de Bartlett (B), la prueba original de Levene (L), la modificación de la prueba Levene propuesta por [Brown y Forsythe \(1974\)](#) basada en medianas (Lmed), la No paramétrica (NP) y la prueba de Levene transformada (LT) todas ellas descritas en los capítulos [2](#) y [3](#).

### 4.1.2. Función de distribución de los errores

Se generaron números aleatorios de seis distribuciones diferentes, las cuales se clasifican en tres rubros, simétricas, simétricas con colas pesadas y distribuciones sesgadas. Dentro de las distribuciones simétricas se consideraron la normal y la uniforme. Para las simétricas con colas pesadas están la t de student y la doble exponencial. La chi-cuadrada y la exponencial se consideraron dentro de las asimétricas.

### 4.1.3. Tamaño de muestra

El estudio de simulación se basa en 4 tratamientos. Se consideraron tamaños de muestra pequeños (4 y 5) que son frecuentemente usados en experimentos biológicos o agrícolas, tamaño mediano (10) y grande (20).

### 4.1.4. Tamaño de efecto

Con respecto al tamaño de efecto de tratamiento, se consideraron cuatro clasificaciones diferentes: cero, para el caso en el cual no existe efecto de tratamiento; pequeño para el caso donde el efecto del tratamiento es 0.1, es decir, la diferencia entre medias está a 0.1 desviaciones típicas; mediano, cuando el efecto es cercano o igual a 0.25, es decir la diferencia entre medias se encuentra a un cuarto de una desviación estándar y grande cuando el efecto es de 0.4 (Cohen, 1977).

## 4.2. Potencia y Tamaño de la prueba

Siempre que se toma una decisión de rechazar o no rechazar una hipótesis nula ( $H_0$ ) existe la probabilidad de cometer dos tipos de errores:

Error tipo I: rechazar  $H_0$  cuando  $H_0$  es verdadera

Error tipo II: no rechazar  $H_0$  cuando  $H_0$  es falsa

Lo deseable sería que los dos tipos de error tuvieran probabilidades pequeñas, sin embargo esto no es posible, ya que cuando se minimiza la probabilidad de alguno de los errores aumenta la probabilidad de cometer el otro error, por esta razón en la práctica se fija la probabilidad de cometer el error tipo I y se trata de minimizar la probabilidad del error tipo II.

### 4.2.1. Potencia de la prueba

*Función de potencia* Casella y Berger (1990): la función de potencia de una prueba de hipótesis con región de rechazo  $X_R$  es la función de  $\theta$  definida por:

$$\beta(\theta) = P(\text{No rechazar } H_0 \text{ cuando } H_0 \text{ es falsa}) = P(X \in X_R)$$

O lo que es lo mismo,

$$\beta(\theta) = 1 - P(\text{Error tipo II})$$

Para evaluar potencia, se generaron datos en situaciones donde la hipótesis nula no era cierta, esto es, existía heterogeneidad de varianzas.

Para calcular la potencia estimada de la prueba  $LT$  usando simulación Monte Carlo se tomaron en cuenta los siguientes pasos:

1. Se generaron los datos bajo las condiciones estudiadas en la simulación, con situaciones en donde existía heterogeneidad de varianzas. Para probar (2.3)

### 4.3. Determinación de la robustez

---

2. Obtener la función de distribución de los errores para las diferentes distribuciones descritas en la sección 4.1 tomando en cuenta la heterogeneidad de varianzas.
3. Con los datos generados realizar las pruebas de heterogeneidad de varianzas que vamos a comparar.
4. Repetir los paso del 2 a 3 hasta obtener  $m$  número de simulaciones.
5. Registrar para cada una de las  $m$  repeticiones si se rechaza la hipótesis nula o no.
6. Finalmente, se estima la potencia de la prueba ( $\hat{\beta}(\theta)$ ) como el cociente del número de rechazos entre el número total de las  $m$  repeticiones de las pruebas de heterogeneidad de varianzas.

#### 4.2.2. Tamaño de la prueba

*Tamaño de la prueba:* Para  $0 < \alpha < 1$ , una prueba con función de potencia  $\beta(\theta)$  es una prueba de tamaño  $\alpha$  si:

$$\sup_{\theta} \beta(\theta) = \alpha$$

donde  $\alpha = P(\text{Rechazar } H_0, \text{ cuando } H_0 \text{ es verdadera})$

Para calcular el tamaño de prueba estimado ( $\hat{\alpha}$ ), se siguen los mismos pasos que para el cálculo de la potencia estimada. Por lo que, la estimación del tamaño de prueba está dado como el cociente del número de rechazos entre el número total de simulaciones Monte Carlo realizadas

### 4.3. Determinación de la robustez

La robustez de la prueba propuesta se evaluó mediante la tasa de error Tipo I. Para esto, se utilizó el criterio de robustez de [Bradley \(1978\)](#).

### 4.3. Determinación de la robustez

---

Este criterio establece que una prueba es robusta si las tasas empíricas de error Tipo I se encuentran comprendidas entre el intervalo  $0.5\alpha \leq \hat{\alpha} \leq 1.5\alpha$ . En el trabajo, se consideró  $\alpha = 0.05$ , por lo que el intervalo para la robustez es  $[0.025, 0.075]$ .

Usando este criterio se considera que el procedimiento es liberal cuando la tasa error Tipo I empírica es mayor que el límite superior y que el procedimiento es conservador cuando está por debajo del límite inferior.

En el Anexo A se muestran las tablas que contienen los tamaños de prueba estimados para cada una de las pruebas consideradas en el estudio de simulación.

# Capítulo 5

## Resultados y discusión

En este apartado se presentan los resultados del estudio de simulación mediante el uso de figuras, con el fin de mostrar el comportamiento de la prueba propuesta (LT) en comparación con las pruebas citadas en la literatura, el análisis se divide en dos secciones, uno para el tamaño de la prueba y el otro para la potencia.

