



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN
EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E
INFORMÁTICA ESTADÍSTICA

**Prueba de homogeneidad de varianzas para muestras
normales censuradas**

Violeta De La Huerta Contreras

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA
OBTENER EL GRADO DE:

MAESTRA EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO
2012

La presente tesis titulada: **Prueba de homogeneidad de varianzas para muestras normales censuradas**, realizada por la alumna: **Violeta De La Huerta Contreras**, 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. Humberto Vaquera Huerta

ASESOR



Dr. Gilberto Rendon Sánchez

ASESOR



Dr. Ignacio Méndez Ramírez

ASESOR



Dra. Martha Elva Ramírez Guzmán

ASESOR



Dr. Alfonso Hernández Garay

Prueba de homogeneidad de varianzas para muestras normales censuradas

Violeta De La Huerta Contreras

Colegio de Postgraduados, 2011

El ANOVA es ampliamente usado en diversos campos de la investigación, es por ello que verificar el cumplimiento de los supuestos es muy importante; especialmente el de homogeneidad de varianzas. Actualmente existen una buena cantidad de pruebas para éste último, sin embargo se tiene una limitante en el ámbito práctico cuando la muestra es censurada. El objetivo de este trabajo fue investigar el comportamiento de la prueba de Bartlett bajo censura por la derecha del tipo II. Se utiliza el método de máxima verosimilitud para estimar las varianzas, se investiga el comportamiento asintótico (método Delta) del estadístico de prueba bajo censura y por medio del método Monte Carlo a 10,000 repeticiones (donde tanto el tamaño y censura de n_1 sea $=$ ó \neq a n_2), para obtener la distribución del estadístico de Bartlett bajo censura. Encontramos que contar con esta metodología es útil ya que conforme aumenta la censura en las muestras, la media y varianza de la distribución del estadístico de Bartlett aumenta considerablemente. Las tablas de valores de la distribución fueron generadas así como el tamaño, potencia y robustez de la prueba. Se generó también la metodología para la prueba de Levene para muestras censuradas por la derecha tipo I. Finalmente se ejemplifica el uso de ambas pruebas en casos de reproducción ovina.

Palabras clave: ANOVA; Censura por la derecha; Método Delta; Prueba de Bartlett; Prueba de Levene; Reproducción animal.

A variance homogeneity test in censored normal samples

Violeta De La Huerta Contreras

Colegio de Postgraduados, 2011

A variance homogeneity test for right censored type II sample is proposed. The test is based on Bartlett's statistic. By using simulation the critical values of the null distribution of Bartlett's statistic for testing homogeneity of variances of two normal populations is obtained when the sample sizes and censoring levels are not equal. Also, we investigated the properties of the proposed test (size, power and robustness). The asymptotic distribution of the test statistic is also obtained. Results shows that the distribution of the test statistic distribution depends on censure level. The test keeps the size at different significance levels, the test is consistent. Furthermore, a methodology for Levene's test under type I right censure is proposed. Finally, some example of application in animal reproduction research is also given for both Bartlett's and Levene's test under censure.

Key words: ANOVA; Bartlett's test; Delta method; Levene's test; Right censure type II; Animal reproduction.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la buena administración y canalización de los recursos destinados para el crecimiento profesional de nuestro México.

Al Colegio de Postgraduados, por su entusiasta espíritu de superación a través de su dirección, su brillante academia y por supuesto sus valiosos recursos humanos que coadyuvan a lograr nuestro objetivo.

Al Programa de Estadística: por la paciente formación brindada de cada uno de sus doctores, por sus secretarías que hacen grata nuestra estancia, y al personal de servicios que su labor es darnos un espacio propio para el estudio.

A los integrantes de mi Consejo de investigación:

De forma muy especial a mi Consejero el **Dr. Humberto Vaquera Huerta** por su gentil, paciente y atinado acompañamiento en todo este proceso.

Al **Dr. Gilberto Rendón Sánchez** por las buenas bases cimentadas en su tiempo y su cordial atención durante el logro de esta meta.

Al **Dr. Ignacio Méndez Ramírez** y **Ph.D. Barry C. Arnold** por galardonar este trabajo.

Finalmente y sin lugar a dudas a la **Dra. Martha Elva Ramírez Guzmán** y el **Dr. Alfonso Hernández Garay** por sus atinados consejos; así como al **Dr. Jaime Gallegos Sánchez** y su equipo de estudiantes en la aportación de datos para ilustrar este trabajo y sobretodo su impecable revisión.

Y a ti que siempre estás presente...

DEDICATORIA

A **DIOS** y todos sus ángeles en el camino.

A mis **padres** y **hermanas** por que son un regalo del cielo.

A mis geniales compañeros de clase y un "adoptado" así como sin lugar a dudas a todos mis compañeros de casa.

A ti **Irma, Said, José, Edgar y Julio** por su energía de activación así como a mis inigualables amigos **Ara, Cuba, Violeta, Jazmín, Fiordalizo**: la leen ehh!!

Y a ti **amor mío**, coprotagonista de mi vida, mi **A⁻¹** que juntos somos la **I**.

Índice

1. Introducción	1
2. Objetivos	3
3. Marco teórico	4
3.1. Pruebas para homogeneidad de varianzas	4
3.1.1. Prueba de Bartlett	5
3.1.2. Prueba de Levene	7
3.2. Datos Censurados	8
3.2.1. Censura por la derecha Tipo I	9
3.2.2. Censura por la derecha aleatoria independiente	10
3.2.3. Censura por la derecha Tipo II	11
3.3. Métodos asintóticos	12
3.3.1. Método Delta	12
4. Metodología para probar homocedasticidad bajo censura usando el estadístico de Bartlett	14
4.1. Estimación de la varianza en muestras normales censuradas	14
4.1.1. Obtención de estimadores de Máxima Verosimilitud	15

Índice

4.2. Uso de la prueba de Bartlett bajo censura	18
4.2.1. Propiedad de invarianza del estadístico	18
4.3. Distribución del estadístico de Bartlett para dos muestras censuradas	21
5. Distribución para muestras grandes del estadístico de Bartlett modificado h	29
5.0.1. Varianza asintótica	29
5.0.2. Distribución para muestras grandes	31
6. Evaluación del método propuesto con la prueba de Bartlett	36
6.1. Potencia de la prueba	36
6.2. Tamaño de la prueba	38
6.3. Determinación de la robustez de la prueba	39
7. Metodología para determinar homogeneidad de varianzas en muestras censuradas empleando la prueba de Levene	44
7.1. Estimación de los residuales de muestras censuradas por la derecha .	44
7.1.1. Estimadores de Buckley-James	45
7.1.2. Estimación de los residuales bajo censura	47
7.2. Distribución del estadístico de Levene bajo censura	49
8. Ejemplos de aplicación	50
8.1. Efecto macho en la activación del estro usando la prueba de Bartlett .	51
8.2. Efecto de neurotransmisores en la reproducción ovina usando la prueba de Levene	53
9. Conclusiones	55

Índice

10. Apéndice	56
Referencias	64

Índice de tablas

4.1. Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 = n_2$ y $c_1 = c_2$	23
4.2. Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 = n_2$ y $c_1 \neq c_2$	24
4.3. Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 \neq n_2$ y $c_1 \neq c_2$	26
6.1. Estimaciones del tamaño de la prueba	39
6.2. Valores de σ^2 empleadas en las distribuciones Uniforme, Exponencial, Lognormal y Normal	40
6.3. Tamaños empíricos y potencia ($\alpha = 0.05$) en las distribuciones Uniforme, Exponencial, Lognormal y Normal	40
6.4. Valores de σ^2 empleadas en las distribuciones Logística, t student, Laplace y Normal	42
6.5. Tamaños empíricos y potencia ($\alpha = 0.05$) en las distribuciones Logística, t student, Laplace y Normal	42
8.1. Datos del efecto macho en la entrada al estro en ovejas.	52
8.2. Valores generados en el ejemplo del efecto macho en la reproducción ovina.	53
8.3. Datos del efecto del glutamato en reproducción ovina.	54

Índice de figuras

4.1. Eficiencia de los estimadores lineales y de máxima verosimilitud de la distribución Normal para valores k/n	17
4.2. Distribuciones empíricas estimadas del estadístico de Bartlett bajo censura de la distribución Normal	21
4.3. Gráfico comparativo de la distribución del estadístico de Bartlett para dos muestras aleatorias censuradas de tamaño $n_1 = n_2 = 30$ a diferentes niveles de censura.	22
6.1. Potencia del estadístico de Bartlett bajo censura generada con 10,000 pares de muestras de una población con distribución normal a diferentes tamaños de muestra y niveles de censura especificados en la leyenda.	38
7.1. Gráfico comparativo de la distribución del estadístico de Levene para dos muestras aleatorias censuradas de tamaño $n_1 = n_2 = 50$ a diferentes niveles de censura.	49

Capítulo 1

Introducción

El Análisis de Varianza (ANOVA), es un método estadístico ampliamente usado en la investigación científica. Uno de los supuestos del ANOVA es que las varianzas de las diferentes poblaciones deben ser iguales (homogeneidad de varianzas). A lo largo de la historia muchas pruebas de homocedasticidad han sido propuestas por diversos autores, entre ellas: [Neyman y Pearson \(1931\)](#), [Bartlett \(1937\)](#), [Cochran \(1941\)](#), [Hartley \(1950\)](#), [Box \(1953\)](#), [Cadwell \(1953\)](#), [Levene \(1960\)](#), [Brown y Forsythe \(1974\)](#), y recientemente [Bhandary y Dai \(2009\)](#). Entre éstas, la prueba de Bartlett es parte de varios programas de cómputo estadísticos. La prueba de Bartlett se usa en: ciencias médicas ([Goto *et al.* \(2011\)](#), [Krasny *et al.* \(2011\)](#), [Altun *et al.* \(2011\)](#)); veterinaria y zootecnia ([Yilmaz *et al.* \(2011\)](#), [Powers y Powers \(1990\)](#)); genética ([Struchalin *et al.* \(2010\)](#); [Calsbeek y Goodnight \(2009\)](#)); ingeniería ([Ho y Romero \(2011\)](#)); biología ambiental ([Kobayashi *et al.* \(2011\)](#)); y climatología ([Dasgupta y De \(2010\)](#)).

La prueba propuesta por [Bartlett \(1937\)](#) se basa en un estadístico de razón de verosimilitud la cual es una ligera modificación a la prueba desarrollada por Neyman y Pearson (N&P). Aunque tanto la de Bartlett como la de N&P son consistentes contra cualquier alternativa, la prueba de Bartlett es preferible porque donde sea es insesgada para cualquier tamaño de muestra. En términos de potencia, esta prueba es considerada por ser el mejor resultado posible. No óptima, como la uniformemente insesgada más potente de las pruebas que existen para el problema de la homogeneidad de varianzas, sin embargo, estudios hechos por simulación Monte Carlo han demostrado la superioridad en potencia de la prueba de Bartlett (asumiendo normalidad) relativa a pruebas competidoras por [Cochran \(1941\)](#), [Hartley \(1950\)](#), [Cadwell \(1953\)](#), [Box \(1953\)](#), y el método jackknife. [E. \(2006\)](#).

Todas las pruebas anteriormente mencionadas, están diseñadas para muestras completas de una población con distribución Normal. Sin embargo, hay circunstancias en que los datos generados en algunas investigaciones presentan observaciones censuradas. Al respecto, varios autores han investigado sobre la censura por la derecha tipo

1. Introducción

II en muestras provenientes de una distribución Normal en diferentes aspectos. En el tema de estimación de parámetros de poblaciones normales bajo censura Gupta (1952) propuso los estimadores de Máxima Verosimilitud para la media y la desviación estándar, Sarhan y Greenberg (1955) estimó los parámetros de localidad y escala a través de las estadísticas de orden, Ali y Chan (1967) propuso los mejores estimadores lineales asintóticos normales, Raqab (1997) trabajó en los predictores futuros de máxima verosimilitud modificados de estadísticas de orden. Otros autores investigaron sobre estimación en regresión lineal en muestras normales censuradas, por ejemplo Buckley y James (1979), posteriormente Lai y Ying (1991) propuso los estimadores de Buckley-James modificados para muestras grandes. Klein *et al.* (1999) trabajó en modelar efectos aleatorios para datos censurados por un modelo de regresión y Linder y Nagaraja (2003) en modelos bivariados. Sin embargo, Stare *et al.* (2000) indica que hasta el momento no hay buenos procedimientos para verificar el supuesto de homocedasticidad del modelo bajo censura.

Por lo que es importante el desarrollar una metodología para probar homogeneidad de varianzas bajo censura. En el presente trabajo se presenta una metodología para dos muestras normales bajo diferentes niveles de censura a diferentes tamaños de muestra obteniendo las tablas de la distribución generada a diferentes valores de significancia en el apartado 4.3. Además, la propiedad de invarianza de los parámetros en el punto 4.2.1 y el resultado teórico para muestras grandes se describen en el capítulo 5. Más adelante en el capítulo 6 se obtuvieron las propiedades de la prueba donde, la potencia muestra ser consistente y el tamaño indica que es exacta. Adicionalmente en el capítulo 7 se empleó la prueba de Levene, la censura considerada es por la derecha tipo I. El propósito de desarrollar esta prueba es plantear los fundamentos teóricos para calcular los residuales censurados y el procedimiento por simulación Monte Carlo para encontrar la distribución del estadístico de Levene, de principio para dos muestras, cuando éstas se encuentran censuradas. Finalmente, en el capítulo 8 se describen ejemplos reales en la ciencia animal sobre reproducción ovina en tratamientos para el inicio del estro aplicando ambas metodologías.

Capítulo 2

Objetivos

General

Desarrollar procedimientos estadísticos para probar homogeneidad de varianzas en dos poblaciones normales independientes bajo censura por la derecha.

Particulares

1. Obtener las tablas de valores críticos para la prueba Bartlett bajo censura.
2. Investigar el comportamiento asintótico de la prueba de Bartlett bajo censura.
3. Evaluar el tamaño, la potencia y robusticidad de la prueba para Bartlett.
4. Implementar la prueba de Levene bajo censura.
5. Mostrar ejemplos de aplicación para la prueba de Bartlett y Levene.
6. Generar programas de computo para implementar las pruebas estudiadas.

Capítulo 3

Marco teórico

3.1. Pruebas para homogeneidad de varianzas

Para verificar la igualdad de varianzas ($H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$), [Conover *et al.* \(1981\)](#) menciona hasta ese momento la existencia de cincuenta y seis pruebas para la igualdad de varianzas, la mayoría de las cuales son variaciones de las más populares y más usadas tanto de las paramétricas como de las no paramétricas pruebas disponibles para probar las k varianzas ($k \geq 2$). Las pruebas paramétricas las clasifica en aquellas que son esencialmente modificaciones a la prueba de razón de verosimilitud o que de otro modo trabajan bajo el supuesto de normalidad (grupo 1), las que son modificaciones a éstas empleando un estimador de la kurtosis (grupo 2) y las que son modificaciones a la prueba F para medias (grupo 3). Entre las pruebas no paramétricas se incluyen las de clase de rango lineal, la modificación consiste en usar la media o la mediana muestral en lugar de la media poblacional cuando se calcula el estadístico de prueba (grupo 4).

- Grupo 1: Neyman-Pearson (1931), Bartlett (1937), Cochran (1941), Bartlett y Kendall (1946), F-max por Hartley (1950), Cadwell (1953), Dixon y Massey (1969), Samuiddin (1976), Gartside (1972) y Lehmann (1959).
- Grupo 2: Box (1953) y Scheffé(1959).

3.1. Pruebas para homogeneidad de varianzas

- Grupo 3: Levene (1960) posteriormente tres modificaciones una de ellas por Brown y Forsythe (1974) y Miller (1969).
- Grupo 4: Mood (1954), Freund-Ansay-Bradley (1957-1960), Barton y David (1958), Siegel y Tukey (1960), Capon (1961), Klotz (1962), Fligne y Killen (1976), Talwar y Gentle (1977) y Shorak y et al. (1965-1978).

Las pruebas citadas más frecuentes y usadas son los métodos propuestos por Bartlett (1937), Hartley (1950), Levene (1960), y Box (1953) [Bhandary y Dai \(2009\)](#).

3.1.1. Prueba de Bartlett

Bartlett, al estudiar las propiedades de suficiencia para toda prueba de significancia en muestras pequeñas desarrolló la idea de variación condicional y demostró su comportamiento en pruebas comunes dependiendo ya sea si son de variación continua o discontinua. En particular, reexaminó el uso de la χ^2 y otros criterios de similares; de esto propone una nueva aplicación de la χ^2 para probar la homogeneidad de una serie de varianzas [Bartlett \(1937\)](#). Esta prueba es frecuentemente llamada como la prueba de Bartlett o la prueba de Neyman-Pearson-Bartlett [E. \(2006\)](#) ya que Bartlett modificó la prueba Neyman-Pearson para corregir el sesgo la prueba resultante es probablemente la más usada [Conover *et al.* \(1981\)](#). Esta prueba también tiene aplicaciones como: verificar las interacciones de alto orden en experimentos factoriales y en el análisis de series de tiempo para investigar fluctuaciones en periodogramas lineales. [E. \(2006\)](#).

Específicamente, la prueba es la siguiente [Bhandary y Dai \(2009\)](#):

Para $t \geq 2$ muestras de tamaño n_i ; $n = 1, 2 \dots$

Sean:

3.1. Pruebas para homogeneidad de varianzas

$$\gamma_i = n_i - 1; i = 1, 2, \dots, t; \gamma = \sum_{i=1}^t \gamma_i$$

Definir:

$$\hat{\sigma}^2 = \sum_{i=1}^t \frac{\gamma_i \hat{\sigma}_i^2}{\gamma}$$

Sea:

$$c = 1 + \frac{1}{3(t-1)} \left(\sum_{i=1}^t \frac{1}{\gamma_i} - \frac{1}{\gamma} \right)$$

y

$$U = \frac{1}{c} \log \frac{(\hat{\sigma}^2)^\gamma}{\prod_{i=1}^t (\hat{\sigma}_i^2)^{\gamma_i}} \quad (3.1)$$

De modo que rechazar H_0 si $U \geq c$, la constante c es determinada del hecho que $U \sim \chi_{t-1}^2$. Por lo que la región de rechazo para la prueba es: rechazar H_0 si $U \geq \chi_{\alpha, t-1}^2$. Donde $\chi_{\alpha, t-1}^2$ es el punto superior α % de la distribución χ^2 con $t - 1$ grados de libertad. [Bhandary y Dai \(2009\)](#).

Las propiedades del estadístico generado por Bartlett se describen a continuación [E. \(2006\)](#):

Sea $L_1 = \prod_{i=1}^t (\hat{\sigma}_i^2)^{n_i/n} / \sum_{i=1}^k (n_i/n) \hat{\sigma}_i^2$ el estadístico de razón de verosimilitud propuesto por Neyman y Pearson del cual Bartlett hace una suave modificación obteniendo el estadístico $L = \prod_{i=1}^t (S_i^2)^{n_i/n} / \sum_{i=1}^t (n_i/n) S_i^2$;

donde $n = \sum_{i=1}^t n_i$ y $n_i = n_i - 1$. La región crítica para la prueba es $0 < L < A$, donde A es determinada por el tamaño de la prueba. En particular, el tamaño α como valor crítico Bartlett, A , es determinado por la ecuación $P_{H_0} 0 < L < A = \alpha$ [E. \(2006\)](#). Por lo que:

3.1. Pruebas para homogeneidad de varianzas

1. Aunque tanto L como L_1 son pruebas consistentes contra todas las alternativas, la prueba de Bartlett es preferible por que la prueba L_1 es sesgada a menos que los tamaños de muestra sean iguales, en cambio la prueba L es insesgada para cualquier tamaño de muestra.
2. En términos de potencia, la prueba de Bartlett es considerada a ser el mejor resultado posible. No óptimo, tal como la uniformemente insesgada más potente. Sin embargo, estudios por el método Monte Carlo han demostrado la superioridad en potencia de esta prueba (asumiendo normalidad) relativa a tales competidores como las pruebas de Cochran, Hartley, Cadwell, Box y el método jackknife.
3. La mayor objeción de usar la prueba de Bartlett es su sensibilidad al supuesto de normalidad. Si las poblaciones no son aproximadamente normalmente distribuidas, sin tomar en cuenta el tamaño de muestra, el nivel de significancia real del procedimiento de Bartlett puede diferir enormemente del nivel de significancia nominal. En particular, el grado de disparidad depende de la kurtosis de la población.

3.1.2. Prueba de Levene

La prueba estándar para verificar la homogeneidad de varianzas es la de Bartlett, la cual es una efectiva herramienta sólo si la población es aproximadamente normalmente distribuida. Cuando el supuesto de normalidad es violado en un tamaño tantas veces más grande que el nivel de significancia nominal, un procedimiento que es relativamente insensible a partidas de normalidad es la prueba de Levene [E. \(2006\)](#). Ésta es una prueba de F de un análisis de varianza de una vía basada en [\(3.2\)](#) la expresión es descrita por [Brown y Forsythe \(1974\)](#).

$$z_{ij} = |x_{ij} - \bar{x}_i| \quad (3.2)$$

3.2. Datos Censurados

Donde z_{ij} son los valores absolutos de los residuales.

$$W_0 = \frac{\sum_{i=1}^t n_i (\bar{z}_i - \bar{z}_{..})^2 / (t-1)}{\sum_{i=1}^t \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2 / \sum_i n_i - 1} \quad (3.3)$$

Donde:

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

La regla de decisión es rechazar H_0 si $W_0 > F_{(I-1)/(N-I)}^{(I-1)}$.

3.2. Datos Censurados

A modo de ejemplo, suponga que se desea estimar el promedio de vida de las lámparas eléctricas producidas en una fábrica. El método clásico sería tomar un cierto número de lámparas de forma aleatoria y tenerlas prendidas hasta que se fundan a fin de obtener los datos requeridos para el análisis. De otro modo, en lugar de gastar las lámparas, sería decidido parar el experimento cuando un número fijado de lámparas se hayan fundido. La muestra obtenida de ese modo sería una muestra censurada tipo II. Por otro lado, en lugar de detener el experimento después de que un número fijado haya fallado, una decisión podría ser terminar el experimento después de 1,000 horas; la muestra en este caso sería censurada tipo I. [Gupta \(1952\)](#).

Suponer que el tiempo de vida de n individuos es representado por las variables aleatorias T_1, \dots, T_n . Sin embargo, en lugar de los valores observados para cada tiempo de vida, se tiene un tiempo t_i que puede ser tiempo de vida o tiempo censurado. Sea $\delta_i = I(T_i = t_i)$ una variable aleatoria llamada censura o indicadora del *status* de t_i donde si $T_i = t_i = 1$ y $T_i > t_i = 0$. Por lo que nos indica si t_i es un tiempo de vida observado ($\delta_i = 1$) o es un tiempo censurado ($\delta_i = 0$). De esta

3.2. Datos Censurados

forma los datos observados consistirán en $(t_i, \delta_i); i = 1, \dots, n$. Donde $t_i = \min(T_i, C_i)$, $\delta_i = I(T_i \leq C_i)$. Lawless (2003).

3.2.1. Censura por la derecha Tipo I

El tipo I de censura se presenta cuando un estudio es conducido sobre un periodo específico de tiempo. Este mecanismo se aplica cuando cada individuo tiene un tiempo potencial fijado de censura $C_i > 0$ tal que T_i es observada si $T_i \leq C_i$; de otro modo solo sabemos que $T_i > C_i$. El tipo I frecuentemente toma lugar cuando un estudio es conducido sobre un periodo de tiempo específico. La función de verosimilitud para una muestra censurada de este tipo es basada en la distribución de probabilidad de (t_i, δ_i) , $i = 1, \dots, n$. Tanto t_i y δ_i son variables aleatorias y su p.d.f. es:

$$f(t_i)^{\delta_i} Pr(T_i > C_i)^{1-\delta_i} \quad (3.4)$$

Nota que C_i son constantes fijadas y que t_i puede tomar valores i con

$$\begin{aligned} Pr(t_i = C_i, \delta_i = 0) &= Pr(T_i > C_i) \\ Pr(t_i, \delta_i = 1) &= f(t_i); t_i \leq C_i \end{aligned}$$

Donde Pr en la segunda expresión denota ya sea una p.d.f. o la función de masa de probabilidad de acuerdo a si la distribución de T_i es continua o discreta en t_i . Asumiendo que los tiempos de vida de T_1, \dots, T_n son estadísticamente independientes, la función de verosimilitud de (3.4) es:

$$L = \prod_{i=1}^n f(t_i)^{\delta_i} S(t_i+)^{1-\delta_i} \quad (3.5)$$

3.2. Datos Censurados

El término $S(t_i+)$ aparece en (3.5) porque equivale a $Pr(T_i > t_i)$ en general; si $S(t)$ es continua en t_i , entonces $S(t_i+) = S(t - i)$. Lawless (2003).

3.2.2. Censura por la derecha aleatoria independiente

Un proceso de censura aleatoria es uno en el cual cada individuo supone tener un tiempo de vida T y un tiempo de censura C , donde T y C son variables aleatorias continuas independientes, con funciones de supervivencia $S(t)$ y $G(t)$ respectivamente. Todos los tiempos de vida y censura son asumidos mutuamente independientes y se supone que $G(t)$ no depende de ninguno de los parámetros de $S(t)$. Como en el tipo I, donde $t_i = \min(T_i, C_i)$ y $\delta_i = 1$ si $T_i \leq C_i$ y $\delta_i = 0$ si $T_i > C_i$. Los datos observados en n individuos suponen parejas (t_i, δ_i) , $i = 1, \dots, n$. La f.d.p. de (t_i, δ_i) se obtiene partiendo de que si $f(t)$ y $g(t)$ son las f.d.p. de T_i y C_i entonces:

$$Pr(t_i = t, \delta_i = 0) = Pr(C_i = t, T_i > C_i) = g(t)S(t)$$

$$Pr(t_i = t, \delta_i = 1) = Pr(T_i = t, T_i \leq C_i) f(t)G(t)$$

Esto puede ser combinado en la expresión siguiente:

$$Pr(t_i = t, \delta_i) = [f(t)G(t)]^{\delta_i} [g(t)S(t)]^{1-\delta_i}$$

y así la función de distribución de (t_i, δ_i) , $i = 1, \dots, n$ es

$$\prod_{i=1}^n [f(t_i)G(t_i)]^{\delta_i} [g(t_i)S(t_i)]^{1-\delta_i}$$

Dado que $G(t)$ y $g(t)$ no involucran ninguno de los parámetros en $f(t)$, pueden ser eliminados y la función de verosimilitud es:

$$L = \prod_{i=1}^n f(t_i)^{\delta_i} S(t_i)^{1-\delta_i}$$

Esta función resulta ser la misma para el Tipo I. Aunque el modelo de variables aleatorias censuradas es razonable, en muchas situaciones el proceso de censura es ligado a fallas en el tiempo de proceso. Supongamos, por ejemplo, que la fecha de termino para un ensayo médico no se fijó antes de que el estudio iniciara, pero se elige después, con la elección influenciada por los resultados del estudio sobre el tiempo. En tales circunstancias sería difícil escribir un modelo que represente el proceso completo; en estas circunstancias esta función resulta útil.

3.2.3. Censura por la derecha Tipo II

Se refiere a la situación donde sólo k tiempos de vida más pequeños $t_{(1)} \leq \dots \leq t_{(k)}$ en una muestra aleatoria de n observaciones. Aquí k es un entero específico entre 1 y n . Este esquema de censura tiene lugar cuando n individuos inician un estudio al mismo tiempo, el estudio termina una vez que k fallas o tiempos de vida hayan sido observados. El valor de k es elegido antes que los datos sean colectados. Para distribuciones continuas el tiempo de vida más pequeño r se denota como $T_{(1)} < T_{(2)} < \dots < T_{(r)}$. Si T_i tiene f.d.p. $f(t)$ y función de supervivencia $S(t)$, entonces la f.d.p. conjunta de $T_{(1)} < T_{(2)} < \dots < T_{(r)}$ es (3.6). Lawless (2003):

$$\frac{n!}{(n-k)!} \left\{ \prod_{i=1}^k f(t_{(i)}) \right\} S(t_{(k)})^{n-k} \quad (3.6)$$

La función de verosimilitud es basada en (3.6). Cayendo la constante $\frac{n!}{(n-k)!}$ y notando que en términos de la notación (δ_i, t_i) tenemos $\delta_i = 0$ y $t_i = t_{(r)}$ para aquellos individuos que son censurados, podemos ver entonces que (3.6) da una verosimilitud de la misma forma que (??) como para la censura tipo I. Lawless (2003).

3.3. Métodos asintóticos

3.3.1. Método Delta

El método Delta consiste en usar la expansión de Taylor para aproximar un vector de la forma $\phi(T_n)$ por la polinomial $\phi(\theta)(T_n - \theta) + \dots$ en $T_n - \theta$. Es un simple pero útil método para deducir la ley del límite de $\phi(T_n) - \phi(\theta)$ proveniente éste de $T_n - \theta$. Las aplicaciones incluyen la no robustez de la prueba de *chi cuadrada* para varianzas normales y la varianza de transformaciones estables. [Vaart~A.W. \(1998\)](#).

El resultado básico es: supóngase es posible un estimador T_n para un parámetro θ , pero la cantidad de interés es $\phi(\theta)$ para alguna función conocida ϕ . Un estimador natural es $\phi(T_n)$. ¿Cómo las propiedades asintóticas de $\phi(T_n)$ continúan de aquellas en T_n ? [Vaart~A.W. \(1998\)](#).

Un primer resultado es una inmediata consecuencia del teorema de mapeo-continuo. Si la secuencia T_n converge en probabilidad a θ y ϕ es continua en θ , entonces $\phi(t_n)$ converge en probabilidad a $\phi(\theta)$ [Vaart~A.W. \(1998\)](#).

De los intereses mayores es que este resultado también concierne al límite de distribuciones. En particular si $\sqrt{n}(T_n - \theta)$ converge debilmente al límite de la distribución, entonces si ϕ es diferenciable es verdad dicha afirmación también para $\sqrt{n}(\phi(T_n) - \phi(\theta))$. Entonces se tiene (3.7)

$$\sqrt{n}(\phi(T_n) - \phi(\theta)) \approx \phi'(\theta)\sqrt{n}(T_n - \theta) \quad (3.7)$$

y si (3.8)

$$\sqrt{n}(T_n - \theta) \xrightarrow{D} T \quad (3.8)$$

para alguna variable T , entonces esperamos (3.9)

3.3. Métodos asintóticos

$$\sqrt{n}(\phi(T_n) - \phi(\theta)) \xrightarrow{D} \phi'(\theta)T \quad (3.9)$$

En particular, si $\sqrt{n}(T_n - \theta)$ es asintóticamente normal $N(0, \sigma^2)$ entonces esperamos (3.10)

$$\sqrt{n}(\phi(T_n) - \phi(\theta)) \xrightarrow[n \rightarrow \infty]{D} N(0, \phi'(\theta)^2 \sigma^2) \quad (3.10)$$

[Vaart~A.W. \(1998\)](#)

El uso del método Delta en inferencia estadística es frecuente en la investigación del comportamiento asintótico de estadísticas (estimadores, estadísticos de prueba, etc.).

Capítulo 4

Metodología para probar homocedasticidad bajo censura usando el estadístico de Bartlett

Sea x_1, \dots, x_n una muestra aleatoria con valores independientes e idénticamente distribuidos que siguen una distribución normal y sea x_1, \dots, x_k la muestra censurada. Sea también $\hat{\sigma}^2$ el estimador de máxima verosimilitud de la varianza la cual contiene la información de la muestra dentro del estadístico de Bartlett en la ecuación (3.1). Sea entonces $\hat{\theta}$ el estimador de máxima verosimilitud de la varianza que contendrá la información de la muestra censurada.

4.1. Estimación de la varianza en muestras normales censuradas

A propósito de definición, el problema de la estimación de una distribución normal truncada, según Gupta (1952), expone dos situaciones de diferente clase, dependiendo de: a) si la población de la cual la muestra ha sido tomada es truncada, o b) si la muestra por si misma es truncada cuando la población es completa. Para distinguir entre estos dos casos, Hald (1949) citado por el mismo autor, los llamó truncada y censurada respectivamente. De este modo, las muestras que son censuradas,

pueden serlo en dos diferentes formas:

1. Observaciones por abajo o arriba de un punto dado (también llamado punto de truncamiento) pueden ser censuradas (tipo I).
2. Las $(n - k)$ observaciones menores o mayores fuera de una muestra de tamaño n pueden ser censuradas (tipo II).

El cálculo de los estimadores de máxima verosimilitud de la media y la desviación estándar en los dos tipos son aproximadamente idénticos. [Hald \(1949\)](#) ha construido la tabla apropiada para los estimadores del tipo I y [Gupta \(1952\)](#) ha planteado y desarrollado los estimadores de máxima verosimilitud para la censura por la derecha tipo II.

4.1.1. Obtención de estimadores de Máxima Verosimilitud

Sea x_1, x_2, \dots, x_n una muestra aleatoria de tamaño n de una población normal con media μ y desviación estándar σ , y sea x_1, x_2, \dots, x_k la muestra censurada de tamaño k en el cual $x_{(k)}$ es la observación mayor y $(n - k)$ las observaciones censuradas. La obtención de los estimadores por el método de máxima verosimilitud fue planteada por [Gupta \(1952\)](#), el desarrollo es el siguiente:

La función de máxima verosimilitud de la muestra es:

$$\{n!/(k-1)!(n-k)!\} (\sqrt{2\pi}\sigma)^{-k} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^k (t_i - \mu)^2\right\} \left\{(\sqrt{2\pi}\sigma)^{-1} \int_{t_k}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}(t - \mu)^2\right\} dt\right\}^{n-k} \quad (4.1)$$

El logaritmo de la función de verosimilitud puede ser escrita como:

$$\log L = C - k \log \sigma - \frac{1}{2\sigma^2} \sum (x_i - \mu)^2 + (n - k) \log \Phi(\eta) \quad (4.2)$$

Donde C es la constante:

$$\frac{n!}{(k-1)!(n-k)!}$$

$$\eta = \frac{x_k - \mu}{\sigma} \text{ y } \Phi(\eta) = \frac{1}{\sqrt{2\pi}} \int_{\eta}^{\infty} e^{-\frac{1}{2}t^2} dt$$

Para derivar (4.2) sean:

$$A = \frac{\phi(\eta)}{\Phi(\eta)} \text{ donde } \phi(\eta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\eta^2}$$

Las ecuaciones de verosimilitud son entonces:

$$\frac{\delta \log L}{\delta \mu} = -\frac{1}{\sigma^2} \sum (x_i - \mu) + (n-k) \frac{A}{\sigma} = 0, \quad (4.3)$$

y

$$\frac{\delta \log L}{\delta \sigma} = -\frac{k}{\sigma} + \frac{1}{\sigma^3} \sum (x_i - \mu)^2 + (n-k) \eta \frac{A}{\sigma} = 0. \quad (4.4)$$

Substituyendo el valor de A de (4.3) en (4.4) se obtiene:

$$\mu_c = \bar{x} + (\theta - s^2)/d \quad (4.5)$$

Donde \bar{x} y s^2 son los estimadores de máxima verosimilitud para μ y σ respectivamente y $d = (x_k - \bar{x})$. El estimador de σ se obtiene entonces:

$$\sigma = d/z \quad (4.6)$$

donde

4.1. Estimación de la varianza en muestras normales censuradas

$$z = \eta + \left(\frac{1}{p} - 1\right) A, \text{ donde } p = \frac{k}{n}$$

De forma que

$$\psi \equiv \frac{s^2}{s^2 + d^2} = \frac{1 + \eta z - z^2}{1 + \eta z} \quad (4.7)$$

Con (4.7) se pueden obtener los valores de z para así obtener (4.6), de modo que [Gupta \(1952\)](#) elabora tablas para diferentes valores de ψ y p que corresponde a la propoción de observados en la muestra censurada.

[Gupta \(1952\)](#) determina las propiedades de $\hat{\mu}_c$ y $\hat{\theta}$ afirmando que son estimadores consistentes y eficientes (4.1), y así para cuando n es grande el problema de estimación es completamente resuelto, excepto para dificultades que podrían presentarse cuando los datos son agrupados. Cuando n es pequeña los estimadores de máxima verosimilitud podrían, por supuesto, ser sesgados y la fórmula asintótica para varianzas y covarianzas no aplicaría estrictamente.