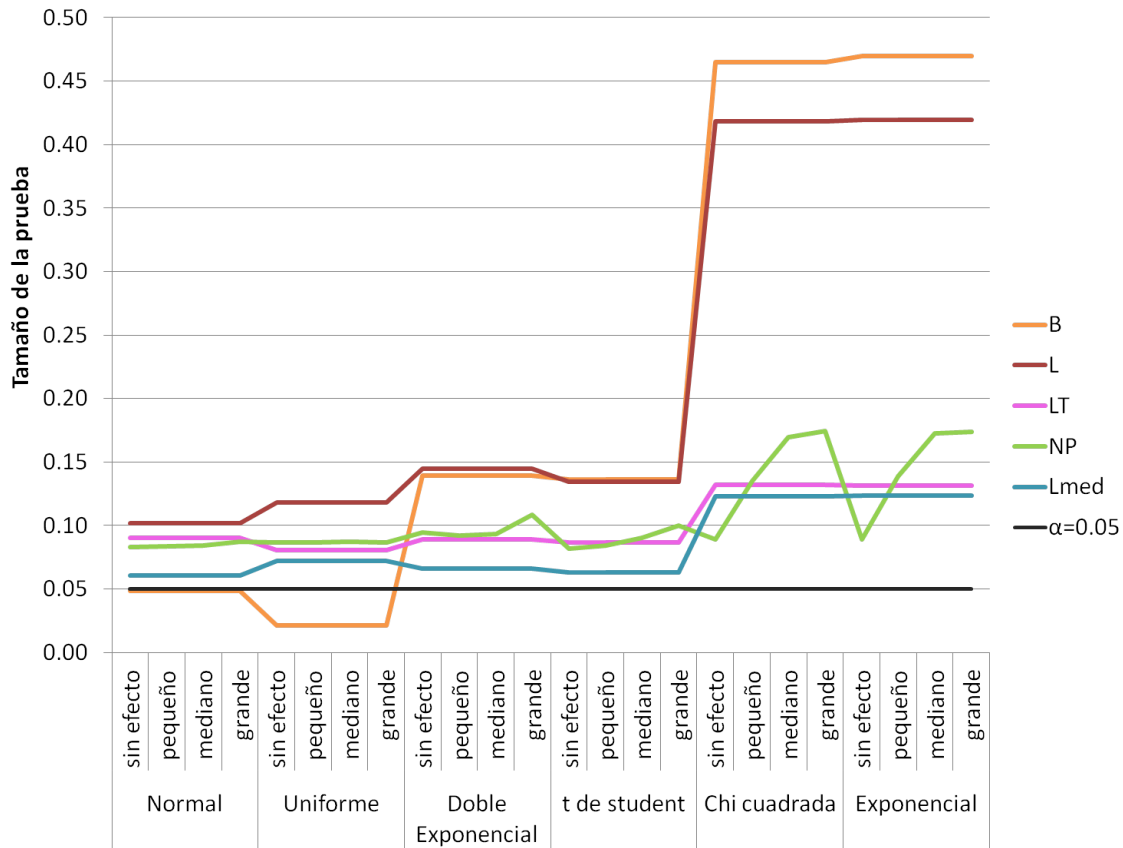
### 5.1. Tamaño de la prueba

Para el caso de los datos que provienen de una distribución normal, la mejor prueba en cuanto a tamaño de prueba es la B al mostrar tasas de error tipo I de 0.05 para todos los tamaños de muestra y todos los tamaños de efecto, coincidiendo con lo reportado por [Parra-Frutos \(2009\)](#), y [Mendes y Özcaya Turhan \(2006\)](#). Se observó también que esta prueba es muy sensible a desviaciones de normalidad en los datos, coincidiendo con lo que establece la literatura ([Conover y Iman \(1981\)](#); [Lim y Loh \(1996\)](#)), ocasionando con esto tamaños de prueba muy superiores al nivel nominal de significancia (valores mayores a 0.45) cuando los datos provienen de distribuciones asimétricas ([Sharma y Golam-Kibria \(2012\)](#)). Para muestras grandes ( $n=20$ ) prácticamente todas las pruebas tuvieron tamaños de prueba cercanos al nominal (5.4), para muestras pequeñas ( $n=4$  y  $5$ ) la prueba de L mostró los tamaños de prueba estimados más alejados a los valores nominales, para tamaño de muestra  $n=5$  la prueba Lmed mostró un comportamiento demasiado conservador con valores que oscilan entre 0.001 y 0.0244 (figura 5.2).



## 5.1. Tamaño de la prueba

La figura 5.1, reporta los resultados para el tamaño de muestra  $n=4$ , bajo este tamaño de muestra la prueba Lmed tuvo un comportamiento robusto (Bradley (1978)) o cuando se aplica a distribuciones simétricas y simétricas con colas pesadas, al presentar tamaños de prueba de entre 0.06 y 0.07 para estas distribuciones, mientras que para las distribuciones asimétricas (chi-cuadrada y exponencial) la tasa de error tipo I fue superior a 0.12. Las pruebas restantes (L, LT, NP), no pueden ser consideradas robustas puesto que las tasas de error tipo I son superiores al 0.08 para todos los casos; se observa que el tamaño de prueba para L supera el 0.1 cuando los datos son asimétricos (chi-cuadrada y exponencial) y que los valores de la tasa de error tipo I, para la prueba NP, oscilan entre 0.08 y 0.11 cuando los datos provienen de distribuciones simétricas, pero en el caso de las distribuciones asimétricas la tasa de error tipo I aumenta conforme aumenta el tamaño del efecto.



**Figura 5.1:** Tamaño de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de muestra  $n=4$ .

Por otro lado, cuando el tamaño de muestra se aumenta en una unidad ( $n=5$ ) la

## 5.1. Tamaño de la prueba

---

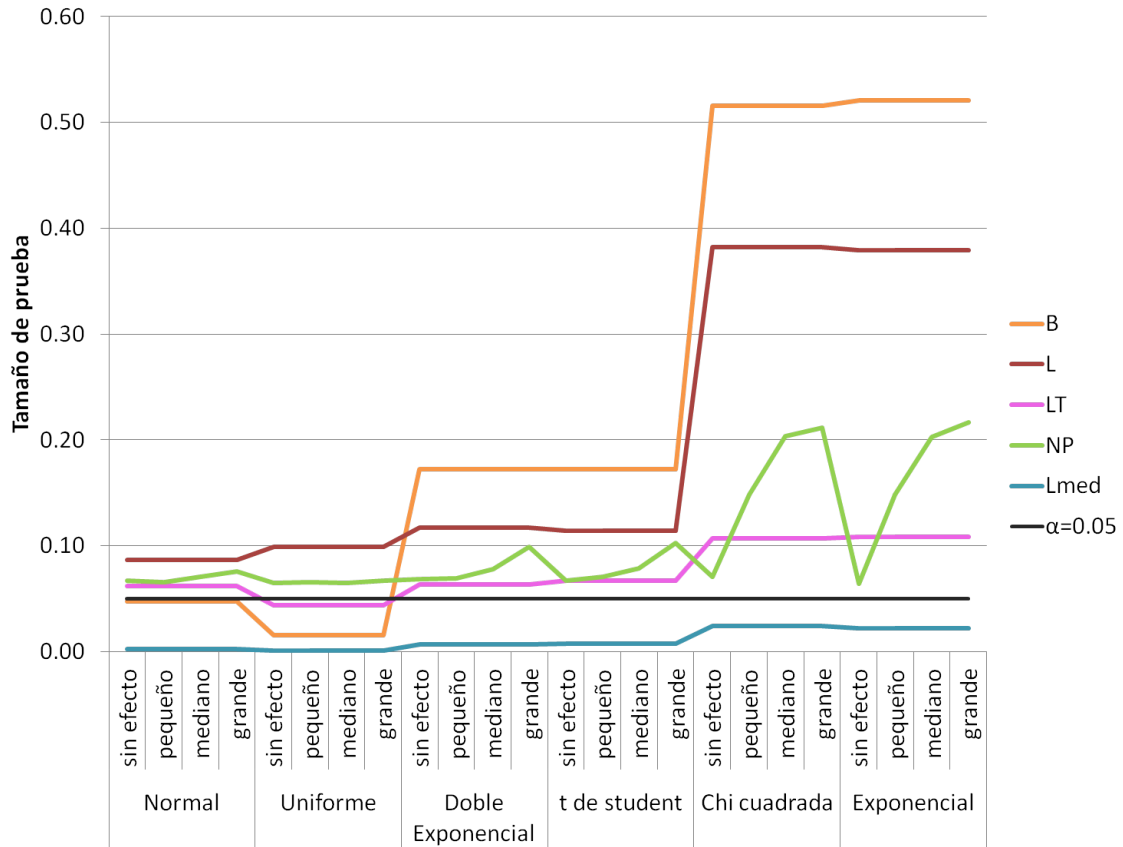
mayoría de las pruebas reducen su tasa de error tipo I (figura 5.2). El comportamiento de las pruebas B, L, LT y NP es muy semejante al observado en la figura 5.1 (tamaño de muestra igual a 4), pero se puede resaltar que la prueba LT mantiene valores del tamaño de prueba estimado entre 0.0438 y 0.0669 para los datos que provienen de las distribuciones normal, uniforme, doble exponencial y t de student ocasionando con esto una prueba robusta. Por otro lado, la prueba Lmed se vuelve una prueba conservadora, al mostrar un tamaño de prueba estimado por debajo del valor nominal, con valores que van de 0.001 a 0.0244, un resultado similar encuentra [Parra-Frutos \(2012\)](#), quien encontró que la prueba de Levene basada en medianas es demasiado conservadora cuando los tamaños de muestra son pequeños e impares.

En este caso, la prueba NP muestra cierto grado de robustez cuando los datos provienen de distribuciones simétricas y de colas pesadas, sin embargo cuando se habla de distribuciones sesgadas la tasa de error tipo I de esta prueba tienden a aumentar conforme crece el tamaño del efecto llegando a un tamaño superior al 20 % cuando el tamaño de efecto considerado en el modelo es grande (Figura 5.2). En general, en el caso de muestras pequeñas ( $n=4$  y  $n=5$ ), las pruebas tomadas en el análisis de simulación muestran un comportamiento malo en la mayoría de las condiciones estudiadas, a excepción de la prueba B aplicada a datos normales.

Por otro lado, si el tamaño de muestra considerado es de  $n=10$ , la prueba L muestra una mejoría cuando los datos son normales o uniformes, presentando tamaños de prueba de 0.0638 y 0.0661 respectivamente; los tamaños de prueba para LT, NP y Lmed también mejoran, las pruebas LT y NP vuelven a ser robustas para datos de poblaciones simétricas (Figura 5.3). Sin embargo, la prueba NP vuelve a mostrar una alta dependencia de la prueba con el tamaño de efecto, en la figura 5.3 se observa (sobre todo en las distribuciones chi cuadrada y exponencial), que el tamaño de prueba de NP aumenta conforme aumenta el tamaño de efecto, por ejemplo, para la distribución chi-cuadrada se puede ver que si en el modelo no se toma en cuenta el efecto las tasas de error tipo I son de 0.0547 mientras que cuando el tamaño del efecto es pequeño las tasas de error tipo I ya superan el 0.2.