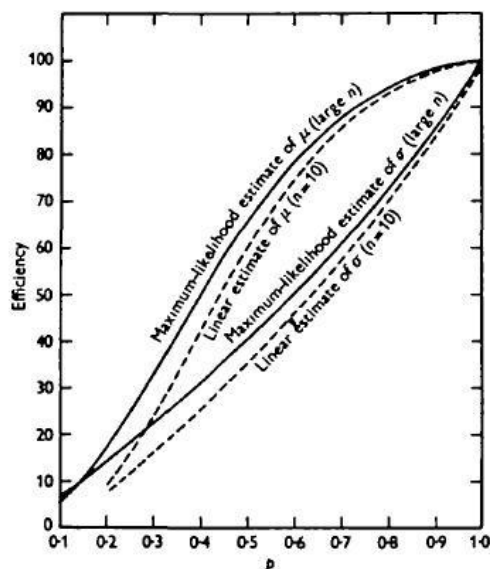


Figura 4.1: Eficiencia de los estimadores lineales y de máxima verosimilitud, de las constantes de la distribución Normal para diferentes valores de k/n [Gupta \(1952\)](#)

Para obtener θ en la práctica, se empleó el paquete R `Fistdistrplus` elaborado por [Delignette-Muller et al. \(2011\)](#). Ver Apéndice.

4.2. Uso de la prueba de Bartlett bajo censura

Sea $x_1, x_2, \dots, x_{k_1}, \dots, x_{n_1}$ y $y_1, y_2, \dots, y_{k_2}, \dots, y_{n_2}$ dos muestras aleatorias de tamaño n_i censuradas en el valor x_{k_i} de una población normal con media μ_i y desviación estándar σ_i y sean $\hat{\mu}_{ci}$ y $\hat{\theta}_i$ los estimadores respectivos de máxima verosimilitud para las muestras censuradas.

Sea h la prueba de Bartlett (3.1) para verificar el supuesto de homocedasticidad de estas dos muestras censuradas. De modo que la prueba consistiría entonces en:

$$h = \frac{1}{c} \log \frac{\left(\sum_{i=1}^t \frac{\gamma_i \hat{\theta}_i}{\gamma} \right)^\gamma}{\prod_{i=1}^t (\hat{\theta}_i)^{\gamma_i}} \quad (4.8)$$

Donde: $i = 1, 2$ y $t = 2$.

El valor de $\hat{\theta}_i$ fue obtenido, como se describe en la sección anterior. Posteriormente este valor se emplea en (4.8) y por medio del método Monte Carlo a 10,000 replicas, se obtuvieron los valores que describen la distribución de este estadístico para dos muestras censuradas.

4.2.1. Propiedad de invarianza del estadístico

Sea $x_1 < x_2 < \dots < x_k$ el estadístico de orden de una muestra de tamaño n de una población $N(\mu, \theta)$. La verosimilitud es (4.9), los estimadores de máxima verosimilitud correspondientes son denotados por $\hat{\mu}(\underline{x})$ and $\hat{\theta}(\underline{x})$.

4.2. Uso de la prueba de Bartlett bajo censura

$$\ell(\mu, \theta, \underline{x}) \propto \theta^{-k/2} \exp \left\{ -\frac{1}{2\theta} \sum_{i=1}^k (x_i - \mu)^2 \right\} \left\{ 1 - \Phi \left(\frac{x_k - \mu}{\theta^{1/2}} \right) \right\}^{n-k} \quad (4.9)$$

Estableciendo que

$$\widehat{\mu}(c\underline{x} + d\underline{1}) = c\widehat{\mu}(\underline{x}) + d$$

y

$$\widehat{\theta}(c\underline{x} + d\underline{1}) = c^2\widehat{\theta}(\underline{x}) \text{ for } c > 0$$

tenemos

$$\ell(\mu, \theta, c\underline{x} + d\underline{1}) \propto \theta^{-k/2} \exp \left\{ -\frac{1}{2\theta} \sum_{i=1}^k (cx_i + d - \mu)^2 \right\} \left\{ 1 - \Phi \left(\frac{cx_k + d - \mu}{\theta^{1/2}} \right) \right\}^{n-k}$$

Reparametrizando $\nu = (\mu - d)/c$ and $\xi = \theta/c^2$; por lo que

$$\begin{aligned} \ell(\nu, \xi, c\underline{x} + d\underline{1}) &\propto c^k \xi^{-k/2} \exp \left[-\frac{1}{2c^2\xi} \sum_{i=1}^k \{cx_i + d - (c\nu + d)\}^2 \right] \left[1 - \Phi \left\{ \frac{cx_k + d - (c\nu + d)}{c\xi^{1/2}} \right\} \right]^{n-k} \\ &\propto \xi^{-k/2} \exp \left\{ -\frac{1}{\xi} \sum_{i=1}^k (x_i - \nu)^2 \right\} \left\{ 1 - \Phi \left(\frac{cx_k - \nu}{\xi^{1/2}} \right) \right\}^{n-k} \end{aligned}$$

Ahora, suponemos que tenemos t muestras completas de una población normal con n_i observaciones en la i -ésima muestra con media μ_i y varianza θ_i . El estadístico de Bartlett es (4.8). Si tenemos muestras censuradas con k_i observaciones de la i -ésima población podemos usar los estimadores de máxima verosimilitud basados en las muestras

4.2. Uso de la prueba de Bartlett bajo censura

censuradas los cuales denotaremos por $\tilde{\theta}_i$ en lugar de $\hat{\theta}_i$.

El estadístico de Bartlett modificado es (4.10) donde $\underline{X}^{(i)}$ denota el vector de la k_i observaciones de la i -ésima población. Usaremos la respectiva notación para cada i , $\tilde{\mu}_i(c\underline{x}^{(i)} + d\underline{1}) = c\tilde{\mu}_i(\underline{x}^{(i)}) + d$ y $\tilde{\theta}_i(c\underline{x}^{(i)} + d\underline{1}) = c^2\tilde{\mu}_i(\underline{x}^{(i)})$ para cualquier c mayor a 0 y $d \in (-\infty, \infty)$.

$$h(\underline{\tilde{\theta}}) = \frac{1}{c} \log \frac{\left\{ \sum_{i=1}^t \frac{\gamma_i \tilde{\theta}_i(\underline{X}^{(i)})}{\gamma} \right\}^\gamma}{\prod_{i=1}^t \tilde{\theta}_i^{\gamma_i}(\underline{X}^{(i)})} \quad (4.10)$$

Entonces tenemos que (4.10) es equivalente a (4.11). Así que, bajo la hipótesis nula $\theta_i = \theta_2 = \dots = \theta_k = \theta$, tenemos (4.11) el cual es el parámetro libre, i.e. no depende de θ o de cada μ_i .

$$\begin{aligned} h(\underline{\tilde{\theta}}) &= \frac{1}{c} \log \frac{\left\{ \sum_{i=1}^t \frac{\gamma_i \tilde{\theta}_i(\underline{X}^{(i)})}{\gamma} \right\}^\gamma}{\prod_{i=1}^t \tilde{\theta}_i^{\gamma_i}(\underline{X}^{(i)})} \\ &= \frac{1}{c} \log \frac{\left\{ \sum_{i=1}^t \frac{\gamma_i \tilde{\theta}_i(\sqrt{\theta_i} \underline{Z}^{(i)} + \mu_i \underline{1})}{\gamma} \right\}^\gamma}{\prod_{i=1}^t \tilde{\theta}_i^{\gamma_i}(\sqrt{\theta_i} \underline{Z}^{(i)} + \mu_i \underline{1})} \\ &= \frac{1}{c} \log \frac{\left\{ \sum_{i=1}^t \frac{\gamma_i \theta_i \tilde{\theta}_i(\underline{Z}^{(i)})}{\gamma} \right\}^\gamma}{\prod_{i=1}^t \theta_i^{\gamma_i} \tilde{\theta}_i^{\gamma_i}(\underline{Z}^{(i)})} \\ h(\underline{\tilde{\theta}}) &= \frac{1}{c} \log \frac{\theta^\gamma \left\{ \sum_{i=1}^t \frac{\gamma_i \tilde{\theta}_i(\underline{Z}^{(i)})}{\gamma} \right\}^\gamma}{\theta_i^{\sum \gamma_i} \prod_{i=1}^t \tilde{\theta}_i^{\gamma_i}(\underline{Z}^{(i)})} = \frac{1}{c} \log \frac{\left\{ \sum_{i=1}^t \frac{\gamma_i \tilde{\theta}_i(\underline{Z}^{(i)})}{\gamma} \right\}^\gamma}{\prod_{i=1}^t \tilde{\theta}_i^{\gamma_i}(\underline{Z}^{(i)})} \quad (4.11) \end{aligned}$$

4.3. Distribución del estadístico de Bartlett para dos muestras censuradas

Gráficamente se muestra esta propiedad por medio de la simulación en la figura 4.2.

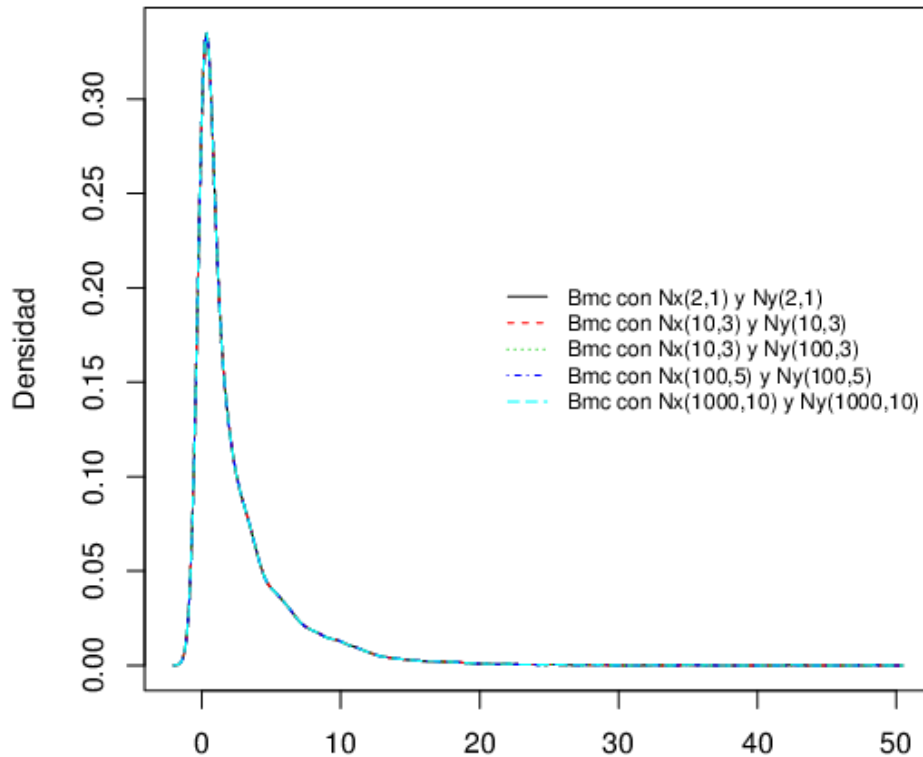


Figura 4.2: Distribuciones empíricas estimadas del estadístico de Bartlett bajo censura generadas con 10,000 muestras de tamaño $n = 50$ con 50% de censura de la distribución Normal para los parámetros especificados en la gráfica.

Sección elaborada con el apoyo del Ph.D. Barry C. Arnold. Riverside California E.U.

4.3. Distribución del estadístico de Bartlett para dos muestras censuradas

Los cuantiles de las distribuciones obtenidas por la metodología propuesta para verificar el supuesto de homogeneidad de varianzas para dos grupos de tamaño $n_1 =$ y n_2 a diferentes niveles de censura empleando

4.3. Distribución del estadístico de Bartlett para dos muestras censuradas

el estadístico de Bartlett, son contenidos en las tablas siguientes clasificadas éstas por tres tipos de casos: 1) grupos homogéneos: tabla 4.1, 2) proporción de censura: tabla 4.2 y 3) proporción de tamaño de muestra y censura: tabla 4.3. El procedimiento por simulación Monte Carlo se muestra en el Apéndice.

Además, por medio de la simulación fue posible investigar el efecto de la censura en el estadístico de estudio que corresponde a la ecuación (4.8). En la figura 4.3 se ilustra la distribución para dos muestras normales censuradas de tamaño $n_1 = n_2 = 30$ a diferentes niveles de censura mostrando que a mayor incremento del nivel de censura la cola izquierda de la distribución se vuelve más pesada.

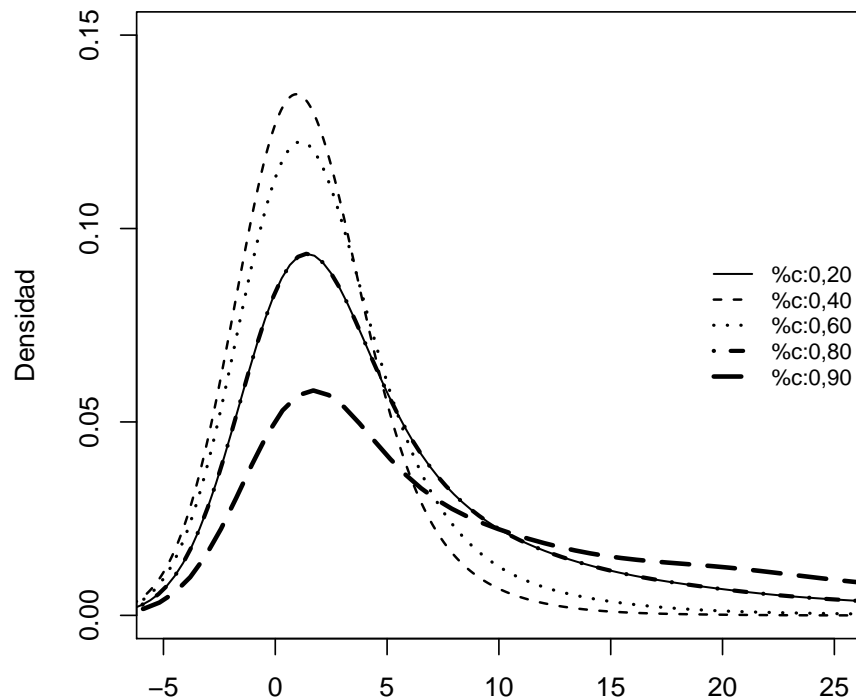


Figura 4.3: Gráfico comparativo de la distribución del estadístico de Bartlett para dos muestras aleatorias censuradas de tamaño $n_1 = n_2 = 30$ a diferentes niveles de censura.

4.3. Distribución del estadístico de Bartlett para dos muestras censuradas

Tabla 4.1: Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 = n_2$ y $c_1 = c_2$

n	c	Nivel de significancia α						
		0.005	0.01	0.02	0.05	0.1	0.15	0.5
10	0	7.9197	6.5201	5.5780	3.9268	2.7279	2.0779	0.4610
	10	9.4733	7.8327	6.4338	4.5669	3.2360	2.4595	0.5355
	20	10.6338	8.8730	7.2034	5.0520	3.5918	2.7082	0.5790
	30	13.5477	11.2893	9.2750	6.6045	4.6769	3.6269	0.7887
	40	15.8043	13.3677	11.1072	8.2349	5.9360	4.5747	1.0012
	50	21.5000	17.5495	14.3302	10.4370	7.5699	5.9565	1.3362
	60	28.1174	24.7797	20.9300	14.7292	10.6507	8.3431	1.9441
20	0	8.0329	6.6940	5.5441	3.7858	2.6408	2.0240	0.4457
	10	9.3128	7.8897	6.3836	4.5107	3.1721	2.4157	0.5414
	20	10.9760	9.1739	7.5837	5.3106	3.7537	2.8685	0.6213
	30	13.1887	10.9982	8.9405	6.5011	4.5839	3.4650	0.7622
	40	15.5769	13.0733	11.0584	8.0404	5.5533	4.2427	0.9243
	50	19.7554	17.1637	14.4139	10.1782	7.2631	5.5076	1.1747
	60	25.9860	22.5646	18.8915	13.5653	9.5542	7.4058	1.6232
30	0	7.7194	6.6906	5.4093	3.8290	2.6780	2.0330	0.4495
	10	9.1169	7.6453	6.3799	4.4336	3.1713	2.4538	0.5559
	20	10.5406	9.1416	7.2400	5.3941	3.7682	2.8982	0.6407
	30	11.9960	10.5764	8.7705	6.2354	4.4431	3.4352	0.7591
	40	15.0555	12.6939	10.7049	7.7730	5.5121	4.2781	0.9240
	50	20.3084	17.1318	13.9947	9.8433	7.0305	5.3511	1.1992
	60	25.6800	22.5438	18.6449	13.0674	9.3376	7.2308	1.6280
40	0	7.7194	6.6906	5.4093	3.8290	2.6780	2.0330	0.4495
	10	9.1169	7.6453	6.3799	4.4336	3.1713	2.4538	0.5559
	20	10.5406	9.1416	7.2400	5.3941	3.7682	2.8982	0.6407
	30	11.9960	10.5764	8.7705	6.2354	4.4431	3.4352	0.7591
	40	15.0555	12.6939	10.7049	7.7730	5.5121	4.2781	0.9240
	50	20.3084	17.1318	13.9947	9.8433	7.0305	5.3511	1.1992
	60	25.6800	22.5438	18.6449	13.0674	9.3376	7.2308	1.6280
50	0	7.7194	6.6906	5.4093	3.8290	2.6780	2.0330	0.4495
	10	9.1169	7.6453	6.3799	4.4336	3.1713	2.4538	0.5559
	20	10.5406	9.1416	7.2400	5.3941	3.7682	2.8982	0.6407
	30	11.9960	10.5764	8.7705	6.2354	4.4431	3.4352	0.7591
	40	15.0555	12.6939	10.7049	7.7730	5.5121	4.2781	0.9240
	50	20.3084	17.1318	13.9947	9.8433	7.0305	5.3511	1.1992
	60	25.6800	22.5438	18.6449	13.0674	9.3376	7.2308	1.6280
60	0	7.7194	6.6906	5.4093	3.8290	2.6780	2.0330	0.4495
	10	9.1169	7.6453	6.3799	4.4336	3.1713	2.4538	0.5559
	20	10.5406	9.1416	7.2400	5.3941	3.7682	2.8982	0.6407
	30	11.9960	10.5764	8.7705	6.2354	4.4431	3.4352	0.7591
	40	15.0555	12.6939	10.7049	7.7730	5.5121	4.2781	0.9240
	50	20.3084	17.1318	13.9947	9.8433	7.0305	5.3511	1.1992
	60	25.6800	22.5438	18.6449	13.0674	9.3376	7.2308	1.6280
70	0	7.7194	6.6906	5.4093	3.8290	2.6780	2.0330	0.4495
	10	9.1169	7.6453	6.3799	4.4336	3.1713	2.4538	0.5559
	20	10.5406	9.1416	7.2400	5.3941	3.7682	2.8982	0.6407
	30	11.9960	10.5764	8.7705	6.2354	4.4431	3.4352	0.7591
	40	15.0555	12.6939	10.7049	7.7730	5.5121	4.2781	0.9240
	50	20.3084	17.1318	13.9947	9.8433	7.0305	5.3511	1.1992
	60	25.6800	22.5438	18.6449	13.0674	9.3376	7.2308	1.6280
80	0	7.7194	6.6906	5.4093	3.8290	2.6780	2.0330	0.4495
	10	9.1169	7.6453	6.3799	4.4336	3.1713	2.4538	0.5559
	20	10.5406	9.1416	7.2400	5.3941	3.7682	2.8982	0.6407
	30	11.9960	10.5764	8.7705	6.2354	4.4431	3.4352	0.7591
	40	15.0555	12.6939	10.7049	7.7730	5.5121	4.2781	0.9240
	50	20.3084	17.1318	13.9947	9.8433	7.0305	5.3511	1.1992
	60	25.6800	22.5438	18.6449	13.0674	9.3376	7.2308	1.6280
90	0	7.7194	6.6906	5.4093	3.8290	2.6780	2.0330	0.4495
	10	9.1169	7.6453	6.3799	4.4336	3.1713	2.4538	0.5559
	20	10.5406	9.1416	7.2400	5.3941	3.7682	2.8982	0.6407
	30	11.9960	10.5764	8.7705	6.2354	4.4431	3.4352	0.7591
	40	15.0555	12.6939	10.7049	7.7730	5.5121	4.2781	0.9240
	50	20.3084	17.1318	13.9947	9.8433	7.0305	5.3511	1.1992
	60	25.6800	22.5438	18.6449	13.0674	9.3376	7.2308	1.6280
100	0	7.7194	6.6906	5.4093	3.8290	2.6780	2.0330	0.4495
	10	9.1169	7.6453	6.3799	4.4336	3.1713	2.4538	0.5559
	20	10.5406	9.1416	7.2400	5.3941	3.7682	2.8982	0.6407
	30	11.9960	10.5764	8.7705	6.2354	4.4431	3.4352	0.7591
	40	15.0555	12.6939	10.7049	7.7730	5.5121	4.2781	0.9240
	50	20.3084	17.1318	13.9947	9.8433	7.0305	5.3511	1.1992
	60	25.6800	22.5438	18.6449	13.0674	9.3376	7.2308	1.6280

Sigue ...

4.3. Distribución del estadístico de Bartlett para dos muestras censuradas

Tabla 4.1: Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 = n_2$ y $c_1 = c_2$ (continuación)

n	c	Nivel de significancia α							
		0.005	0.01	0.02	0.05	0.1	0.15	0.5	
	30	12.9934	10.9968	9.0592	6.3734	4.3882	3.3727	0.7513	
	40	16.1334	13.5854	11.2351	7.7410	5.4907	4.1946	0.8996	
	50	19.9232	17.3704	13.9260	9.6450	6.6274	5.2412	1.1120	
	60	25.7464	21.7038	17.5967	12.4828	8.8592	6.8578	1.5163	
	70	38.0027	31.1990	25.5477	18.4378	13.0137	10.0007	2.1067	
	80	63.5176	52.2203	42.2302	29.9573	21.2501	15.9071	3.5416	
	90	127.1085	110.1533	90.9777	65.7207	47.3390	36.7070	8.3133	
	80	0	7.9394	6.5859	5.3750	3.9534	2.7136	2.0833	0.4612
		10	8.8714	7.5999	6.4024	4.5349	3.2020	2.4253	0.5369
20		10.5716	8.8897	7.5708	5.3234	3.7597	2.8823	0.6140	
30		12.6929	10.9363	8.9160	6.4580	4.6247	3.5005	0.7739	
40		15.5753	13.0240	10.9681	7.7391	5.5317	4.2203	0.9375	
50		20.3371	16.9974	13.6625	9.8363	6.9344	5.3938	1.1884	
60		27.4458	21.6448	17.8916	12.3412	8.9471	6.8797	1.6048	
70		37.7688	31.5865	25.7608	18.3306	12.5571	9.7620	2.1826	
80		60.6929	49.9705	41.2577	28.8356	20.6643	15.5977	3.5694	
90	126.4004	107.5004	87.3046	63.6409	46.4209	35.7203	8.1450		
90	0	7.9368	6.4754	5.2840	3.8079	2.7051	2.0459	0.4420	
	10	9.4749	7.8197	6.2567	4.3809	3.0961	2.3486	0.5394	
	20	10.4346	8.9611	7.3011	5.1792	3.6689	2.8494	0.6148	
	30	12.7584	10.7602	8.8534	6.4063	4.5540	3.5148	0.7780	
	40	15.4900	13.1864	10.7608	7.5985	5.4008	4.1374	0.9070	
	50	20.6293	16.0716	13.2870	9.5651	6.9551	5.3370	1.1686	
	60	26.3889	21.6854	17.8847	12.8882	9.0425	6.8809	1.5359	
	70	35.9286	30.5243	24.7959	18.1278	12.7508	9.8412	2.1017	
	80	58.7263	50.0047	40.6996	28.5208	20.1036	15.6492	3.4966	
90	127.9252	105.2058	87.1813	65.2406	47.0458	35.9025	7.9492		
100	0	7.6634	6.5852	5.4426	3.8668	2.6706	2.0445	0.4596	
	10	9.1285	7.7133	6.2841	4.4039	3.1689	2.4122	0.5361	
	20	10.5078	9.1043	7.2134	5.3088	3.6597	2.8250	0.6210	
	30	13.2659	10.7556	8.8241	6.3625	4.5006	3.4090	0.7365	
	40	15.7232	12.9146	10.8147	7.8310	5.5322	4.1904	0.8841	
	50	19.2022	16.5183	13.9014	9.6863	6.9850	5.3354	1.1219	
	60	25.5423	21.3219	17.6306	13.0039	8.9803	6.8268	1.4860	
	70	36.8502	31.2871	25.0613	17.9273	12.5962	9.7443	2.0733	
	80	59.2361	48.5416	40.3481	28.6637	20.1678	15.2159	3.4393	
90	121.3253	105.7845	86.7215	65.2475	46.2702	34.5349	7.6781		

Fin

Tabla 4.2: Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 = n_2$ y $c_1 \neq c_2$

n	c	Nivel de significancia α						
		0.005	0.01	0.02	0.05	0.1	0.15	0.5
	0,20	9.0769	7.7209	6.3477	4.5762	3.2302	2.4971	0.5547
	0,40	12.2044	10.2544	8.4451	6.1429	4.2459	3.2670	0.7198
	0,60	20.5339	16.5658	13.3884	9.2427	6.5277	4.9646	1.0310
	0,80	51.2935	43.1343	35.2856	24.4205	16.6202	12.3554	2.4903
	0,90	135.1597	120.3753	97.8934	73.6348	53.5708	40.6497	8.2159
	10,20	9.4127	8.2740	6.8363	4.9409	3.4902	2.7009	0.5887
	10,40	12.4364	10.6752	8.8005	6.4117	4.5326	3.4209	0.7529
	10,60	21.0225	16.9686	13.6655	9.5018	6.7025	5.0535	1.1089
	10,80	51.6528	44.0154	35.2999	24.5428	16.5497	12.3234	2.5164

Sigue ...

4.3. Distribución del estadístico de Bartlett para dos muestras censuradas

Tabla 4.2: Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 = n_2$ y $c_1 \neq c_2$ (continúa)

n	c	Nivel de significancia α						
		0.005	0.01	0.02	0.05	0.1	0.15	0.5
50	10,90	136.7063	120.0084	97.9297	73.7820	53.3370	40.6536	8.1461
	30,20	11.1951	9.6564	7.8187	5.8235	4.0735	3.0979	0.7105
	30,40	14.0001	11.9856	9.8816	7.1748	5.0381	3.7939	0.8543
	30,60	21.8674	18.5407	14.3312	10.0326	7.0456	5.3036	1.1894
	30,80	51.5341	43.7046	36.2110	24.8377	16.6365	12.4098	2.5691
	30,90	133.4885	118.3410	98.1520	73.4194	52.7454	39.9610	7.9205
	50,20	15.2699	12.9940	10.5562	7.5451	5.3995	4.1506	0.9275
	50,40	17.2553	14.9486	12.4201	8.7243	6.2182	4.8689	1.0824
	50,60	24.2592	20.7212	16.5326	11.3440	8.1287	6.2619	1.4218
	50,80	51.2403	44.2092	36.2820	25.3125	16.8854	12.7758	2.7087
	50,90	133.6382	115.1462	96.8274	73.1176	52.1902	39.4010	7.7292
	70,20	29.8496	25.3457	19.8814	13.9294	9.4384	7.0123	1.4608
	70,40	30.7695	26.2902	20.9509	14.6236	9.9219	7.6194	1.5807
	70,60	32.8527	28.0973	23.1714	16.5738	11.4958	8.7206	1.9772
	70,80	52.0177	45.1800	38.3333	27.6181	18.5998	14.4761	3.0957
	70,90	134.7483	114.9096	95.2097	69.8350	49.8586	37.9545	7.5880
	0,20	9.0856	7.6903	6.3275	4.5516	3.2359	2.4986	0.5557
	0,40	11.8092	10.3179	8.2320	5.7842	4.1135	3.1941	0.7181
	0,60	20.0535	15.9257	12.6415	9.1085	6.3891	4.8388	1.0240
	0,80	46.5855	39.2753	30.7765	20.3167	13.7752	10.2993	2.2904
0,90	109.5003	96.1034	79.4997	56.3019	39.2702	28.9581	5.6127	
10,20	9.9903	8.1311	6.7974	4.8995	3.4532	2.6458	0.5806	
10,40	12.3874	10.5662	8.7398	6.0141	4.3175	3.3028	0.7548	
10,60	20.0545	16.5677	13.3502	9.3422	6.5425	4.9530	1.0487	
10,80	47.3413	39.0693	31.1338	20.4558	13.9923	10.4303	2.2761	
10,90	109.8600	96.9642	79.0662	56.5193	38.8468	29.2032	5.5794	
30,20	11.4314	9.9248	8.0846	5.7687	4.0980	3.0986	0.7039	
30,40	14.0457	11.9444	9.8121	6.9338	4.8827	3.7647	0.8426	
30,60	21.3285	17.3144	14.4233	9.8700	7.0791	5.2831	1.1413	
30,80	48.4444	38.7941	30.4827	20.9924	14.0564	10.7514	2.3554	
30,90	110.6925	96.2851	80.5586	55.9445	38.9310	28.8405	5.5938	
50,20	15.3620	13.0580	10.4955	7.7146	5.3398	4.0181	0.8717	
50,40	16.7639	14.7766	11.9403	8.6263	6.0038	4.6264	1.0397	
50,60	22.9194	19.6112	16.1596	11.4061	7.9626	6.2641	1.3173	
50,80	50.8131	40.9028	31.3069	21.7652	14.7994	11.2747	2.5216	
50,90	111.8126	95.8556	80.7409	56.0799	39.0009	28.8991	5.7947	
70,20	27.0070	21.6010	17.7574	12.2567	8.5431	6.5292	1.4185	
70,40	29.1754	23.7048	18.5208	13.0185	9.0499	7.1500	1.5840	
70,60	32.9840	26.9701	21.7643	15.5587	11.2588	8.5347	1.8478	
70,80	52.8847	43.1469	34.9404	24.4334	17.0438	13.2779	3.0642	
70,90	113.8207	98.5280	79.0701	55.5012	38.2647	28.7211	5.9389	
0,20	9.1374	7.7974	6.4877	4.7218	3.2364	2.4600	0.5374	
0,40	11.9248	10.1674	8.1594	6.0103	4.1488	3.2019	0.6875	
0,60	18.8651	14.6063	12.0757	8.5037	5.9874	4.6318	1.0028	
0,80	40.7693	33.1983	27.5210	18.8287	12.7645	9.5187	2.0220	
0,90	98.2457	82.1020	65.5031	44.7614	31.4147	23.3787	4.6868	
10,20	9.7413	8.3007	6.9556	5.0147	3.4223	2.6092	0.5643	
10,40	12.7317	10.4523	8.6410	6.3872	4.3448	3.3537	0.7204	
10,60	19.1933	15.3616	12.5477	8.8118	6.2310	4.8077	1.0581	
10,80	41.2588	33.4392	27.8621	19.0890	12.9261	9.7850	2.0412	
10,90	99.3977	81.0780	64.5011	45.0303	31.2900	23.4389	4.7120	
30,20	11.3530	9.8323	8.1046	6.0145	4.2070	3.2112	0.7039	
30,40	13.9499	12.2607	9.8198	7.2135	5.1115	3.8770	0.8627	
30,60	20.7163	16.8414	13.7903	9.6909	6.8750	5.3359	1.1664	
30,80	41.3289	35.2463	28.7763	19.8453	13.4882	10.0787	2.1619	
30,90	98.7603	82.7770	65.3156	44.9810	31.5667	23.7152	4.7534	
50,20	15.7413	12.7921	10.4270	7.5594	5.4445	4.1306	0.9045	
50,40	17.9738	14.9848	11.8812	8.7033	6.2253	4.7735	1.0520	
50,60	23.2530	19.2732	15.9186	11.3156	8.0499	6.1286	1.3882	
50,80	43.2438	37.4211	29.4293	21.0573	14.3146	10.8811	2.4226	

Sigue ...

4.3. Distribución del estadístico de Bartlett para dos muestras censuradas

Tabla 4.2: Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 = n_2$ y $c_1 \neq c_2$ (continúa)

n	c	Nivel de significancia α						
		0.005	0.01	0.02	0.05	0.1	0.15	0.5
	50,90	101.3096	83.1403	64.7427	45.8544	32.0327	24.4162	4.9765
	70,20	26.0647	21.8323	16.9591	12.0712	8.4188	6.4022	1.3991
	70,40	27.4213	22.3046	18.5578	13.2568	9.1816	7.0038	1.5320
	70,60	32.4488	26.9651	21.7042	15.2812	10.8851	8.4190	1.8352
	70,80	49.8281	42.6747	33.1720	23.8211	17.0395	13.0182	2.8792
	70,90	99.2440	82.3169	67.7419	47.2529	33.3101	25.0126	5.3900
	0,20	9.5322	8.0658	6.4591	4.5568	3.1714	2.4643	0.5545
	0,40	11.8356	10.2467	8.3902	5.8747	4.1192	3.1675	0.6794
	0,60	17.6331	15.0128	12.4778	8.8473	6.0675	4.6408	0.9501
	0,80	38.1610	31.4267	25.5352	17.8021	12.4610	9.3926	1.9760
	0,90	91.6650	79.4322	64.8447	43.2382	29.5637	22.1961	4.6377
	10,20	9.9260	8.5079	6.8085	4.8340	3.4111	2.5966	0.5938
	10,40	12.9671	10.4739	8.5712	6.1653	4.3784	3.3644	0.7072
	10,60	18.5796	15.2219	12.7828	9.0424	6.3452	4.7984	1.0050
	10,80	38.5167	31.1279	25.5282	18.0550	12.8949	9.6477	1.9925
	10,90	94.4398	79.2201	64.4366	43.4813	29.5361	21.9834	4.6935
	30,20	11.8507	10.0359	8.1196	5.8596	4.0946	3.1874	0.6795
	30,40	15.1886	12.1448	10.0228	7.0743	4.9789	3.8655	0.8278
100	30,60	20.3934	17.0118	13.4470	9.7049	6.9907	5.3724	1.1221
	30,80	40.2357	32.6961	26.7542	19.0235	13.4452	10.0860	2.1221
	30,90	93.6834	79.6054	64.0449	43.2395	30.0731	22.6706	4.7133
	50,20	14.8810	12.8230	10.5903	7.6564	5.2956	4.1002	0.8670
	50,40	17.8123	15.0107	12.1955	8.7196	6.1817	4.7119	0.9788
	50,60	22.6559	19.3879	15.9218	11.3735	7.9871	6.0901	1.3275
	50,80	42.7920	35.5933	28.7388	20.0447	14.2607	10.8893	2.2638
	50,90	95.2532	82.0208	64.5110	45.0334	30.7892	23.0035	4.8161
	70,20	25.5990	21.0626	16.5467	11.7281	8.1646	6.2521	1.3910
	70,40	27.7487	22.8666	18.2046	12.7050	9.0499	6.8884	1.5308
	70,60	30.2652	26.0007	21.1629	15.4808	10.9859	8.3021	1.8276
	70,80	48.7258	39.7597	32.9634	23.0132	16.9629	12.8447	2.7057
	70,90	97.6407	82.3806	65.7906	46.8790	32.1533	24.3478	5.2658

Fin

Tabla 4.3: Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 \neq n_2$ y $c_1 \neq c_2$

n	c	Nivel de significancia α						
		0.005	0.01	0.02	0.05	0.1	0.15	0.5
	0,20	9.3818	7.9368	6.2508	4.5085	3.1784	2.4367	0.5311
	0,40	11.7260	9.6278	7.8970	5.5752	3.9852	3.0539	0.6712
	0,60	18.4784	15.3181	12.3283	8.2619	5.7520	4.2860	0.9606
	10,20	9.9328	8.0973	6.7324	4.7636	3.4875	2.6477	0.5855
	10,40	11.7868	10.3119	8.1172	5.9644	4.1752	3.2515	0.7159
	10,60	19.0219	15.4852	12.8171	8.5958	5.8401	4.4362	1.0207
30, 40	30,20	11.7031	9.7639	7.8765	5.7405	4.1407	3.1739	0.6914
	30,40	13.7867	11.8746	9.4998	6.7899	4.7585	3.7550	0.8390
	30,60	20.3579	17.1623	13.5824	9.4466	6.4065	4.8893	1.1470
	50,20	16.2307	13.8187	11.0688	7.9184	5.7234	4.3717	0.9896
	50,40	17.6499	15.1247	12.2561	8.9057	6.3067	4.8961	1.1011
	50,60	22.9112	19.4372	15.9594	11.2318	7.8743	6.0110	1.3922
	0,20	8.2581	7.1390	5.9630	4.2582	2.9845	2.2836	0.5102
	0,40	9.7201	8.3172	6.9629	4.9316	3.4415	2.6869	0.6027
	0,60	13.3938	11.3024	9.0511	6.4244	4.5403	3.3998	0.7401

Sigue ...