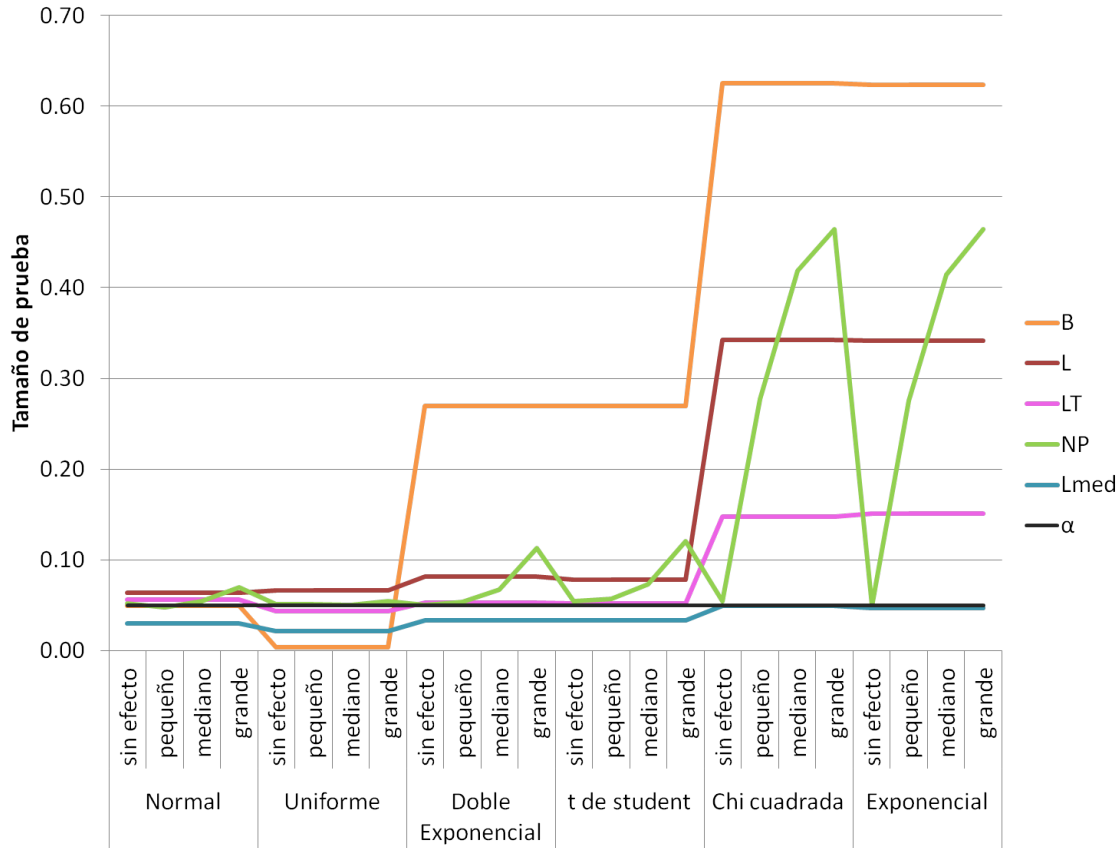
En el caso del tamaño de la prueba Lmed éste se vuelve aceptable, siendo conservadora para los datos uniformes, con valores de 0.0217 ([Parra-Frutos \(2012\)](#) obtiene tamaños de prueba de 0.033 para esta distribución) y presentando tamaños de prueba robustos para las demás distribuciones (valores entre 0.03 y 0.049).

## 5.1. Tamaño de la prueba



**Figura 5.2:** Tamaño de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de muestra  $n=5$ .

## 5.1. Tamaño de la prueba

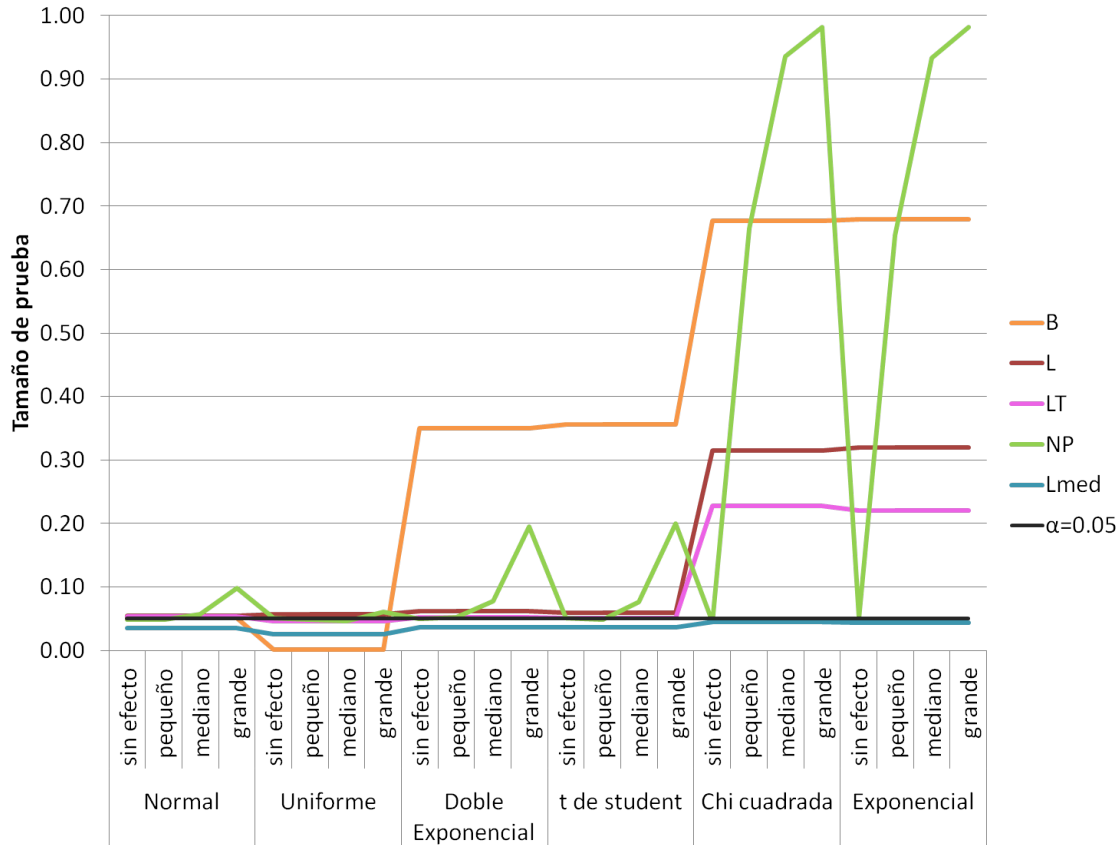


**Figura 5.3:** Tamaño de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de muestra  $n=10$ .

Si el tamaño de muestra aumenta a 20 unidades, las pruebas L y LT mantienen tasas de error tipo I robustas para datos de poblaciones simétricas, presentando valores de entre 0.055 y 0.0626 para la prueba L y de 0.0469 y 0.0522 para la prueba LT (figura 5.4). La prueba NP es considerada robusta en la distribución uniforme, con valores de 0.0485 y 0.0606, en las demás distribuciones el tamaño de prueba aumenta conforme crecen los tamaños de efecto, de hecho, cuando las distribuciones consideradas son sesgadas los tamaños de prueba superan el 0.6 para tamaños de efecto pequeños llegando a ser casi 1 cuando el efecto es grande (figura ??). La prueba Lmed, vuelve a ser la de mejor comportamiento con tamaños ligeramente inferiores al nivel nominal de significancia, Parra-Frutos (2012) obtiene para esta misma prueba tamaños de error tipo I entre 0.04 y 0.055.