4.3. Distribución del estadístico de Bartlett para dos muestras censuradas

Tabla 4.3: Valores críticos de la distribución del estadístico de Bartlett para dos muestras censuradas $n_1 \neq n_2$ y $c_1 \neq c_2$ (continúe)

n	c	Nivel de significancia α						
		0.005	0.01	0.02	0.05	0.1	0.15	0.5
	10,20	10.0445	8.6365	7.1395	4.8883	3.5518	2.7171	0.6065
	10,40	13.3325	11.9379	9.5215	6.5428	4.7226	3.6263	0.7838
	10,60	21.8850	18.4622	14.8440	10.3364	7.0700	5.5163	1.2314
	30,20	12.0481	9.9568	8.0554	5.6681	4.0200	3.0979	0.6777
	30,40	15.4120	12.5930	10.1900	7.3230	5.1040	3.9420	0.8690
	30,60	23.4767	19.3839	15.5010	10.7503	7.5601	5.8621	1.2855
	50,20	14.2603	12.1798	9.8677	7.0483	4.8395	3.7531	0.8134
	50,40	17.1033	14.5266	11.8464	8.3928	5.9697	4.5513	0.9807
	50,60	24.2645	20.7444	16.4982	11.6237	8.4619	6.4662	1.3991
	70, 80	0,20	9.1431	7.9357	6.2471	4.4735	3.2157	2.4485
0,40		12.4933	10.1802	8.3524	5.7230	3.9686	2.9696	0.6603
0,60		18.2928	14.8898	11.8518	8.2783	5.7895	4.4594	0.9306
10,20		10.0786	8.2379	6.7417	4.8082	3.3793	2.6057	0.5847
10,40		13.0688	10.8954	8.5273	5.9483	4.2346	3.1795	0.7028
10,60		18.4498	15.4000	12.0566	8.6821	6.0303	4.6432	0.9556
30,20		12.1873	10.2293	8.3862	5.7093	4.1385	3.1257	0.7046
30,40		14.9547	12.5897	10.1297	6.8894	4.8798	3.6621	0.8167
30,60		20.2341	16.9495	13.6383	9.5807	6.6251	5.0153	1.1036
50,20		19.9409	16.4909	13.5524	9.4855	6.6744	5.1507	1.1002
70, 100	0,20	8.7656	7.3652	6.1325	4.4437	3.1021	2.3561	0.5184
	0,40	11.7102	9.4520	7.7550	5.4046	3.8069	2.8852	0.6152
	0,60	16.3011	13.4152	10.9684	7.6794	5.4744	4.2434	0.8731
	10,20	9.7079	8.0176	6.5827	4.7937	3.3075	2.5423	0.5851
	10,40	12.3123	10.0548	8.1184	5.8399	4.0793	3.1294	0.6867
	10,60	17.0408	14.0514	11.2934	8.0934	5.7730	4.4493	0.9175
	30,20	11.8174	10.1010	8.3062	5.8067	4.1279	3.2252	0.6944
	30,40	14.4728	12.0459	9.9177	7.0286	4.7908	3.6240	0.8077
	30,60	19.9330	15.8779	12.8507	9.1122	6.3841	4.8736	1.0643
	50,20	17.0695	13.7470	11.1019	7.7353	5.5011	4.2586	0.9504
50,40	18.8355	15.3660	12.5130	8.8075	6.2761	4.7914	1.0098	
50,60	22.3158	19.2162	15.4699	11.1866	7.8096	5.9755	1.2628	

Fin

Capítulo 5

Distribución para muestras grandes del estadístico de Bartlett modificado h

5.0.1. Varianza asintótica

Notar que [Gupta \(1952\)](#) obtiene la distribución asintótica de los estimadores de máxima verosimilitud (mle) de μ and σ , i.e. la media y la desviación estándar respectivamente. Por lo que en la i -ésima población tenemos una muestra de tamaño n_i en la cual observamos solo las más pequeñas k_i observaciones. Notación: el total del tamaño de muestras en todas las t poblaciones será denotada por n , y eventualmente será $n \rightarrow \infty$ en tal caso una forma en que $\frac{n_i}{n} \rightarrow \delta_i$, o $n_i = \delta_i n$, y las proporciones de observaciones no censuradas en las poblaciones satisface $\frac{k_i}{n_i} \rightarrow p_i$, así que $k_i = p_i n_i = p_i \delta_i n$.

Ahora, [Gupta \(1952\)](#) define $\Phi(\eta) = \int_{\eta}^{\infty} \phi(t) dt$. Habitualmente esto sería denotado como $\bar{\Phi}(\eta)$ ya que es la función de supervivencia no la función de distribución de la distribución normal estándar. En este caso usaremos la notación $\Phi(\eta) = \int_{-\infty}^{\eta} \phi(t) dt$ and $\bar{\Phi}(\eta) = 1 - \Phi(\eta) = \int_{\eta}^{\infty} \phi(t) dt$. Por lo que el término A de [Gupta \(1952\)](#) will be: $A(\eta) = \frac{\phi(\eta)}{\bar{\Phi}(\eta)}$, reconocida como la función *hazard*.

5. Distribución para muestras grandes del estadístico de Bartlett modificado h

En la i -ésima población, los mle μ_i y σ_i , i.e., $\hat{\mu}_i$ y $\hat{\sigma}_i$ son obtenidos numéricamente, o reportados por Gupta (1952) en la Tabla 1. Sin embargo, éstos mismos pueden ser obtenidos por un programa de optimización. La varianza y covarianza asintótica de $(\hat{\mu}_i, \hat{\sigma}_i)$ es entonces dada por (5.1), donde $\Sigma = V^{-1}$ y V es (5.2), en el cual cada v_{ij} es definida a partir de la ecuación (5.3) hasta (5.5).

$$\frac{\sigma_i^2}{n_i} \Sigma = \frac{\sigma_i^2}{n_i} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \quad (5.1)$$

$$V = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \quad (5.2)$$

$$v_{11} = p_i + \phi(\Phi^{-1}(p_i)) \left[\frac{\phi(\Phi^{-1}(p_i))}{1 - p_i} - \Phi^{-1}(p_i) \right] \quad (5.3)$$

$$v_{12} = v_{21} = \phi(\Phi^{-1}(p_i)) \left[\Phi^{-1}(p_i) \left\{ \frac{\phi(\Phi^{-1}(p_i))}{1 - p_i} - \Phi^{-1}(p_i) \right\} - 1 \right] \quad (5.4)$$

$$v_{22} = 2p_i - \Phi^{-1}(p_i) \phi(\Phi^{-1}(p_i)) + [\Phi^{-1}(p_i)]^2 \phi(\Phi^{-1}(p_i)) \left[\frac{\phi(\Phi^{-1}(p_i))}{1 - p_i} - \Phi^{-1}(p_i) \right] \quad (5.5)$$

Notar que en artículo de Gupta (1952) él da las expresiones en términos de $\hat{\eta}$ que satisface $\Phi(\hat{\eta}) = \int_{-\infty}^{\hat{\eta}} \phi(t) dt = p_i$, i.e., $\hat{\eta} = \Phi^{-1}(p_i)$. En este caso hemos hecho esta sustitución en sus expresiones para los v_{ij} 's. En este caso el interés es estimar σ_i^2 . Así, notar primero que:

$$\text{var}(\hat{\sigma}) = \frac{\sigma_i^2}{n_i} \sigma_{22} = \frac{\sigma_i^2}{n_i} \left[\frac{v_{11}}{v_{11}v_{22} - v_{12}^2} \right]$$

Ahora $\hat{\sigma}^2 = (\hat{\sigma})^2$ y podemos usar el método delta con $g(u) = u^2$ para

5. Distribución para muestras grandes del estadístico de Bartlett modificado h

obtener la varianza asintótica de $\widehat{\sigma}^2$, i.e. (5.6).

$$\text{var}(\widehat{\sigma}^2) = \frac{4\sigma_i^4}{n_i}\sigma_{22} = \frac{4\sigma_i^4}{n_i} \left[\frac{v_{11}}{v_{11}v_{22} - v_{12}^2} \right] \quad (5.6)$$

Notar que las v_{ij} 's en todas las funciones de p_i así que la función $\kappa(p) = \left[\frac{v_{11}(p)}{v_{11}(p)v_{22}(p) - v_{12}^2(p)} \right]$ podría ser tabulada, y tendríamos $\text{var}(\widehat{\sigma}^2) = \frac{4\sigma_i^4}{n_i}\sigma_{22} = \frac{4\sigma_i^4}{n_i}\kappa(p_i)$. De modo que usaremos esta notación considerando la distribución del estadístico de Bartlett modificado donde θ_i denota σ_i^2 ; así que $\text{var}(\widehat{\theta}_i) = \frac{4\theta_i^2}{n_i}\kappa(p_i)$. Notar que en ausencia de censura $p_i = 1$ y $\kappa(1) = 1/2$ son $\text{var}(\widehat{\theta}_i) = \frac{2\theta_i^2}{n_i}$.

5.0.2. Distribución para muestras grandes

Asumimos que tenemos $k_i = p_i n_i = p_i \delta_i n$ observaciones no censuradas de la muestra de tamaño $n_i = \delta_i n$ de la i -ésima población la cual tiene una distribución $N(\mu_j, \theta_j)$. Sea n que denota el número total de observaciones en el total de t poblaciones, entonces $\sum_{i=1}^t \delta_i = 1$. Según

[Gupta \(1952\)](#), $\tilde{\theta}_i \sim N(\theta_i, \frac{\theta_i}{n\delta_i}\sigma_{22}^{(i)})$, para $i = 1, 2, \dots, t$. Ahora si H_0 es verdadera entonces $\tilde{\theta}_i \sim N(\theta, \frac{\theta}{n\delta_i}\sigma_{22}^{(i)})$, para $i = 1, 2, \dots, t$. En general tenemos:

$$\begin{aligned} T_n &= h(\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_t) \\ &= \frac{1}{c} \log \frac{\left(\sum_{i=1}^t \frac{\gamma_i \tilde{\theta}_i}{\gamma} \right)^\gamma}{\prod_{i=1}^t (\tilde{\theta}_i)^{\gamma_i}} \end{aligned}$$

5. Distribución para muestras grandes del estadístico de Bartlett modificado h

Considerar la función $h(\theta_1, \theta_2, \dots, \theta_t) = \frac{1}{c} \log \frac{\left(\sum_{i=1}^t \frac{\gamma_i \theta_i}{\gamma} \right)^\gamma}{\prod_{i=1}^t (\theta_i)^{\gamma_i}}$

Para cada $j \in \{1, 2, \dots, t\}$, tenemos

$$\begin{aligned} \frac{\partial}{\partial \theta_j} h(\theta_1, \theta_2, \dots, \theta_t) &= \frac{\partial}{\partial \theta_j} \left[\frac{1}{c} \log \frac{\left(\sum_{i=1}^t \frac{\gamma_i \theta_i}{\gamma} \right)^\gamma}{\prod_{i=1}^t (\theta_i)^{\gamma_i}} \right] \\ &= \frac{1}{c} \left\{ \gamma \frac{\partial}{\partial \theta_j} \log \left(\sum_{i=1}^t \frac{\gamma_i \theta_i}{\gamma} \right) - \sum_{i=1}^t \gamma_i \frac{\partial}{\partial \theta_j} \log (\theta_i) \right\} \\ &= \frac{\gamma_j}{c} \left\{ \left(\sum_{i=1}^t \frac{\gamma_i \theta_i}{\gamma} \right)^{-1} - \theta_j^{-1} \right\} \end{aligned}$$

Por lo que para una n grande tenemos:

$$T_n \sim N(h(\underline{\theta}), \psi(\underline{\theta})/n) \quad \text{or} \quad \sqrt{n}(T_n - h(\underline{\theta})) \xrightarrow{d} N(0, \psi(\underline{\theta}))$$

$$\text{donde } \psi(\underline{\theta}) = \frac{1}{c^2} \sum_{j=1}^t \gamma_j^2 \left[\frac{\gamma \theta_j}{\left(\sum_{i=1}^t \gamma_i \theta_i \right)} - 1 \right]^2 \frac{1}{\delta_j} \sigma_{22}^{(j)}$$

Sin embargo si $\theta_1 = \theta_2 = \dots = \theta_t = \theta$ entonces para cada $j \in \{1, 2, \dots, t\}$ tenemos:

$$\frac{\partial}{\partial \theta_j} h(\theta_1, \theta_2, \dots, \theta_t) |_{\theta_1=\theta_2=\dots=\theta_t=\theta} = \frac{\gamma_j}{c} \left\{ \left(\sum_{i=1}^t \frac{\gamma_i \theta}{\gamma} \right)^{-1} - \theta^{-1} \right\} = 0$$

Así que necesitamos un método delta de segundo orden para obtener la distribución asintótica de $n(T_n - h(\theta, \theta, \dots, \theta))^2$ Notar que $h(\theta, \theta, \dots, \theta) =$

5. Distribución para muestras grandes del estadístico de Bartlett modificado h

0. Usando la expansión de las series de Taylor y el hecho de que $\tilde{\theta}_i$'s son independientes. Tenemos:

$$nT_n = nh(\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_t) \stackrel{d}{=} \frac{1}{2} \sum_{i=1}^t [\sqrt{n}(\tilde{\theta}_i - \theta)]^2 \frac{\partial^2}{\partial \theta_i^2} h(\theta_1, \theta_2, \dots, \theta_t) |_{\theta_1 = \theta_2 = \dots = \theta_t = \theta}$$

Sin embargo, sabemos que cuando $\theta_1 = \theta_2 = \dots = \theta_t = \theta$, $var(\tilde{\theta}_i) = \frac{4\theta^2}{n_i} \kappa(p_i)$

$$\begin{aligned} \frac{\partial^2}{\partial \theta_j^2} h(\theta_1, \theta_2, \dots, \theta_t) &= \frac{\partial}{\partial \theta_j} \frac{\gamma_j}{c} \left\{ \left(\sum_{i=1}^t \frac{\gamma_i \theta_i}{\gamma} \right)^{-1} - \theta_j^{-1} \right\} \\ &= \frac{\gamma_j}{c} \left\{ - \left(\sum_{i=1}^t \frac{\gamma_i \theta_i}{\gamma} \right)^{-2} \frac{\gamma_j}{\gamma} + \theta_j^{-2} \right\} \end{aligned}$$

$$\begin{aligned} \text{Así que } \frac{\partial^2}{\partial \theta_j^2} h(\theta_1, \theta_2, \dots, \theta_t) |_{\theta_1 = \theta_2 = \dots = \theta_t = \theta} &= \theta^{-2} \frac{\gamma_j}{c} \left\{ - \left(\sum_{i=1}^t \frac{\gamma_i}{\gamma} \right)^{-2} \frac{\gamma_j}{\gamma} + 1 \right\} = \\ &= \frac{\gamma_j}{c\theta^2} \left\{ 1 - \frac{\gamma_j}{\gamma} \right\} \end{aligned}$$

Consecuentemente

$$nT_n \stackrel{d}{=} \frac{1}{2} \sum_{i=1}^t \{ \sqrt{n}(\tilde{\theta}_i - \theta) \}^2 \frac{\gamma_i}{c\theta^2} \left(1 - \frac{\gamma_i}{\gamma} \right).$$

Pero recordando que $var(\tilde{\theta}_i) = \frac{4\theta^2}{n_i} \kappa(p_i)$ and $n_i = \delta_i n$; así que

5. Distribución para muestras grandes del estadístico de Bartlett modificado h

$$\begin{aligned} nT_n &\stackrel{d}{=} \frac{1}{2} \sum_{i=1}^t \frac{4\theta^2 \kappa(p_i)}{\delta_i} \left\{ \frac{\sqrt{n_i}(\hat{\theta}_i - \theta)}{\sqrt{4\theta^2 \kappa(p_i)}} \right\}^2 \frac{\gamma_i}{c\theta^2} \left(1 - \frac{\gamma_i}{\gamma} \right) \\ &\stackrel{d}{=} \sum_{i=1}^t \frac{2\kappa(p_i)}{c\delta_i} \gamma_i \left(1 - \frac{\gamma_i}{\gamma} \right) U_i \end{aligned}$$

Es una combinación lineal de t i.i.d. χ_1^2 variables aleatorias, U_1, U_2, \dots, U_t .
Eliminando el subíndice n tenemos

$$T_n \stackrel{d}{=} \sum_{i=1}^t \frac{2\kappa(p_i)}{cn_i} \gamma_i \left(1 - \frac{\gamma_i}{\gamma} \right) U_i$$

Usando los pesos usuales en $\gamma_i = n_i - 1 \approx n_i = \delta_i$ se obtiene

$$T_n \stackrel{d}{=} \sum_{i=1}^t \frac{2\kappa(p_i)}{c} \left(1 - \frac{n_i}{n} \right) U_i$$

Usando muestra completas entonces $k(1) = 1/2$ por lo que tenemos

$$T_n \stackrel{d}{=} \sum_{i=1}^t \frac{2\kappa(1)}{c} \left(1 - \frac{n_i}{n} \right) U_i = \frac{1}{c} \sum_{i=1}^t \left(1 - \frac{n_i}{n} \right) U_i$$

De modo que si no usamos la constante c de Bartlett tendríamos

$$\widetilde{T}_n = cT_n \stackrel{d}{=} \sum_{i=1}^t \left(1 - \frac{n_i}{n} \right) U_i$$

Así que $E(\widetilde{T}_n) = E(cT_n) = \sum_{i=1}^t \left(1 - \frac{n_i}{n} \right) = t - 1$, la cual es la media de una χ_{t-1}^2 variable aleatoria.

**5. Distribución para muestras grandes del estadístico de Bartlett
modificado h**

Sección elaborada con el apoyo del Ph.D. Barry C. Arnold. Riverside
California E.U.

Capítulo 6

Evaluación del método propuesto con la prueba de Bartlett

6.1. Potencia de la prueba

La función de potencia de una prueba de hipótesis con región de rechazo R es la función en este caso de h el estadístico de Bartlett bajo censura definida por (6.1). Casella (1990).

$$\beta(h) = P_h(X \in R) \quad (6.1)$$

Donde:

$$P_h(X \in R) = \begin{cases} \Pr(\text{Error I}) & \text{si } h \in h_0 \\ 1 - \Pr(\text{Error II}) & \text{si } h \in h_0^c \end{cases} \quad (6.2)$$

La potencia de la prueba se ilustra en la figura 6.1, realizándose por medio de la simulación Monte Carlo bajo el siguiente procedimiento:

1. Se establece el juego de hipótesis:

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ vs } H_0 : \sigma_1^2 \neq \sigma_2^2 \quad (6.3)$$

6.1. Potencia de la prueba

2. Se obtienen los pares de grupos aleatoriamente a un determinado tamaño de muestra y de censura para una distribución normal.
3. Se generan 10,000 pares de muestras por el método Monte Carlo para obtener el valor del estadístico de Bartlett en condiciones de censura.
4. Al conjunto de valores obtenidos se aplica la regla de decisión: si el valor obtenido es mayor o igual al reportado por la distribución correspondiente a un $\alpha = 0.05$ entonces se rechaza la hipótesis nula, en caso de ser menor se acepta ésta.
5. Obtener el valor de potencia al sumar el total de casos rechazados y dividirlo entre el número de repeticiones Monte Carlo.
6. Se ejecutó la prueba bajo la H_0 y bajo la H_1 para un radio de varianzas para otros 30 diferentes valores de σ^2 en el grupo dos.

A fin de valorar el efecto de la censura en la prueba de homogeneidad de varianzas por el estadístico de Bartlett en estas condiciones, se presentan tres gráficos. En la figura (a) se puede apreciar que la prueba es consistente dado que a medida que el tamaño de muestra aumenta de igual modo la potencia se incrementa, en este caso la censura es constante, ambos grupos tienen el mismo número de observaciones y tamaño de muestra. En la figura (b) se varía el porcentaje de censura para ver su efecto en la potencia de la prueba la cual desciende conforme la censura aumenta; el tamaño de muestra fue constante y al igual que el anterior ambos grupos son homogéneos en tamaño de muestra y porcentaje de observados. Finalmente en la figura (c) donde los pares de grupos son heterogéneos tanto en tamaño de muestra como en porcentaje de censura prevalece el efecto de la censura sobre el tamaño de muestra.

6.2. Tamaño de la prueba

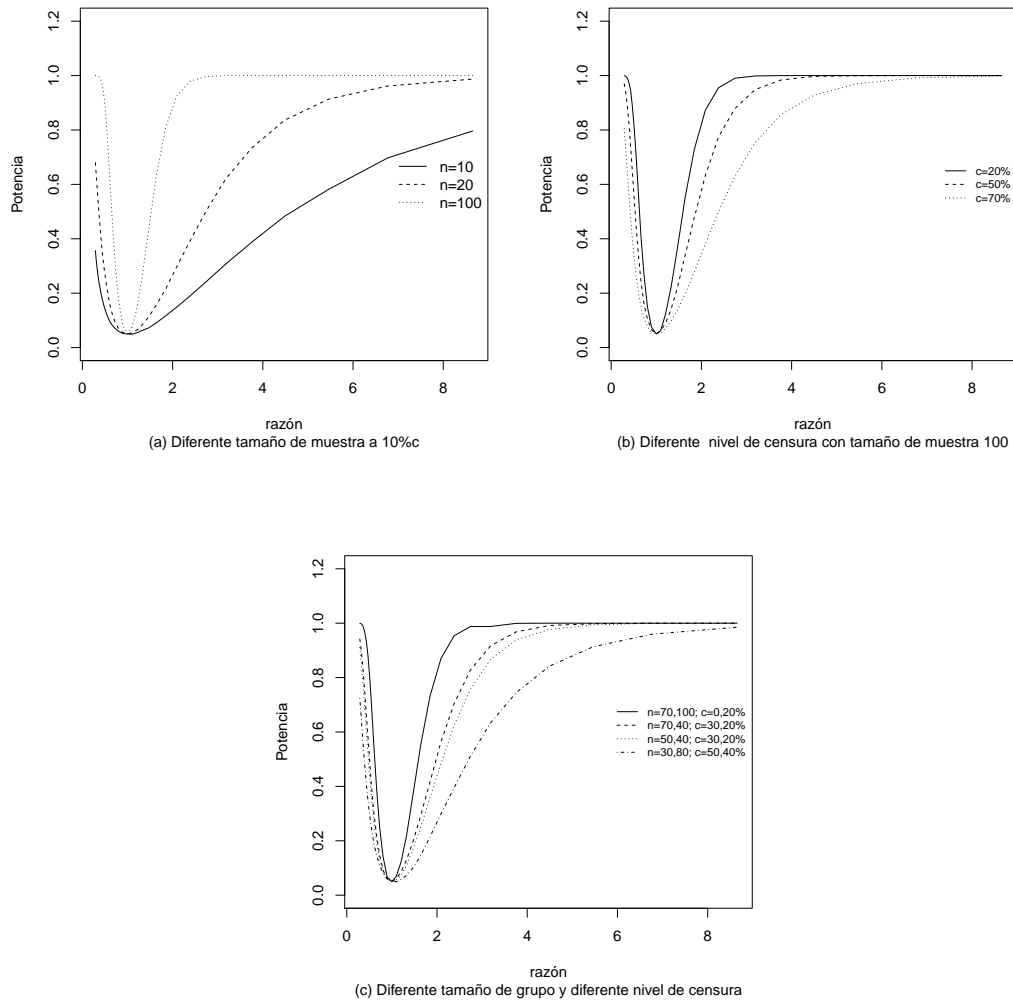


Figura 6.1: Potencia del estadístico de Bartlett bajo censura generada con 10,000 pares de muestras de una población con distribución normal a diferentes tamaños de muestra y niveles de censura especificados en la leyenda.

6.2. Tamaño de la prueba

El tamaño de la prueba se emplea para medir la capacidad de discriminación de una estadística de prueba cuando H_0 es verdadera. Es decir, con ella sabremos la probabilidad de la nueva prueba de Bartlett para

6.3. Determinación de la robustez de la prueba

muestras censuradas cuando ambos grupos tienen varianzas iguales sea $\leq \alpha$.

El tamaño de la prueba está definido por:

$$\beta(\underline{\theta}) = \Pr (\text{Rechazar } H_0 \mid H_0 \text{ es verdadera}) \quad (6.4)$$

Por lo que la prueba h es de tamaño α si $Pr \{\text{Error tipo I usando } h\} \leq \alpha$. El juego de hipótesis es nuevamente (6.3). El procedimiento para conocer el tamaño de la prueba fue idéntico que para la potencia de la prueba bajo la H_0 . En este caso $\alpha = 0.01, 0.025, 0.05, 0.1$.

Tabla 6.1: Estimaciones del tamaño de la prueba

Condiciones		Nivel de significancia α			
n_1, n_2	$\%c_1, \%c_2$	0.01	0.025	0.05	0.10
10,10	10,10	0.0100	0.0251	0.0500	0.1000
20,20	10,10	0.0100	0.0251	0.0502	0.1001
100,100	10,10	0.0100	0.0250	0.0502	0.1000
100,100	20,20	0.0101	0.0248	0.0499	0.1003
100,100	50,50	0.0100	0.0252	0.0501	0.1002
100,100	30,30	0.0101	0.0250	0.0501	0.1001
70,100	0,20	0.0100	0.0250	0.0500	0.1001
70,40	30,20	0.0101	0.0252	0.0500	0.1002
50,40	30,20	0.0101	0.0251	0.0499	0.1000
30,80	50,40	0.0101	0.0251	0.0501	0.0998

6.3. Determinación de la robustez de la prueba

Lim y W.Y. (1996) and Conover *et al.* (1981) describen el proceso de una prueba de robustez a un nivel de significancia 5% en un experimento de simulación para comparar diferentes pruebas de igualdad de varianzas, entre ellas la prueba de Bartlett usando varias distribuciones simétricas y asimétricas. Las distribuciones empleadas en este caso son: Normal, t student, Uniforme, Exponencial y Lognormal, además la Logística y Laplace (doble exponencial). Éstas representan un rango de simétricas a asimétricas, de colas ligeras a pesadas y de baja a

6.3. Determinación de la robustez de la prueba

elevada kurtosis [Lim y W.Y. \(1996\)](#). Las pruebas se corrieron en tres categorías: 1) cuando $n_1 = n_2$ & $\%c_1 = \%c_2$; 2) cuando $n_1 = n_2$ & $\%c_1 \neq \%c_2$; y 3) cuando $n_1 \neq n_2$ & $\%c_1 \neq \%c_2$. Los valores respectivos fueron: $n_1 = 10, 20$ y $n_1 = 10, 30$ & $n_2 = 20$ en el tercer grupo, para $\%c_1 = 0, 10, 30$ & $\%c_2 = 20$.

Las distribuciones se dividieron en dos partes, la primera conformada por las distribuciones uniforme, exponencial, lognormal y normal con una razón de varianzas bajo H_0 en 1:1 y verificando la potencia bajo H_1 en 1:2 y 1:3. La segunda parte está conformada por las distribuciones logística, t student, laplace y normal como punto de referencia; la razón de varianzas bajo H_0 es 1:1 y verificando la potencia bajo H_1 a 1:1+0.5 y 1:2. El proceso de prueba para la robustez fue similar que en [6.1](#). En las tablas [6.2](#) y [6.4](#) se muestran las varianzas empleadas en cada distribución. Si el valor obtenido es ≤ 0.05 entonces la prueba es robusta para esa distribución.

Como resultado, la prueba es robusta solamente para la distribución uniforme en los 18 pares de grupos planteados, sin embargo, las distribuciones simétricas son mayores en potencia que la normal, la cual aumenta conforme crece el tamaño de muestra y disminuye con el aumento de la censura. La tablas [6.3](#) y [6.5](#) muestran los resultados obtenidos.

Tabla 6.2: Valores de σ^2 empleadas en las distribuciones Uniforme, Exponencial, Lognormal y Normal

Identificador	Uniforme	Exponencial	Lognormal	Normal
1	0.083	1	255.016	1
2	0.167	2	510.031	2
3	0.250	3	765.047	3

Tabla 6.3: Tamaños empíricos y potencia ($\alpha = 0.05$) en las distribuciones Uniforme, Exponencial, Lognormal y Normal

Condiciones			Distribuciones			
n_1, n_2	$\%c_1, \%c_2$	σ_1^2, σ_2^2	Uniforme	Exponencial	Lognormal	Normal
10,10	0,0	1:1	0.0124	0.2227	0.3901	0.0500
10,10	0,0	1:2	0.0746	0.3033	0.4316	0.1536
10,10	0,0	1:3	0.2585	0.4057	0.4644	0.3281

Sigue ...

6.3. Determinación de la robustez de la prueba

Tabla 6.3: Tamaños empíricos y potencia ($\alpha = 0.05$) (continuación)

Condiciones			Distribuciones			
n_1, n_2	$\%c_1, \%c_2$	σ_1^2, σ_2^2	Uniforme	Exponencial	Lognormal	Normal
10,10	10,10	1:1	0.0186	0.1884	0.2979	0.0500
10,10	10,10	1:2	0.0782	0.2653	0.3323	0.1409
10,10	10,10	1:3	0.2258	0.3686	0.3594	0.2823
10,10	30,30	1:1	0.0350	0.1498	0.1902	0.0500
10,10	30,30	1:2	0.0882	0.2062	0.2147	0.1086
10,10	30,30	1:3	0.1769	0.2880	0.2289	0.2009
20,20	0,0	1:1	0.0041	0.2778	0.4887	0.0500
20,20	0,0	1:2	0.0514	0.4251	0.5419	0.3107
20,20	0,0	1:3	0.3452	0.5925	0.5761	0.6519
20,20	10,10	1:1	0.0180	0.2214	0.3515	0.0500
20,20	10,10	1:2	0.1947	0.3752	0.4009	0.2677
20,20	10,10	1:3	0.6047	0.5552	0.4360	0.5728
20,20	30,30	1:1	0.0287	0.1802	0.2256	0.0500
20,20	30,30	1:2	0.1726	0.2938	0.2535	0.1947
20,20	30,30	1:3	0.4166	0.4469	0.2788	0.4178
10,10	0,20	1:1	0.0198	0.2626	0.4203	0.0500
10,10	0,20	1:2	0.0777	0.1937	0.3932	0.1090
10,10	0,20	1:3	0.2425	0.2167	0.3833	0.2372
10,10	10,20	1:1	0.0223	0.1959	0.2906	0.0500
10,10	10,20	1:2	0.0714	0.2020	0.2867	0.1187
10,10	10,20	1:3	0.1986	0.2662	0.2963	0.2376
10,10	30,20	1:1	0.0316	0.1711	0.2295	0.0500
10,10	30,20	1:2	0.1475	0.2842	0.2749	0.1226
10,10	30,20	1:3	0.2043	0.3969	0.3058	0.2300
20,20	0,20	1:1	0.0113	0.3800	0.5996	0.0500
20,20	0,20	1:2	0.3057	0.2366	0.5299	0.2399
20,20	0,20	1:3	0.7049	0.2804	0.4284	0.5386
20,20	10,20	1:1	0.0133	0.2422	0.3564	0.0500
20,20	10,20	1:2	0.2012	0.2524	0.3426	0.2335
20,20	10,20	1:3	0.5723	0.3874	0.4091	0.5188
20,20	30,20	1:1	0.0247	0.2183	0.2841	0.0500
20,20	30,20	1:2	0.2071	0.4375	0.3641	0.2203
20,20	30,20	1:3	0.4853	0.6180	0.3641	0.4678
10,20	0,20	1:1	0.0157	0.2997	0.4686	0.0500
10,20	0,20	1:2	0.1608	0.2189	0.4288	0.1779
10,20	0,20	1:3	0.4507	0.2665	0.4143	0.3818
10,20	10,20	1:1	0.0191	0.2129	0.3159	0.0500
10,20	10,20	1:2	0.1356	0.2389	0.3172	0.1819
10,20	10,20	1:3	0.3508	0.3325	0.3281	0.3691
10,20	30,20	1:1	0.0335	0.1798	0.2467	0.0500
10,20	30,20	1:2	0.1384	0.3348	0.3085	0.1579
10,20	30,20	1:3	0.2888	0.4775	0.3440	0.3027
30,20	0,20	1:1	0.0098	0.4175	0.6600	0.0500
30,20	0,20	1:2	0.3694	0.2306	0.5760	0.2665
30,20	0,20	1:3	0.7767	0.2704	0.5387	0.5900
30,20	10,20	1:1	0.0119	0.2526	0.3835	0.0500
30,20	10,20	1:2	0.2470	0.2565	0.3545	0.2616
30,20	10,20	1:3	0.6594	0.4059	0.3498	0.5779

Sigue ...

6.3. Determinación de la robustez de la prueba

Tabla 6.3: Tamaños empíricos y potencia ($\alpha = 0.05$) (continuación)

Condiciones			Distribuciones			
n_1, n_2	$\%c_1, \%c_2$	σ_1^2, σ_2^2	Uniforme	Exponencial	Lognormal	Normal
30,20	30,20	1:1	0.0256	0.2221	0.2990	0.0500
30,20	30,20	1:2	0.2431	0.4930	0.3886	0.2609
30,20	30,20	1:3	0.5937	0.6855	0.4467	0.5536

Fin

Tabla 6.4: Valores de σ^2 empleadas en las distribuciones Logística, t student, Laplace y Normal

Identificador	Logística	t student	Laplace	Normal
1	13.160	1.500	8.000	1.000
2	13.166	2.000	8.500	1.500
3	26.319	3.000	16.000	2.000

Tabla 6.5: Tamaños empíricos y potencia ($\alpha = 0.05$) en las distribuciones Logística, t student, Laplace y Normal

Condiciones			Distribuciones			
n_1, n_2	$\%c_1, \%c_2$	σ_1^2, σ_2^2	Logística	t student	Laplace	Normal
10,10	0,0	1:1	0.0864	0.1068	0.1539	0.0500
10,10	0,0	1:1+0.5	0.0853	0.1357	0.1548	0.0805
10,10	0,0	1:2	0.1968	0.1852	0.2497	0.1536
10,10	10,10	1:1	0.0756	0.0890	0.1381	0.0500
10,10	10,10	1:1+0.5	0.0750	0.1125	0.1392	0.0812
10,10	10,10	1:2	0.1722	0.1449	0.2254	0.1409
10,10	30,30	1:1	0.0698	0.0842	0.1237	0.0500
10,10	30,30	1:1+0.5	0.0702	0.1054	0.1245	0.0689
10,10	30,30	1:2	0.1329	0.1301	0.1779	0.1086
20,20	0,0	1:1	0.1047	0.1344	0.1864	0.0500
20,20	0,0	1:1+0.5	0.1074	0.1862	0.1882	0.1369
20,20	0,0	1:2	0.3571	0.2822	0.3917	0.3107
20,20	10,10	1:1	0.0917	0.1056	0.1697	0.0500
20,20	10,10	1:1+0.5	0.0914	0.1406	0.1717	0.1207
20,20	10,10	1:2	0.3157	0.2074	0.3570	0.2677
20,20	30,30	1:1	0.0835	0.1027	0.1552	0.0500
20,20	30,30	1:1+0.5	0.0833	0.1353	0.1573	0.0952
20,20	30,30	1:2	0.2361	0.1914	0.2893	0.1947
10,10	0,20	1:1	0.0798	0.0941	0.1473	0.0500
10,10	0,20	1:1+0.5	0.0759	0.1085	0.1433	0.0644
10,10	0,20	1:2	0.0697	0.1303	0.1743	0.1090

Sigue ...

6.3. Determinación de la robustez de la prueba

Tabla 6.5: Tamaños empíricos y potencia ($\alpha = 0.05$) (continuación)

Condiciones			Distribuciones			
n_1, n_2	$\%c_1, \%c_2$	σ_1^2, σ_2^2	Logística	t student	Laplace	Normal
10,10	10,20	1:1	0.0796	0.0843	0.1338	0.0500
10,10	10,20	1:1+0.5	0.0757	0.1072	0.1325	0.0705
10,10	10,20	1:2	0.0696	0.1335	0.1958	0.1187
10,10	30,20	1:1	0.1337	0.0767	0.1252	0.0500
10,10	30,20	1:1+0.5	0.1457	0.1022	0.1269	0.0778
10,10	30,20	1:2	0.1465	0.1285	0.1921	0.1226
20,20	0,20	1:1	0.0962	0.1198	0.1882	0.0500
20,20	0,20	1:1+0.5	0.0847	0.1444	0.1797	0.1062
20,20	0,20	1:2	0.0848	0.2006	0.2795	0.2399
20,20	10,20	1:1	0.0933	0.1055	0.1695	0.0500
20,20	10,20	1:1+0.5	0.0853	0.1342	0.1702	0.1062
20,20	10,20	1:2	0.0864	0.1987	0.3166	0.2335
20,20	30,20	1:1	0.2627	0.1034	0.1624	0.0500
20,20	30,20	1:1+0.5	0.2767	0.1357	0.1646	0.1109
20,20	30,20	1:2	0.2610	0.1933	0.3096	0.2203
10,20	0,20	1:1	0.0854	0.1403	0.1611	0.0500
10,20	0,20	1:1+0.5	0.0857	0.1636	0.1591	0.0930
10,20	0,20	1:2	0.2076	0.2228	0.2410	0.1779
10,20	10,20	1:1	0.0826	0.1161	0.1511	0.0500
10,20	10,20	1:1+0.5	0.0813	0.1556	0.1528	0.0978
10,20	10,20	1:2	0.2161	0.2216	0.2657	0.1819
10,20	30,20	1:1	0.0690	0.0831	0.1302	0.0500
10,20	30,20	1:1+0.5	0.0712	0.1169	0.1311	0.0932
10,20	30,20	1:2	0.1835	0.1762	0.2355	0.1579
30,20	0,20	1:1	0.0924	0.1182	0.1850	0.0500
30,20	0,20	1:1+0.5	0.0912	0.1410	0.1801	0.1067
30,20	0,20	1:2	0.2674	0.1983	0.2694	0.2665
30,20	10,20	1:1	0.0889	0.1038	0.1656	0.0500
30,20	10,20	1:1+0.5	0.0885	0.1417	0.1656	0.1130
30,20	10,20	1:2	0.2904	0.2074	0.3173	0.2616
30,20	30,20	1:1	0.0914	0.1069	0.1720	0.0500
30,20	30,20	1:1+0.5	0.0917	0.1403	0.1740	0.1170
30,20	30,20	1:2	0.2894	0.2039	0.3255	0.2609

Fin

Capítulo 7

Metodología para determinar homogeneidad de varianzas en muestras censuradas empleando la prueba de Levene

Sea $x_1, x_2, \dots, x_{k_1}, \dots, x_{n_1}$ una muestra censurada $(n_1 - k_1)$ de tamaño n_1 y sea $y_1, y_2, \dots, y_{k_2}, \dots, y_{n_2}$ otra muestra censurada $(n_2 - k_2)$ de tamaño n_2 ambas provenientes de una población con distribución $N(\mu, \sigma^2)$. Sea e_{ij}^* el valor residual censurado o no censurado que se empleará en (3.3) que sustituirá a (3.2) los residuales de las muestras completas, para verificar $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$. El desarrollo de la obtención de los residuales se desglosa a continuación.