En resumen, la prueba de B tiene el mejor comportamiento en la distribución normal, pero muy mal comportamiento en el resto de las distribuciones. En este estudio la

## 5.1. Tamaño de la prueba



**Figura 5.4:** Tamaño de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de muestra  $n=20$ .

prueba L mostró un comportamiento muy malo en las distribuciones sesgadas bajo todas las condiciones estudiadas, Shoemaker (2003) por ejemplo, demostró mediante un estudio de simulación que bajo violaciones de normalidad la tasa de error tipo I de la prueba Levene se incrementa. Mendes y Özcaya Turhan (2006), por su parte, encuentran que el tamaño de la prueba Levene se desvía del nivel predeterminado de alfa cuando la distribución de los datos es chi-cuadrada. En el caso de distribuciones simétricas y tamaños de muestra pequeños la prueba L también fue mala. La prueba NP en general, fue peor que la prueba LT en todas las condiciones, pero además, mostró una gran dependencia del tamaño del efecto, al aumentar el tamaño de efecto, aumenta también el valor del tamaño de la prueba estimado, llegando a un valor cercano a uno (0.9815) para modelos con un tamaño de efecto grande. La prueba Lmed fue la que mostró valores más cercanos al valor nominal cuando se aplica a datos que provienen de distribuciones simétricas de colas pesadas y las asimétricas,

## 5.2. Potencia de la prueba

---

comportamientos similares fueron encontrados por [Conover \*et al.\* \(1981\)](#) y [Carroll y Schneider \(1985\)](#), sin embargo, con el tamaño  $n$  igual a cinco la prueba tuvo un comportamiento demasiado conservador, resultado que concuerda con lo que establece [Parra-Frutos \(2012\)](#) quien encuentra que con un tamaño de muestra pequeño e impar la prueba de Levene basada en medianas es muy conservadora. Por otro lado, si los datos son uniformes la prueba que presenta un mejor tamaño es la LT.

Los resultados del estudio de simulación mostraron que, para la prueba LT, la tasa de error tipo I tiende a disminuir conforme el tamaño de la muestra aumenta en el caso de las distribuciones simétricas (ver figuras [5.1](#), [5.2](#), [5.3](#) y [5.4](#)), sin embargo, en las poblaciones que provienen de distribuciones sesgadas, como son la chi-cuadrada y la exponencial, el tamaño del error tipo I tiende a aumentar superando por mucho el nivel nominal de la prueba de  $\alpha = 0.05$ .

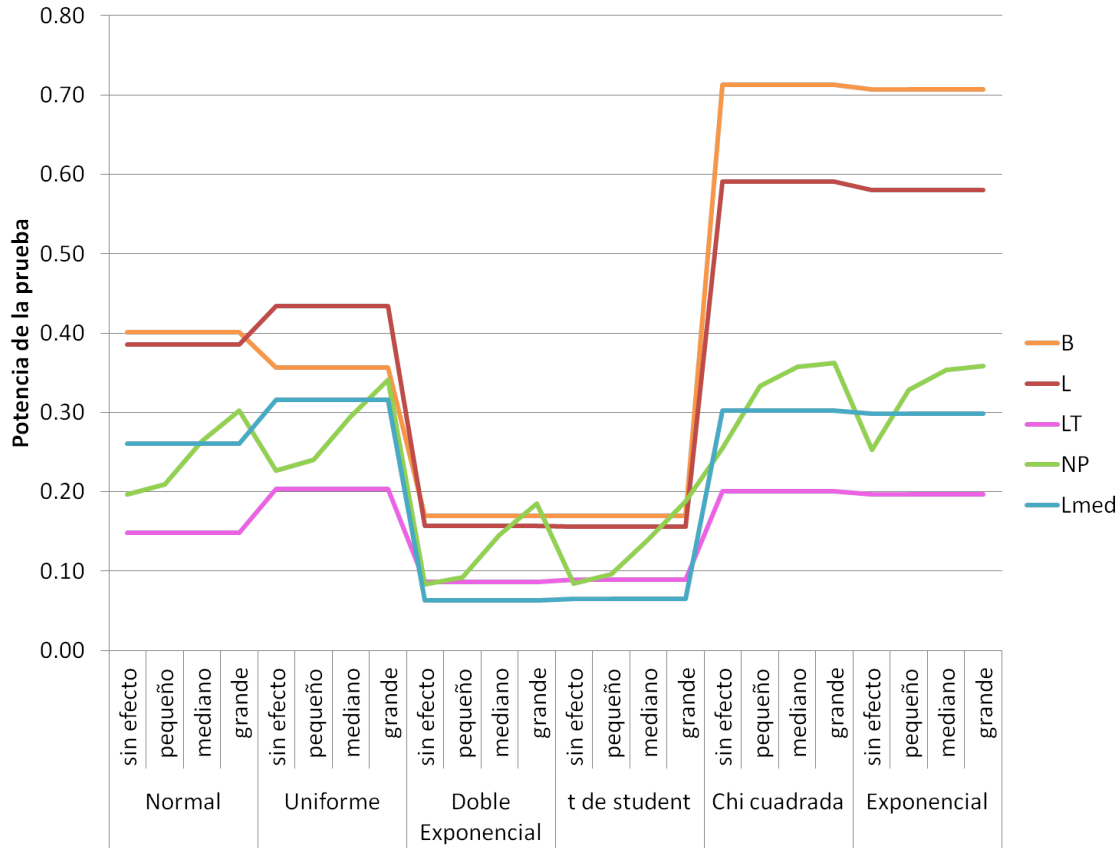
## 5.2. Potencia de la prueba

Bajo un tamaño de muestra de  $n=4$ , todas las pruebas mostraron un comportamiento inaceptable con la distribución doble exponencial y la  $t$  de student. Cuando las distribuciones consideradas en el estudio son las asimétricas, las pruebas que presentan una mejor potencia son la Bartlett y la Levene, siendo más potente la prueba Bartlett, sin embargo, tuvieron valores inaceptables en el tamaño de prueba estimado (figura [5.5](#)).

La prueba Lmed tuvo mayor potencia que la prueba LT en las distribuciones normal, uniforme, chi chadrada y exponencial; la potencia de Lmed se encuentra entre valores de 26.1 y 30.25 %, mientras que para la prueba LT los valores oscilan entre 14.79 y 19.66 %. (figura 6.5). En el caso de las distribuciones de colas pesadas, doble exponencial y  $t$  de student, la prueba LT le gana en potencia a la Lmed. Al igual que en el tamaño de prueba, la potencia de la NP muestra una dependencia del tamaño del efecto, para esta prueba los niveles mas altos de potencia se presentaron con las distribuciones normal, uniforme y las sesgadas.

Si el tamaño de la muestra aumenta en una unidad, figura [5.6](#), los niveles de potencia para la prueba Lmed bajan drásticamente en todas las condiciones estudiadas, en este caso, la potencia de esta prueba es superada por todas las demás. La prueba

## 5.2. Potencia de la prueba



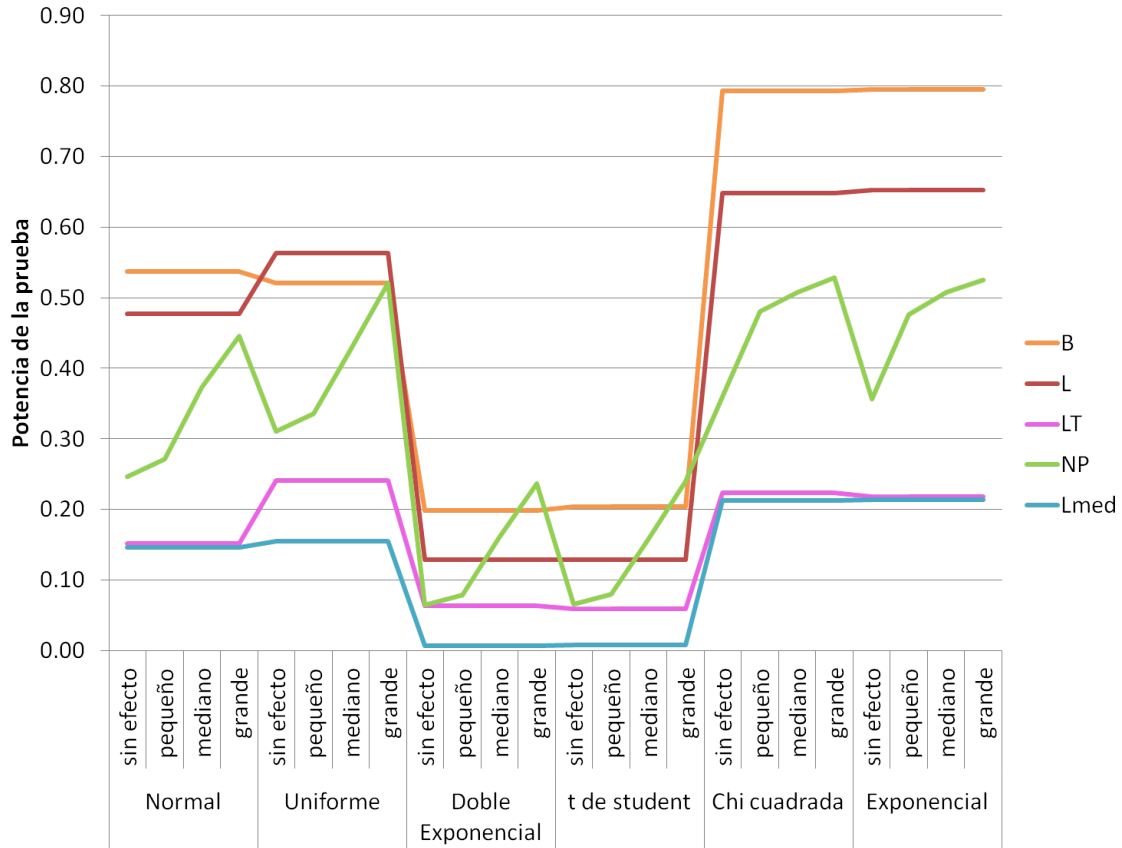
**Figura 5.5:** Potencia de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de prueba  $n=4$ .