7.1. Estimación de los residuales de muestras censuradas por la derecha

Hillis (1995) desarrolla la metodología para la obtención de residuales de muestras censuradas por la derecha empleando para ello los estimadores para la regresión lineal con datos censurados desarrollados por Buckley y James (1979).

7.1.1. Estimadores de Buckley-James

En 1979 Jonathan Buckley e Ian James desarrollaron la obtención del estimador para la regresión lineal con datos censurados. La propuesta se describe a continuación. [Buckley y James \(1979\)](#).

En una muestra censurada, típicamente somos incapaces de observar las variables de interés, los tiempos de supervivencia y_1, \dots, y_n y en su lugar observamos:

$$z_i = \min(y_i, t_i); i = 1, 2, \dots, n$$

dónde t_1, \dots, t_n es el valor de censura, junto con la variable indicadora

$$\delta_i = \begin{cases} 1 & \text{si } y_i \leq t_i \text{ (observados)} \\ 0 & \text{si } y_i > t_i \text{ (censurados)} \end{cases}$$

Consideramos el modelo de regresión lineal

$$y_i = \alpha + \beta x_i + \epsilon_i \tag{7.1}$$

Donde las ϵ_i son i.i.d. con función de distribución no especificada F , media cero y varianza finita.

El modelo de regresión lineal que considera los datos censurados es definido como:

$$y_i^* = y_i \delta_i + E(y_i | y_i < t_i)(1 - \delta_i) \tag{7.2}$$

Entonces $E(y_i^*)$ es igual a (7.2) de modo que:

$$E \left\{ \sum_{i=1}^n (x_i - \bar{x})(y_i^* - \beta x_i) \right\} = 0$$

Por lo que se debe elegir como estimador de la pendiente el valor b de modo que:

7.1. Estimación de los residuales de muestras censuradas por la derecha

$$\sum_{i=1}^n (x_i - \bar{x})(y_i^* - bx_i)$$

A partir de que $E(y_i|y_i > t_i)$ es desconocido, se adopta un estimador autoconsistente aproximado y se reemplazan las observaciones censuradas por:

$$\bar{y}_i(b) = bx_i + \sum_{i=k}^n w_{i,k}(b)(y_k - bx_k) \quad (7.3)$$

Donde

$$w_{i,k}(b) = \begin{cases} v_k(b) / \left\{ 1 - \hat{F}_{0,b}(t_i - bx_i) \right\} & (e_i(0, b) < e_k(0, b)) \\ 0 & \text{d.o.m.} \end{cases} \quad (7.4)$$

En [Stare et al. \(2000\)](#) usan el método de Buckley-James y reescriben (7.3) como:

$$(\epsilon_i | \epsilon_i > t_i - (\alpha + \beta x_i)) = \int_{t_i - (\alpha + \beta x_i)}^{\infty} \epsilon \frac{dF}{1 - F(t_i - (\alpha + \beta x_i))} \quad (7.5)$$

Donde F es la función de distribución de ϵ . Después de sustituir por F su estimador \hat{F} Kaplan-Meier citado por [Stare et al. \(2000\)](#) (uno menos el usual estimador Kaplan-Meier de la función de supervivencia), se tiene entonces:

$$y_i^* = y_i \delta_i + \left[\beta' x_i + \frac{\sum_{\epsilon_k > \epsilon_i} w_k \epsilon_k}{1 - F(\epsilon_i)} \right] (1 - \delta_i) \quad (7.6)$$

Donde w_k son pasos de \hat{F} . Observando (7.6) un estimador razonable sería

$$\hat{\beta} = \frac{(X - \bar{x})' y^*(\hat{\beta})}{(X - \bar{x})'(X - \bar{x})} \quad (7.7)$$

7.1. Estimación de los residuales de muestras censuradas por la derecha

Reemplazando (7.6) por sus estimadores y tomando en cuenta que la estimación depende de β , se necesitan iteraciones. Finalmente se tiene que

$$\alpha = \bar{y}^* - \beta' \bar{x} \quad (7.8)$$

7.1.2. Estimación de los residuales bajo censura

Hillis (1995) propuso residuales para modelos lineales cuando la variable respuesta no es exactamente observada sino censurada por la derecha. Él considera el modelo siguiente:

$$t_i = \beta z_i + \epsilon_i; i = 1, 2, \dots, n \quad (7.9)$$

Las ϵ_i son independientes e indenticamente distribuidas con función de distribución F . t_i es el tiempo de supervivencia del i -ésimo individuo. z_i es el valor de las covariables correspondientes. El tiempo censurado para t_i es denotado como δ_i con el supuesto de que la distribución de las ϵ_i no depende del valor de z_i o c_i . Los datos observados para el modelo (7.9) es el triple (y_i, z_i, δ_i) donde $y_i = \min(t_i, c_i)$ y $\delta_i = I(t_i \leq c_i)$ donde I es la función indicadora.

Para el desarrollo de los residuales en este contexto, Hillis (1995) define una secuencia de variables aleatorias.

$$\epsilon_i^* = \delta_i \epsilon_i + (1 - \delta_i) U_i \quad (7.10)$$

Donde cada U_i viene de la función de distribución F_i definida como

7.1. Estimación de los residuales de muestras censuradas por la derecha

$$F_i(x) = P(\epsilon_i \leq x | \epsilon_i > c_i - z_i\beta) = \begin{cases} 0 & : x \leq c_i - z_i\beta \\ \frac{F(x) - F(c_i - z_i\beta)}{1 - F(c_i - z_i\beta)} & : x > c_i - z_i\beta \end{cases} \quad (7.11)$$

Esto significa que (7.10) igual a $\epsilon : i$ para una observación no censurada, y (7.10) es igual a una observación generada aleatoriamente de la distribución condicional de ϵ_i dado que $\epsilon_i > c_i - z_i\beta$.

Hillis muestra que los ϵ_i^* tiene la misma distribución conjunta que los ϵ_i y sugiere reemplazar β y F con sus estimadores para definir los residuales del modelo (7.9)

$$\hat{e}_i^* = \delta_i \hat{e}_i + (1 - \delta_i) \hat{u}_i \quad (7.12)$$

Donde \hat{u}_i son observaciones generadas aleatoriamente de la distribución \hat{F}_i definida por:

$$\hat{F}_i(x) = \begin{cases} 0 & : x \leq c_i - z_i\hat{\beta} \\ \frac{\hat{F}(x) - \hat{F}(c_i - z_i\hat{\beta})}{1 - \hat{F}(c_i - z_i\hat{\beta})} & : x > c_i - z_i\hat{\beta} \end{cases} \quad (7.13)$$

En esta expresión \hat{F} es el estimador del producto-límite basado en los residuales censurados y no censurados $\hat{e}_i = y_i - z_i\hat{\beta}$, y $\hat{\beta}$ es el estimador Buckley-James para el parámetro β . Si los supuestos del modelo son correctos, la gráfica de los \hat{e}_i^* contra la variable independiente o los valores ajustados exhiben una ubicación aleatoria.

7.2. Distribución del estadístico de Levene bajo censura

Por medio del programa R usando el paquete `rms` se encontró el valor de \hat{e}_i^* y empleando la simulación Monte Carlo a 10,000 replicas se encontró la distribución que sigue el estadístico de Levene cuando dos muestras se encuentran bajo censura por la derecha. El procedimiento se describe en el Apéndice. En la figura 7.1 se grafica el comportamiento de esta prueba bajo el efecto de la censura.

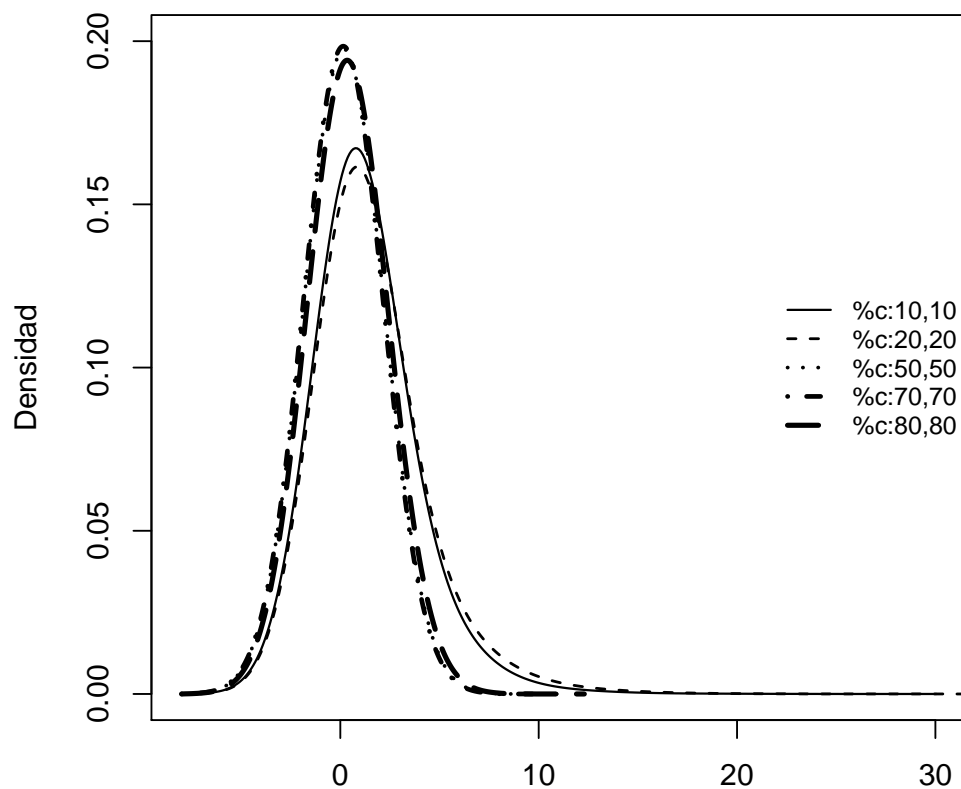


Figura 7.1: Gráfico comparativo de la distribución del estadístico de Levene para dos muestras aleatorias censuradas de tamaño $n_1 = n_2 = 50$ a diferentes niveles de censura.

Capítulo 8

Ejemplos de aplicación

La producción pecuaria nacional, particularmente la de los rumiantes, ante los retos y oportunidades que significa la globalización en el comercio mundial, requiere necesariamente del desarrollo e implementación de nuevas estrategias de producción acordes a las demandas de la sociedad con el fin de competir en igualdad de condiciones con otros países.

Actualmente en el laboratorio de reproducción de ovinos y caprinos (LaROCa) del Colegio de Postgraduados se desarrollan diferentes estrategias para mejorar la actividad reproductiva de los rumiantes, para lograr una producción animal *verde, limpia y ética*. La primera estrategia tiene la finalidad de manipular los eventos reproductivos, a través de *efectos socio-sexuales*, para inducir la ovulación sincronizada en las hembras que, de otra manera, serían anovulatorias (anestro estacional y periodo postparto). La segunda estrategia se basa en el *manejo del amamantamiento* en la cual se manipula la relación madre-cría, con el objetivo de que la madre restablezca la actividad reproductiva lo más rápidamente posible después del parto. La tercera estrategia se basa en la *alimentación focalizada*, es decir, en el conocimiento de las respuestas fisiológicas a la nutrición. Esta estrategia tiene la finalidad de desarrollar programas de *corta* duración de complementos nutricionales balanceados exclusivamente para cada uno de los eventos reproductivos, como la producción de gametos, la sobrevivencia embrionaria y fetal, la programación fetal y la producción de calostro y leche [Martin](#)

8.1. Efecto macho en la activación del estro usando la prueba de Bartlett

et al. (2004).

La eficiencia reproductiva y por lo tanto el ingreso económico de las unidades de producción pecuaria, dependen de la cantidad de crías nacidas, en otras palabras, dependen de la fertilidad y junto con un buen manejo de las crías tener así una oferta de producto al mercado que garantice la permanencia de la unidad de producción M. (1999). En este sentido el estudio y análisis de los eventos fisiológicos que inducen la actividad reproductiva en los ovinos tales como: el efecto macho, la aplicación de selenio y manejo de las señales neuroendócrinas son algunos de los métodos para elevar la productividad ganadera. Este tipo de casos tienen la peculiaridad de que algunas observaciones pueden ser censuradas y verificar el supuesto de homogeneidad de varianzas es una herramienta para su análisis.

8.1. Efecto macho en la activación del estro usando la prueba de Bartlett

Cruz (2011) evaluó el efecto macho en ovejas de la raza Pelibuey con la finalidad de inducir la reactivación reproductiva lo más pronto posible después del parto. Es decir, reducir el intervalo entre partos e incrementar el número de corderos nacidos como respuesta al efecto macho observado en la aceptación al apareamiento por la hembra.

El experimento fue un ANOVA de una vía con dos niveles, formando así dos grupos: sin efecto macho y con efecto macho respectivamente. Los datos fueron censurados después de la observación 63ma. de un total de 73 respuestas esperadas de inicio del estro en ovejas. El tamaño del grupo uno fue $n_1 = 37$ y el tamaño del grupo dos $n_2 = 36$. En el grupo uno la posición $k_1 = 27$, así que fueron $n_1 - k_1$ valores censurados, en el cambio el grupo dos fue completo. Los datos de los grupos se encuentran en la tabla 8.1.

El procedimiento para verificar si $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$ es obtener $\hat{\sigma}_1^2$ & $\hat{\sigma}_2^2$, el estimador de máxima verosimilitud de la varianza

8.1. Efecto macho en la activación del estro usando la prueba de Bartlett

de cada grupo usando el paquete FITDISTRPLUS en R como se describe en el Apéndice. Después, usamos la ecuación (3.1) para tener el valor observado del estadístico de prueba. Este valor será comparado contra el valor crítico de tablas. Para obtener este valor, nos referimos a la tabla 4.2 donde el tamaño de muestra y el porcentaje de censura sean lo más cercano a los valores de la muestra redondeando al inmediato superior. En este caso el valor crítico correspondiente lo encontramos en $n_1 = n_2 = 30$, $\%c_1 = 0$ y $\%c_2 = 20$, cabe hacer notar que el orden de los grupos no afecta. Así que, en este caso, el valor crítico a un nivel de significancia $\alpha = 0.05$ es 4.57, el cual no es muy lejano del valor obtenido por el método descrito en ?? mostrado en el cuadro 8.2. La regla de decisión es: si $\hat{h} \geq h$ entonces H_0 es rechazada.

Tabla 8.1: Datos del efecto macho en la entrada al estro en ovejas. $\delta = 1$ es valor observado & $\delta = 0$ es valor censurado.

Grupo 1	δ	Grupo 1	δ	Grupo 2	δ	Grupo 2	δ
28	1	54	1	23	1	30	1
29	1	56	1	24	1	31	1
30	1	59	1	25	1	31	1
31	1	67	1	26	1	31	1
31	1	72	1	26	1	32	1
32	1	74	1	27	1	32	1
32	1	74	1	27	1	33	1
33	1	75	1	27	1	34	1
34	1	75	0	28	1	36	1
35	1	75	0	28	1	37	1
35	1	75	0	28	1	37	1
37	1	75	0	28	1	37	1
37	1	75	0	28	1	38	1
40	1	75	0	28	1	38	1
42	1	75	0	29	1	38	1
45	1	75	0	29	1	39	1
47	1	75	0	30	1	43	1
48	1	75	0	30	1	43	1
48	1						

La nueva metodología muestra que $\hat{U} \geq h$, por lo que el grupo 1 y el grupo 2 no satisfacen el supuesto de homogeneidad de varianzas.

8.2. Efecto de neurotransmisores en la reproducción ovina usando la prueba de Levene

Tabla 8.2: Valores generados en el ejemplo del efecto macho en la reproducción ovina.

	Valor	Descripción
$\hat{\theta}_1$	457.27	Varianza del grupo 1
$\hat{\theta}_2$	26.46	Varianza del grupo 2
\hat{h}	54.62	Valor observado del estadístico
h	4.84	Valor crítico de la distribución a $\alpha = 0.05$

8.2. Efecto de neurotransmisores en la reproducción ovina usando la prueba de Levene

El control de los eventos reproductivos también se puede hacer desde un nivel cerebral por medio de los neurotransmisores, entre éstos la L-glutamina es uno de los más empleados. En este sentido [Rodríguez~M. \(2011\)](#) utilizó un tratamiento en ovejas de raza Pelibuey en las cuales aplicaron esta sustancia esperando una reducción del tiempo en el que las ovejas mostraran receptividad al macho y un incremento en la actividad ovárica. La respuesta clave de que la oveja ha entrado nuevamente en su proceso de fertilidad es por la detección del celo. En este caso el tiempo máximo de observación y determinación de censura fue de 100 días.

Se tuvieron dos grupos con tamaños de $n_1 = 25$ y $n_2 = 31$ el grupo 1 tuvo un valor $k_1 = 24$ y $n_2 - k_1 = 1$ valores censurados, el grupo 2 tuvo $k_2 = 27$ y $n_2 - k_2 = 4$ valores censurados. Para verificar si $H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$ aplicamos el procedimiento descrito en el capítulo 7 obtenemos los residuales con la ayuda del programa `rms` y el valor crítico de la distribución a un nivel de significancia $\alpha = 0,05$ (ver Apéndice). En la tabla 8.3 se muestran los datos de cada tratamiento con su respectivo residual.