LT alcanza valores de potencia superiores al 20 % para las distribuciones uniforme, chi-cuadrada y exponencial.

La figura 5.7, muestra los resultados para la potencia de la prueba cuando el tamaño de muestra es de 10, bajo esta condición la prueba LT, para distribuciones simétricas con colas pesadas presenta una potencia de 5.48 % superando a la prueba Lmed (esta prueba presenta potencias de 3.64 y 2.85 %). A pesar de que, para las distribuciones asimétricas la potencia de LT aumenta (41.35 y 40.24 %) no logra superar a la prueba Lmed (valores de potencia de 3.07 %).

Algo que se remarca en el caso de la prueba NP, es que al igual que con el tamaño de prueba, la potencia es muy dependiente del tamaño del efecto, por ejemplo, para el caso de la distribución chi-cuadrada, cuando no existe un efecto de tratamiento la potencia de la prueba es de 70.26 % (que es el caso que reportan en Nordstokke y

## 5.2. Potencia de la prueba



**Figura 5.6:** Potencia de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de prueba  $n=5$ .

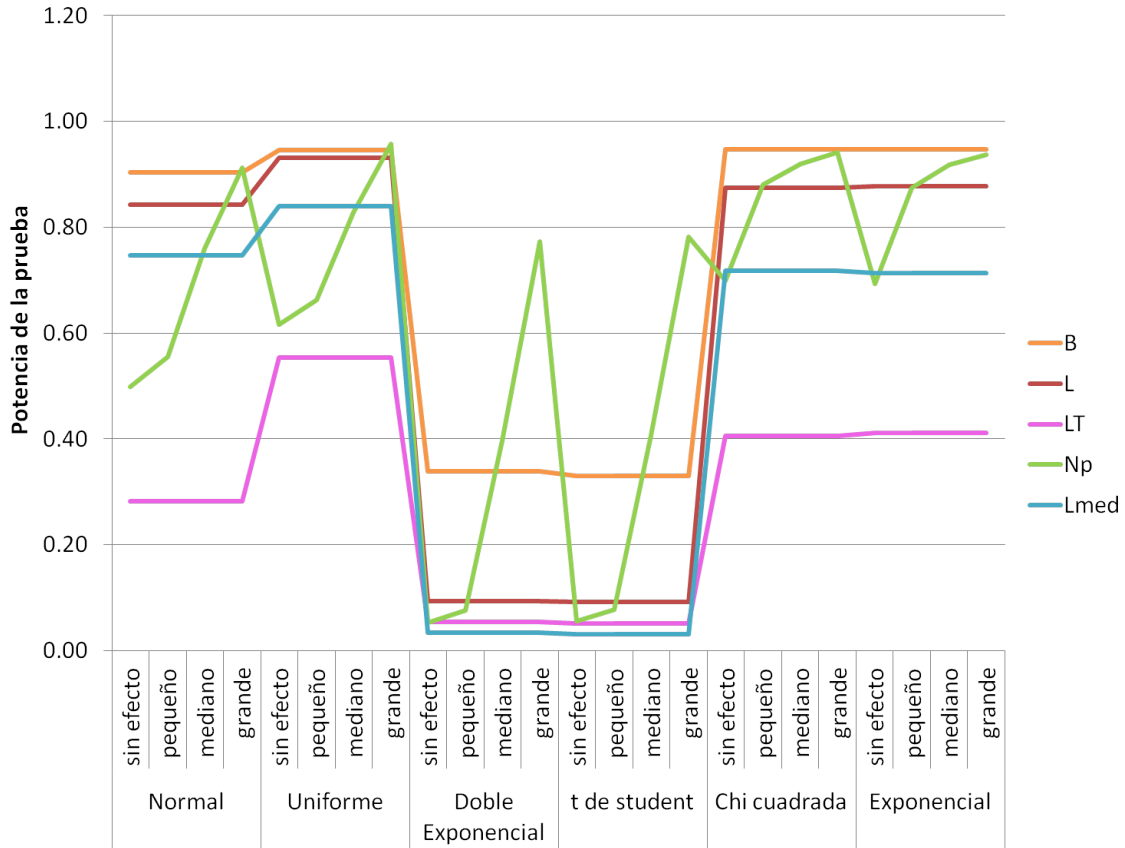
Zumbo (2010)), sin embargo, conforme aumenta el tamaño del efecto la potencia de la prueba tiende a aumentar, llegando a 94.05 % para el caso cuando el efecto es grande . Para el caso de las distribuciones asimétricas con colas pesadas, el comportamiento es similar, cuando aumenta el tamaño del efecto, también aumenta la potencia de la prueba (figura 5.7).

Cuando el tamaño de muestra aumenta a 20, las pruebas NP, B y L siguen teniendo un comportamiento similar al del caso cuando  $n=10$ , solo que las potencias de las pruebas han aumentado (figura 5.8).

La prueba LT, para distribuciones simétricas con colas pesadas presenta una potencia de 5.6 %, mientras que la prueba Lmed presenta una potencia de 4.54 %. Entre estas dos pruebas, la de mejor potencia para la distribución uniforme es la prueba Lmed, con una potencia arriba del 97 % (figura 5.8).



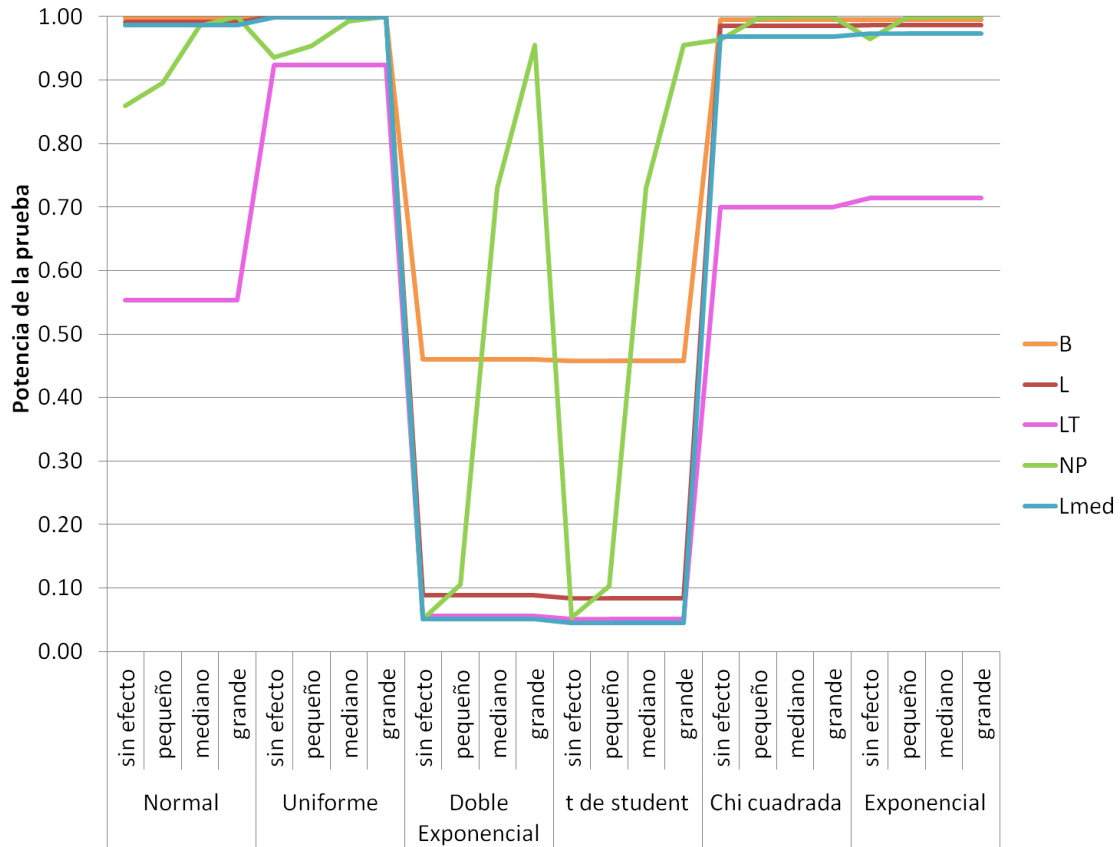
## 5.2. Potencia de la prueba



**Figura 5.7:** Potencia de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de prueba  $n=10$ .

En resumen, se observó que conforme aumenta el tamaño de muestra la potencia de las pruebas estudiadas también aumentaba, como en el caso de la prueba LT. Para el caso de las distribuciones sesgadas las pruebas que presentaron mejor potencia son la B y L (un resultado consistente con lo que reportan [Boos y C. \(1989\)](#)), sin embargo fueron las que obtuvieron un tamaño de prueba inaceptable en estas distribuciones. Al igual que en el tamaño de prueba la potencia de la NP depende mucho del tamaño de efecto en el modelo. La prueba LT supero en potencia a la prueba Lmed en el caso de la distribución doble exponencial con cualquier tamaño de muestra; con  $n=5$  se observó que la prueba LT es más potente que la Lmed en todos los casos, aquí es importante recordar que con este mismo tamaño de muestra la prueba Lmed obtuvo tamaños de prueba muy conservadores, inferiores a 0.0244.

## 5.2. Potencia de la prueba



**Figura 5.8:** Potencia de prueba para los estadísticos Bartlett (B), Levene (L), Levene Transformada (LT), No paramétrico (NP) y Levene con medianas (Lmed). Con un tamaño de prueba  $n=20$ .

# Capítulo 6

## Conclusiones

En esta sección se mencionan las conclusiones de la investigación con base en los resultados obtenidos para el tamaño y potencia de las pruebas en el estudio de simulación.

Sin duda, cuando los datos provienen de la distribución normal la mejor prueba para detectar heterogeneidad de varianzas es la Bartlett, pues en los resultados de simulación las tasas de error tipo I fueron aceptables en todos los tamaños de muestra. Sin embargo, cuando la distribución de los errores deja de ser normal, la prueba se vuelve muy liberal y falla en todos los tamaños de muestra.

Cuando los tamaños de muestra son pequeños, todas las pruebas fallan obteniendo tamaños de prueba superiores al nivel de significancia (excepto la prueba B cuando los datos son normales).

El estadístico NP, tiene un comportamiento muy errático, cuando no se presenta un tamaño de efecto en el modelo, la prueba parece comportarse bien, en cuanto a tamaño y potencia; sin embargo, al aumentar el tamaño de efecto la prueba tiende a presentar tasas de error tipo I muy liberales ocasionando con esto que la prueba falle y no sea recomendable para probar homogeneidad de varianzas. Incluso, se observó que el tamaño de la prueba NP se incrementaba conforme aumentaba el tamaño del efecto.

La prueba Lmed mostró un comportamiento demasiado conservador cuando el tamaño de la muestra es pequeño e impar ( $n=5$ ) con valores que oscilan entre 0.001 y 0.0244.

## 6. Conclusiones

---

Para muestras grandes ( $n=20$ ) prácticamente todas las pruebas tuvieron tamaños de prueba cercanos al nominal.

La prueba LT resultó ser una prueba consistente cuando los datos provienen de distribuciones simétricas y con colas pesadas, se observa que conforme el tamaño de muestra aumenta la tasa de error tipo I tiende a disminuir.

En cuanto a potencia, las mejores pruebas en distribuciones asimétricas, son la L y la B, sin embargo en tamaño de prueba resultaron ser las peores en estas condiciones, por lo que no son pruebas recomendables en este tipo de distribuciones.

En distribuciones con colas pesadas, la prueba LT presentó mayor potencia (aunque con valores bajos) que la prueba Lmed.

Para tamaños de muestra mediano y grande ( $n=10$  y  $n=20$  respectivamente) y distribuciones asimétricas la mejor prueba es la Lmed.

# Anexos

# Anexos A

## Tamaños de prueba

**Tabla A.1:** Tamaño estimado de prueba, n=4

Distribución	Efecto	B	L	Lmed	NP	LT
Normal	sin efecto	0.0487	0.1016	0.0609	0.0831	0.0906
	pequeño	0.0487	0.1016	0.0609	0.0837	0.0906
	mediano	0.0487	0.1016	0.0609	0.0843	0.0906
	grande	0.0487	0.1016	0.0609	0.0871	0.0906
Uniforme	sin efecto	0.0212	0.1182	0.0723	0.0869	0.0804
	pequeño	0.0212	0.1182	0.0723	0.0869	0.0804
	mediano	0.0212	0.1182	0.0723	0.0874	0.0804
	grande	0.0212	0.1182	0.0723	0.0868	0.0804
Doble Exponencial	sin efecto	0.1396	0.1446	0.0661	0.0945	0.0894
	pequeño	0.1396	0.1446	0.0661	0.0919	0.0894
	mediano	0.1396	0.1446	0.0661	0.0934	0.0894
	grande	0.1396	0.1446	0.0661	0.1087	0.0894
t de student	sin efecto	0.1362	0.1345	0.0633	0.0821	0.0866
	pequeño	0.1362	0.1345	0.0633	0.0840	0.0866
	mediano	0.1362	0.1345	0.0633	0.0904	0.0866
	grande	0.1362	0.1345	0.0633	0.0999	0.0866
Chi cuadrada	sin efecto	0.4648	0.4182	0.1228	0.0892	0.1318
	pequeño	0.4648	0.4182	0.1228	0.1349	0.1318
	mediano	0.4648	0.4182	0.1228	0.1697	0.1318
	grande	0.4648	0.4182	0.1228	0.1746	0.1318
Exponencial	sin efecto	0.4697	0.4193	0.1237	0.0894	0.1314
	pequeño	0.4697	0.4193	0.1237	0.1386	0.1314
	mediano	0.4697	0.4193	0.1237	0.1725	0.1314
	grande	0.4697	0.4193	0.1237	0.1738	0.1314

## A. Tamaños de prueba

---

**Tabla A.2:** Tamaño estimado de prueba, n=5

Distribución	Efecto	B	L	Lmed	NP	LT
Normal	sin efecto	0.0476	0.0869	0.0022	0.0669	0.0622
	pequeño	0.0476	0.0869	0.0022	0.0653	0.0622
	mediano	0.0476	0.0869	0.0022	0.0704	0.0622
	grande	0.0476	0.0869	0.0022	0.0754	0.0622
Uniforme	sin efecto	0.0151	0.0987	0.0010	0.0645	0.0438
	pequeño	0.0151	0.0987	0.0010	0.0655	0.0438
	mediano	0.0151	0.0987	0.0010	0.0648	0.0438
	grande	0.0151	0.0987	0.0010	0.0670	0.0438
Doble Exponencial	sin efecto	0.1719	0.1174	0.0066	0.0687	0.0637
	pequeño	0.1719	0.1174	0.0066	0.0692	0.0637
	mediano	0.1719	0.1174	0.0066	0.0781	0.0637
	grande	0.1719	0.1174	0.0066	0.0990	0.0637
t de student	sin efecto	0.1724	0.1144	0.0076	0.0669	0.0669
	pequeño	0.1724	0.1144	0.0076	0.0707	0.0669
	mediano	0.1724	0.1144	0.0076	0.0789	0.0669
	grande	0.1724	0.1144	0.0076	0.1025	0.0669
Chi cuadrada	sin efecto	0.5157	0.3823	0.0244	0.0703	0.1071
	pequeño	0.5157	0.3823	0.0244	0.1482	0.1071
	mediano	0.5157	0.3823	0.0244	0.2037	0.1071
	grande	0.5157	0.3823	0.0244	0.2114	0.1071
Exponencial	sin efecto	0.5207	0.3792	0.0222	0.0643	0.1084
	pequeño	0.5207	0.3792	0.0222	0.1484	0.1084
	mediano	0.5207	0.3792	0.0222	0.2027	0.1084
	grande	0.5207	0.3792	0.0222	0.2169	0.1084

## A. Tamaños de prueba

---

**Tabla A.3:** Tamaño estimado de prueba, n=10

Distribución	Efecto	B	L	Lmed	NP	LT
Normal	sin efecto	0.0496	0.0638	0.0300	0.0517	0.0564
	pequeño	0.0496	0.0638	0.0300	0.0477	0.0564
	mediano	0.0496	0.0638	0.0300	0.0543	0.0564
	grande	0.0496	0.0638	0.0300	0.0695	0.0564
Uniforme	sin efecto	0.0040	0.0661	0.0217	0.0507	0.0438
	pequeño	0.0040	0.0661	0.0217	0.0509	0.0438
	mediano	0.0040	0.0661	0.0217	0.0500	0.0438
	grande	0.0040	0.0661	0.0217	0.0543	0.0438
Doble Exponencial	sin efecto	0.2695	0.0813	0.0333	0.0504	0.0528
	pequeño	0.2695	0.0813	0.0333	0.0532	0.0528
	mediano	0.2695	0.0813	0.0333	0.0672	0.0528
	grande	0.2695	0.0813	0.0333	0.1131	0.0528
t de student	sin efecto	0.2696	0.0778	0.0329	0.0548	0.0516
	pequeño	0.2696	0.0778	0.0329	0.0567	0.0516
	mediano	0.2696	0.0778	0.0329	0.0732	0.0516
	grande	0.2696	0.0778	0.0329	0.1209	0.0516
Chi cuadrada	sin efecto	0.6256	0.3425	0.0490	0.0547	0.1477
	pequeño	0.6256	0.3425	0.0490	0.2768	0.1477
	mediano	0.6256	0.3425	0.0490	0.4188	0.1477
	grande	0.6256	0.3425	0.0490	0.4642	0.1477
Exponencial	sin efecto	0.6235	0.3417	0.0469	0.0492	0.1514
	pequeño	0.6235	0.3417	0.0469	0.2754	0.1514
	mediano	0.6235	0.3417	0.0469	0.4143	0.1514
	grande	0.6235	0.3417	0.0469	0.4642	0.1514



## A. Tamaños de prueba

---

**Tabla A.4:** Tamaño estimado de prueba, n=20

Distribución	Efecto	B	L	Lmed	NP	LT
Normal	sin efecto	0.0508	0.0550	0.0349	0.0484	0.0531
	pequeño	0.0508	0.0550	0.0349	0.0485	0.0531
	mediano	0.0508	0.0550	0.0349	0.0577	0.0531
	grande	0.0508	0.0550	0.0349	0.0979	0.0531
Uniforme	sin efecto	0.0018	0.0571	0.0255	0.0516	0.0469
	pequeño	0.0018	0.0571	0.0255	0.0485	0.0469
	mediano	0.0018	0.0571	0.0255	0.0460	0.0469
	grande	0.0018	0.0571	0.0255	0.0606	0.0469
Doble Exponencial	sin efecto	0.3499	0.0626	0.0369	0.0500	0.0522
	pequeño	0.3499	0.0626	0.0369	0.0522	0.0522
	mediano	0.3499	0.0626	0.0369	0.0781	0.0522
	grande	0.3499	0.0626	0.0369	0.1951	0.0522
t de student	sin efecto	0.3557	0.0597	0.0361	0.0509	0.0510
	pequeño	0.3557	0.0597	0.0361	0.0489	0.0510
	mediano	0.3557	0.0597	0.0361	0.0765	0.0510
	grande	0.3557	0.0597	0.0361	0.1995	0.0510
Chi cuadrada	sin efecto	0.6768	0.3144	0.0446	0.0456	0.2283
	pequeño	0.6768	0.3144	0.0446	0.6648	0.2283
	mediano	0.6768	0.3144	0.0446	0.9353	0.2283
	grande	0.6768	0.3144	0.0446	0.9810	0.2283
Exponencial	sin efecto	0.6795	0.3193	0.0442	0.0511	0.2209
	pequeño	0.6795	0.3193	0.0442	0.6540	0.2209
	mediano	0.6795	0.3193	0.0442	0.9326	0.2209
	grande	0.6795	0.3193	0.0442	0.9815	0.2209

# Anexos B

## Potencia de prueba

**Tabla B.1:** Potencia estimada, n=4

Distribución	Efecto	B	L	Lmed	NP	LT
Normal	sin efecto	0.4012	0.3855	0.1479	0.1964	0.2610
	pequeño	0.4012	0.3855	0.1479	0.2096	0.2610
	mediano	0.4012	0.3855	0.1479	0.2632	0.2610
	grande	0.4012	0.3855	0.1479	0.3027	0.2610
Uniforme	sin efecto	0.3565	0.4337	0.2038	0.2263	0.3158
	pequeño	0.3565	0.4337	0.2038	0.2403	0.3158
	mediano	0.3565	0.4337	0.2038	0.2943	0.3158
	grande	0.3565	0.4337	0.2038	0.3410	0.3158
Doble Exponencial	sin efecto	0.1700	0.1573	0.0869	0.0838	0.0633
	pequeño	0.1700	0.1573	0.0869	0.0920	0.0633
	mediano	0.1700	0.1573	0.0869	0.1451	0.0633
	grande	0.1700	0.1573	0.0869	0.1852	0.0633
t de student	sin efecto	0.1697	0.1561	0.0891	0.0842	0.0650
	pequeño	0.1697	0.1561	0.0891	0.0957	0.0650
	mediano	0.1697	0.1561	0.0891	0.1398	0.0650
	grande	0.1697	0.1561	0.0891	0.1882	0.0650
Chi cuadrada	sin efecto	0.7126	0.5908	0.2004	0.2550	0.3025
	pequeño	0.7126	0.5908	0.2004	0.3337	0.3025
	mediano	0.7126	0.5908	0.2004	0.3572	0.3025
	grande	0.7126	0.5908	0.2004	0.3621	0.3025
Exponencial	sin efecto	0.7071	0.5797	0.1966	0.2525	0.2980
	pequeño	0.7071	0.5797	0.1966	0.3285	0.2980
	mediano	0.7071	0.5797	0.1966	0.3538	0.2980
	grande	0.7071	0.5797	0.1966	0.3584	0.2980

## B. Potencia de prueba

---

**Tabla B.2:** Potencia estimada, n=5

Distribución	Efecto	B	L	Lmed	NP	LT
Normal	sin efecto	0.5368	0.4777	0.1512	0.2465	0.1467
	pequeño	0.5368	0.4777	0.1512	0.2710	0.1467
	mediano	0.5368	0.4777	0.1512	0.3727	0.1467
	grande	0.5368	0.4777	0.1512	0.4453	0.1467
Uniforme	sin efecto	0.5211	0.5636	0.2407	0.3111	0.1544
	pequeño	0.5211	0.5636	0.2407	0.3358	0.1544
	mediano	0.5211	0.5636	0.2407	0.4256	0.1544
	grande	0.5211	0.5636	0.2407	0.5213	0.1544
Doble Exponencial	sin efecto	0.1990	0.1293	0.0636	0.0648	0.0072
	pequeño	0.1990	0.1293	0.0636	0.0785	0.0072
	mediano	0.1990	0.1293	0.0636	0.1603	0.0072
	grande	0.1990	0.1293	0.0636	0.2370	0.0072
t de student	sin efecto	0.2035	0.1292	0.0586	0.0659	0.0076
	pequeño	0.2035	0.1292	0.0586	0.0801	0.0076
	mediano	0.2035	0.1292	0.0586	0.1576	0.0076
	grande	0.2035	0.1292	0.0586	0.2397	0.0076
Chi cuadrada	sin efecto	0.7928	0.6486	0.2236	0.3593	0.2122
	pequeño	0.7928	0.6486	0.2236	0.4808	0.2122
	mediano	0.7928	0.6486	0.2236	0.5072	0.2122
	grande	0.7928	0.6486	0.2236	0.5284	0.2122
Exponencial	sin efecto	0.7948	0.6527	0.2182	0.3568	0.2133
	pequeño	0.7948	0.6527	0.2182	0.4759	0.2133
	mediano	0.7948	0.6527	0.2182	0.5077	0.2133
	grande	0.7948	0.6527	0.2182	0.5251	0.2133

## B. Potencia de prueba

---

**Tabla B.3:** Potencia estimada, n=10

Distribución	Efecto	B	L	Lmed	NP	LT
Normal	sin efecto	0.9036	0.8425	0.2819	0.4986	0.7460
	pequeño	0.9036	0.8425	0.2819	0.5553	0.7460
	mediano	0.9036	0.8425	0.2819	0.7598	0.7460
	grande	0.9036	0.8425	0.2819	0.9119	0.7460
Uniforme	sin efecto	0.9456	0.9313	0.5532	0.6156	0.8397
	pequeño	0.9456	0.9313	0.5532	0.6627	0.8397
	mediano	0.9456	0.9313	0.5532	0.8281	0.8397
	grande	0.9456	0.9313	0.5532	0.9578	0.8397
Doble Exponencial	sin efecto	0.3386	0.0934	0.0548	0.0530	0.0331
	pequeño	0.3386	0.0934	0.0548	0.0756	0.0331
	mediano	0.3386	0.0934	0.0548	0.3973	0.0331
	grande	0.3386	0.0934	0.0548	0.7735	0.0331
t de student	sin efecto	0.3300	0.0920	0.0506	0.0563	0.0307
	pequeño	0.3300	0.0920	0.0506	0.0771	0.0307
	mediano	0.3300	0.0920	0.0506	0.4075	0.0307
	grande	0.3300	0.0920	0.0506	0.7822	0.0307
Chi cuadrada	sin efecto	0.9477	0.8742	0.4054	0.6982	0.7182
	pequeño	0.9477	0.8742	0.4054	0.8806	0.7182
	mediano	0.9477	0.8742	0.4054	0.9189	0.7182
	grande	0.9477	0.8742	0.4054	0.9405	0.7182
Exponencial	sin efecto	0.9468	0.8767	0.4108	0.6925	0.7135
	pequeño	0.9468	0.8767	0.4108	0.8742	0.7135
	mediano	0.9468	0.8767	0.4108	0.9180	0.7135
	grande	0.9468	0.8767	0.4108	0.9371	0.7135

## B. Potencia de prueba

---

**Tabla B.4:** Potencia estimada, n=20

Distribución	Efecto	B	L	Lmed	NP	LT
Normal	sin efecto	0.9969	0.9909	0.5538	0.8598	0.9859
	pequeño	0.9969	0.9909	0.5538	0.8958	0.9859
	mediano	0.9969	0.9909	0.5538	0.9847	0.9859
	grande	0.9969	0.9909	0.5538	0.9998	0.9859
Uniforme	sin efecto	0.9996	0.9994	0.9232	0.9356	0.9987
	pequeño	0.9996	0.9994	0.9232	0.9534	0.9987
	mediano	0.9996	0.9994	0.9232	0.9926	0.9987
	grande	0.9996	0.9994	0.9232	1.0000	0.9987
Doble Exponencial	sin efecto	0.4596	0.0881	0.0560	0.0512	0.0509
	pequeño	0.4596	0.0881	0.0560	0.1059	0.0509
	mediano	0.4596	0.0881	0.0560	0.7302	0.0509
	grande	0.4596	0.0881	0.0560	0.9548	0.0509
t de student	sin efecto	0.4577	0.0839	0.0514	0.0531	0.0454
	pequeño	0.4577	0.0839	0.0514	0.1026	0.0454
	mediano	0.4577	0.0839	0.0514	0.7295	0.0454
	grande	0.4577	0.0839	0.0514	0.9553	0.0454
Chi cuadrada	sin efecto	0.9954	0.9850	0.6998	0.9632	0.9688
	pequeño	0.9954	0.9850	0.6998	0.9976	0.9688
	mediano	0.9954	0.9850	0.6998	0.9993	0.9688
	grande	0.9954	0.9850	0.6998	0.9997	0.9688
Exponencial	sin efecto	0.9947	0.9869	0.7138	0.9649	0.9730
	pequeño	0.9947	0.9869	0.7138	0.9986	0.9730
	mediano	0.9947	0.9869	0.7138	0.9993	0.9730
	grande	0.9947	0.9869	0.7138	0.9998	0.9730

# Referencias

- Bartlett, M. S. (1937). Properties of Sufficiency and Statistical Tests. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 160, 268–282.
- Boos, D. y C., B. (1989). Bootstrap methods for testing homogeneity of variances. *Technometrics*, 31, 6982.
- Bradley, J. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 31, 144–152.
- Brown, M. B. y Forsythe, A. B. (1974). Robust tests for equality of variances. *Journal of the American Statistical Association*, 69, 364367.
- Carroll, R. J. y Schneider, H. (1985). A note of Levene’s test for equality of variances. *Statistics & Probability Letters*, 3, 191–194.
- Casella, R. L. y Berger, G. (1990). *Statistical Inference*.
- Conover, W. y Iman, R. L. (1981). Rank Transformations as a Bridge Between Parametric and Nonparametric Statistics. *The American Statistician*, 35, 3, 124–133.
- Conover, W., Johnson, M. y Johnson, M. (1981). A comparative-study of test for homogeneity of variances, with applications to the outer continental-shelf bidding data. *Technometrics*, 23, 4, 351–361.
- Gastwirth, J. L., Gel, Y. R. y Miao, W. (2009). The Impact of Levenes Test of Equality of Variances on Statistical Theory and Practice. *Statistical Science*, 24, 343360.
- Hartley, H. (1950). The Maximum F-Ratio as a Short-Cut Test for Heterogeneity of Variance. *Biometrika*, 37, 308312.
- Hines, W. G. S. y O’Hara Hines, R. J. (2000). Increased Power with Modified Forms of the Levene (Med) Test for Heterogeneity of Variance. *Biometrics*, 56, 451–454.

## Referencias

---

- Keyes, T. K. y Levy, M. S. (1997). Analysis of Levene's Test Under Design Imbalance. *Journal of Educational and Behavioral Statistics*, 22, 227–236.
- Levene, H. (1960). Robust test for equality of variances. *Contributions to Probability and Statistics*, 278–292.
- Lim, T.-S. y Loh, W.-Y. (1996). A comparison of tests of equality of variances. *Computational Statistics & Data Analysis*, 22, 287–301.
- Loh, W. Y. (1987). Some Modifications of Levene's Test of Variance Homogeneity. *Journal of the Statistical Computation and Similarities*, 28.
- Mendes, M. y Özcaya Turhan, N. (2006). A new alternative in testing for homogeneity of variances. *Journal of Statistical Research*, 40, 65–83.
- Milliken, G. A. y Johnson, D. E. (2002). *Analysis of Messy Data*. CRC Press Taylor & Francis Group.
- Nordstokke, D. W. y Zumbo, B. D. (2010). A New Nonparametric Levene Test for Equal Variances. *Psicológica*, 31, 401–430.
- O'Brien, R. (1978). Robust techniques for testing heterogeneity of variance effects in factorial designs. *Psychometrika*, 43, 327342.
- O'Brien, R. (1981). A simple test for variance effects in experimental designs. *Psychol. Bull.*, 89, 570574.
- Parra-Frutos, I. (2009). The behaviour of the modified Levenes test when data are not normally distributed. *Comput Stat*, 24, 671693.
- Parra-Frutos, I. (2012). Testing homogeneity of variances with unequal sample sizes. *Computational Statistics*, 1–29.
- Sharma, D. y Golam-Kibria, B. M. (2012). On some test statistics for testing homogeneity of variances: a comparative study. *Journal of Statistical Computation and Simulation*, 1–20.
- Shoemaker, L. H. (2003). Fixing the F Test for Equal Variances. *The American Statistician*, 52, 105–114.
- Yitnosumarto, S. y O'Neill, M. E. (1986). On Levene's tests of variance homogeneity. *Australian Journal of Statistics*, 28, 230–241.