La metodología por medio de la prueba de Levene indica que $\hat{W}_o = 3.43$ y el valor teórico $W_o = 5.61$, la regla de decisión es: si $\hat{W}_o \geq W_o$ entonces H_0 es rechazada. En este caso \hat{W}_o no es $\geq W_o$, por lo que el grupo 1 y el grupo 2 satisfacen el supuesto de homogeneidad de varianzas.

8.2. Efecto de neurotransmisores en la reproducción ovina usando la prueba de Levene

Tabla 8.3: Datos del efecto del glutamato en reproducción ovina, $\delta = 1$ es valor observado & $\delta = 0$ es valor censurado.

Grupo 1			Grupo 2		
Valor	δ	e_{ij}^*	Valor	δ	e_{ij}^*
5.50	1	0.98	5.21	1	1.10
12.00	1	0.79	5.30	1	1.10
12.25	1	0.78	5.41	1	1.09
24.10	1	0.43	5.58	1	1.09
24.20	1	0.43	5.66	1	1.08
24.40	1	0.42	11.56	1	0.91
25.05	1	0.41	11.66	1	0.91
25.30	1	0.40	24.05	1	0.54
31.15	1	0.22	24.15	1	0.54
31.18	1	0.22	24.60	1	0.52
31.23	1	0.22	24.65	1	0.52
31.33	1	0.22	24.90	1	0.52
31.40	1	0.22	24.95	1	0.51
31.70	1	0.21	25.00	1	0.51
31.85	1	0.20	25.35	1	0.50
34.03	1	0.14	31.26	1	0.33
34.35	1	0.13	31.46	1	0.32
40.21	1	0.04	31.78	1	0.31
40.46	1	0.05	34.05	1	0.24
40.81	1	0.06	34.28	1	0.24
48.80	1	0.30	34.50	1	0.23
49.83	1	0.33	40.26	1	0.06
89.06	1	1.49	40.55	1	0.05
89.16	1	1.49	48.75	1	0.19
117.5	0	2.33	49.50	1	0.21
			74.4	1	0.95
			95.63	1	1.58
			117.5	0	2.23
			117.50	0	2.23
			117.5	0	2.23
			117.5	0	2.23

Capítulo 9

Conclusiones

- El estadístico de Bartlett tiene cambios considerables en su distribución cuando se presenta el caso censura.
- La prueba cumple con la propiedad de invarianza es decir no cambia el valor del estadístico de prueba ante cambios en los parámetros de la población.
- El comportamiento de la distribución es que a un nivel de censura constante y aumento en el tamaño de muestra, la media y la varianza disminuyen pero a un tamaño de muestra constante y aumento en el nivel de censura, estos valores se hacen mayores.
- Presenta un tamaño de prueba α para cualquier tamaño de muestra y nivel de censura.
- La potencia de la prueba indica que esta es consistente. Es decir, conforme aumenta el tamaño de muestra se aproxima mas a la unidad, sin embargo, reacciona ante los diferentes niveles de censura, ya que conforme la censura aumenta tarda mas en converger a 1.
- La prueba es robusta ante la distribución Uniforme, con las distribuciones simétricas la Logística es cercana al valor de significancia, sin embargo, es mucho mayor la potencia en las no simétricas bajo la hipótesis alterna ($H_1 : \sigma_1^2 \neq \sigma_2^2$).
- En muestras grandes, el estadístico de Bartlett converge a una distribución ji-cuadrada con $t - 1$ grados de libertad.

Capítulo 10

Apéndice

CÓDIGO EN R.

VALOR CRÍTICO DEL ESTADÍSTICO DE BARTLETT BAJO CENSURA POR LA DERECHA TIPO II

```
B = function(w,x,y,z)
nr = 10000

{

n1=w; n2=x; p1=y; p2=z; k1=p1*n1
k2=p2*n2; ncen1=n1-k1; ncen2=n2-k2

library(fitdistrplus)

for (i in 1:nr)
{

set.seed(i)
#grupo 1
gr1=rnorm(n1,2,1)
```

10. Apéndice

```
left=c(sort(gr1)[1:k1],rep(sort(gr1)[k1],ncen1))
right=c(sort(gr1)[1:k1],rep(NA,ncen1))
datg1=data.frame(left,right)
fnopt1=fitdistcens(datg1, "norm")
vari1[i]=(fnopt1$estimate[2])^2

grupo 2
gr2=rnorm(n2,2,1)
left=c( sort(gr2)[1:k2],rep( sort(gr2)[k2],ncen2))
right=c( sort(gr2)[1:k2],rep(NA,ncen2))
datg2=data.frame(left,right)
fnopt2=fitdistcens(datg2, "norm")
vari2[i]=(fnopt2$estimate[2])^2

}

I=2
gi=c(n1-1,n2-1)
g=sum(gi)
c=1+(1/(3*(I-1)))*(sum(1/gi)-(1/g))
bart=matrix(c(vari1,vari2),nr,2)
for(i in 1:nr)
{
sigmai=bart
sigma[i]=sum(gi*(sigmai[i,]))/g
pi[i]=prod(sigmai[i,]^gi)
Uc[i]=(1/c)*log((sigma[i]^g)/(pi[i]))
}
return(Uc)
}

#Entrada de condiciones
w= #:escribir el tamaño del grupo n1
x= #:escribir el tamaño del grupo n2
y= #:escribir en 00/100 los datos observados del grupo 1 p1
```

10. Apéndice

```
z= #:escribir 00/100 los datos observados del grupo 2 p2  
alpha=#:nivel de significancia
```

```
#Salida del valor crítico  
Ucc=B(w,x,y,z)  
pr=1-alpha  
cv=quantile(Ucc,pr)  
cv #:valor crítico
```

VARIANZAS Y VALOR OBSERVADO DEL ESTADÍSTICO DE BARTLETT BAJO CENSURA #Entrada de datos entre

los paréntesis

eg1=c(,) #: cada valor del grupo 1

eg2=c(,) #: cada valor del grupo 2

k1= #: número de datos observados del grupo 1

k2= #: número de datos observados del grupo 2

#Calculando

n1=length(eg1)

n2=length (eg2)

ncen1=n1-k1

ncen2=n2-k2

library(fitdistrplus)

gr1=eg1

left=c(sort(gr1)[1:k1],rep(sort(gr1)[k1],ncen1))

right=c(sort(gr1)[1:k1],rep(NA,ncen1))

datg1=data.frame(left,right)

fnopt1=fitdistcens(datg1, "norm")

vari1=(fnopt1\$estimate[2])^2

gr2=eg2

left=c(sort(gr2)[1:k2],rep(sort(gr2)[k2],ncen2))

right=c(sort(gr2)[1:k2],rep(NA,ncen2))

datg2=data.frame(left,right)

fnopt2=fitdistcens(datg2, "norm")

vari2=(fnopt2\$estimate[2])^2

#Calculando valor observado del estadístico de Bartlett

I=2

gi=c(n1-1,n2-1)

g=sum(gi)

c=1+(1/(3*(I-1)))*(sum(1/gi)-(1/g))

10. Apéndice

```
bart=matrix(c(vari1,vari2),1,2)
sigmai=bart
sigma=sum(gi*(sigmai))/g
pi=prod(sigmai^gi)
Uc=(1/c)*log((sigma^g)/(pi))

#Salida
vari1 #: variance of group 1
vari2 #: variance of group 2
Uc #: observed Bartlett's statistic
```

VALOR CRÍTICO DE LA PRUEBA DE LEVENE BAJO CENSURA POR LA DERECHA TIPO I

```
#Datos de entrada
nr=10000
le=function(w,x,y,z)

m1=w #tamaño de muestra grupo 1
m2=x #tamaño de muestra grupo 2
p1=y # porción de observados grupo 1
p2=z #porción de observados grupo 2
k1=p1*m1 #número de observaciones
k2=p2*m2
a=m1+m2

#Datos de salida
fc=matrix(0,nr,1)

library(rms)
```

10. Apéndice

```
for (i in 1:nr)

set.seed(i)
g1=sort(rnorm(m1,5,1))
g1c=pmin(g1,g1[k1+1]) #muestra censurada grupo 1
g1d=ifelse(g1>g1[k1],0,1) indicador de censura grupo 1

g2=sort(rnorm(m2,5,1))
g2c=pmin(g2,g2[k2+1]) muestra censurada grupo 2
g2d=ifelse(g2>g2[k2],0,1) indicador de censura grupo 2

gdata=c(g1c,g2c) #datos de entrada
delta=c(g1d,g2d)

Z=c(rep(1,m1),rep(0,a),rep(1,m2))
zi=matrix(Z,a,2,dimnames=list(c(1:a),c("b1","b2"))) matriz
de tratamientos

j=psm(Surv(gdata, delta) ~zi, dist="gaussian") regresión censurada
rc = resid(j, 'cens')residuales

nres=dim(rc)[1]
residuales=abs(c(rc[1:nres]))
residuales

nivel=c(rep(1,m1),rep(2,m2))
grupos=factor(nivel)
Lc=lm(residuales grupos) #corriendo anova modelo lineal con
censura
summary(Lc) #obtener estadístico F:prueba de Levene
```


10. Apéndice

```
mua=anova(Lc)
fc[i]=mua$F[1]

}
return(fc)
}
```

```
lf=le(w,x,y,z)
```

```
#w: tamaño del grupo 1
#x: tamaño del grupo 2
#y: razón de observados sobre tamaño del grupo 1
#z: razón de observados sobre tamaño del grupo 2
```

```
qf=quantile(lf,0.95)
qf
```

RESIDUALES VALOR OBSERVADO DEL ESTADÍSTICO DE LA PRUEBA LEVENE BAJO CENSURA

```
#Entrada de datos
```

```
g1r=c( , ) #introducir datos del grupo 1
g2r=c( , ) #introducir datos del grupo 1
m1=length(g1r)
m2=length(g2r)
k1= #número de datos observados del grupo 1
k2= #número de datos observados del grupo 2
a=m1+m2
```

```
library(rms)
```

```
g1dr=ifelse(g1r>g1r[k1],0,1) #indicador de censura grupo 1
```

```
g2dr=ifelse(g2r>g2r[k2],0,1) indicador de censura grupo 2
```

```
gdatar=c(g1r,g2r) #datos de entrada
```

```
deltar=c(g1dr,g2dr)
```

```
Z=c(rep(1,m1),rep(0,a),rep(1,m2))
```

```
zi=matrix(Z,a,2,dimnames=list(c(1:a),c("b1","b2"))) matriz de
tratamientos
```

```
j=psm(Surv(gdatar, deltar) ~zi, dist="gaussian") regresión
```

```
rc = resid(j, 'cens') #obtención de residuales
```

```
nres=dim(rc)[1]
```

```
residualesr=abs(c(rc[1:nres]))
```

```
residualesr
```

```
#Calculando el estadístico de Levene bajo censuran tipo I
```

10. Apéndice

```
nivel=c(rep(1,m1),rep(2,m2))
grupos=factor(nivel)
Lc=lm(residualesr grupos) #corriendo anova modelo lineal con
censura
summary(Lc) #obtener estadístico F:prueba de Levene
mua=anova(Lc)
fc1=mua$F[1]
```

Referencias

- Ali, M. M. y Chan, L. K. (1967). BAN linear estimates of parameters of normal distribution from censored samples. *Annals of the Institute of Statistical Mathematics*, 19, 3, 401–&.
- Altun, I., Cinar, N. D. y Walsh, A. (2011). Psychometric properties of the parents'fever management scale in a Turkish population. *Healthmed*, 5, 3, 567–575.
- Bartlett, M. (1937). Properties of sufficiency and statistical tests. *Proceedings of the royal society of london series a-mathematical and physical sciences*, 160, A901, 0268–0282. ISSN 0080-4630.
- Bhandary, M. y Dai, H. (2009). An Alternative Test for the Equality of Variances for Several Populations When the Underlying Distributions are Normal. *Communications in Statistics: Simulation and Computation*, 38, 109–117(9).
- Box, G. E. P. (1953). Non-normality and tests on variances. *Biometrika*, 40, 318–335.
- Brown, M. B. y Forsythe, A. B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, 69, 346, pp. 364–367. ISSN 01621459.
- Buckley, J. y James, I. (1979). Linear Regression with Censored Data. *Biometrika*, 66, 3, pp. 429–436. ISSN 00063444.
- Cadwell, J. H. (1953). Approximating to the distributions of measures of dispersion by a power of X^2 . *Biometrika*, 40, 336 – –346.
- Calsbeek, B. y Goodnight, C. J. (2009). Empirical comparison of g matrix test statistics: finding biologically relevant change. *Evolution*, 63, 10, 2627–2635.
- Casella, R. L., G.; Berger (1990). *Statistical inference*. Wadsworth Publishing Company.
- Cochran, W. G. (1941). The distribution of the largest of a set of estimated variances as a fraction of their total. *Annals of Human Genetics*, 11, 1, 47–52. ISSN 1469-1809.
- Conover, W., Johnson, M. y Johnson, M. (1981). A comparative-study of tests for homogeneity of variances, with applications to the outer continental-shelf bidding data. *Technometrics*, 23, 4, 351–361. ISSN 0040-1706.
- Cruz, E. F. (2011). *.Efecto macho su relación en anestro postparto en ovejas Pelibuey amantando..* Tesis de Maestría. Colegio de Postgraduados, Campus Montecillo.

Referencias

- Dasgupta, S. y De, U. K. (2010). A statistical study of pre-monsoon weather over south Bengal using descriptive and inferential techniques. *Mausam*, 61, 4, 547–552.
- Delignette-Muller, M. L., Pouillot, R., Denis, J.-B. y Dutang, C. (2011). Help to fit of a parametric distribution to non-censored or censored data: package 'fitdistrplus'.
- E., G. R. (2006). *Encyclopedia of statistical sciences*, tomo 5. John Wiley & Sons, Inc., New Jersey, U.S.A., segunda edición.
- Goto, F., Tsutsumi, T. y Ogawa, K. (2011). The Japanese version of the Dizziness Handicap Inventory as an index of treatment success: Exploratory factor analysis. *Acta otolaryngologica*, 131, 8, 817–825.
- Gupta, A. K. (1952). Estimation of the mean and standard deviation of a normal population from a censored sample. *Biometrika*, 39, 260–273.
- Hald, A. (1949). Maximum likelihood estimation of the parameters of a normal distribution wich is truncated at a known point. *Skandinavisk Aktuarietidskrift*, 119.
- Hartley, H. O. (1950). The maximum F-ratio as a short-cut test for heterogeneity of variance. 37, 308–312.
- Hillis, S. L. (1995). Residual plots for the censored data linear regression model. *Statistics in Medicine*, 14, 18, 2023–2036. ISSN 1097-0258.
- Ho, C.-H. y Romero, P. (2011). Using Asphalt Mixture Beams in the Bending Beam Rheometer Experimental and Numerical Approach. *Road materials and pavement design*, 12, 2, 293–314.
- Klein, J. P., Pelz, C. y Zhang, M.-j. (1999). Modeling Random Effects for Censored Data by a Multivariate Normal Regression Model. *Biometrics*, 55, 2, 497–506. ISSN 1541-0420.
- Kobayashi, K., Pillai, K. S., Guhatakurta, S., Cherian, K. M. y Ohnishi, M. (2011). Statistical tools for analysing the data obtained from repeated dose toxicity studies with rodents: a comparison of the statistical tools used in Japan with that of used in other countries. *Journal of environmental biology*, 32, 1, 11–16.
- Krasny, M., Wysocki, J., Zadurska, M. y Skarzynski, P. H. (2011). Relative nasopharyngeal patency index as possible objective indication for adenoidectomy in children with orthodontic problems. *International journal of pediatric otorhinolaryngology*, 75, 2, 250–255.
- Lai, T. L. y Ying, Z. (1991). Large sample theory of a modified buckley-james estimator for regression-analysis with censored-data. *Annals of Statistics*, 19, 3, 1370–1402.
- Lawless, J. F. (2003). *Statistical models and methods for lifetime data*. John Wiley & Sons, Inc., Hoboken, New Jersey, segunda edición.
- Levene, H. (1960). Robust tests for equality of variances. En *In contributions to probability and statistics (I. Olkin, ed.)*, 278–292. Stanford University Press, Palo Alto CA.
- Lim, T. y W.Y., L. (1996). A comparison of tests of equality of variances. *Computational Statistics & Data Analysis*, 22, 287–301.

Referencias

- Linder, R. y Nagaraja, H. (2003). Impact of censoring on sample variances in a bivariate normal model. *Journal of Statistical Planning and Inference*, 114, 1-2, 145–160.
- M., D. (1999). Examen de fertilidad para la selección en toros de carne. *Tecnovet*, 52.
- Martin, G., Milton, J., Davidson, R., Hunzicker, G. B., Lindsay, D. y Blache, D. (2004). Natural methods for increasing reproductive efficiency in small ruminants. *Animal Reproduction Science*, 82-83, 0, 231 – 245. ISSN 0378-4320. ¿ce:title¿Research and Practice III. 15th International Congress on Animal Reproduction¿/ce:title¿.
- Neyman, J. y Pearson, E. S. (1931). *Bulletin of the Academy Polish of Sciences Letter A*.
- Powers, J. y Powers, T. (1990). Statistical-analysis of pharmacokinetic data with special applications to bioequivalence studies. *Annales de recherches veterinaires*, 21, 1, S87–S92.
- Raqab, M. (1997). Modified maximum likelihood predictors of future order statistics from normal samples. *Computational Statistics & Data Analysis*, 25, 1, 91–106.
- Rodríguez M., J. (2011). *Efecto de la L-glutamina en la respuesta reproductiva de ovejas pelibuey*. Tesis Doctoral, Universidad Autónoma Chapingo.
- Sarhan, A. E. y Greenberg, B. G. (1955). Estimation of location and scale parameters by order statistics from singly and doubly censored samples .1. The normal distribution up to samples of size 10. *Ananals of Mathematical Statistics*, 26, 3, 536.
- Stare, J., Heinzl, H. y Harrell, F. (2000). On the use of Buckley and James least squares regression for survival data. *New approaches in applied statistics*, 16, 125–134.
- Struchalin, M. V., Dehghan, A., Witteman, J. C. M., van Duijn, C. y Aulchenko, Y. S. (2010). Variance heterogeneity analysis for detection of potentially interacting genetic loci: method and its limitations. *BMC genetics*, 11.
- Vaart A.W., V. D. (1998). *Asymptotics Statistics*. Press syndicate of the university of Cambridge.
- Yilmaz, I., Eyduran, E., Kaygisiz, A. y Javed, K. (2011). Estimates of Genetic Parameters for Lactation Shape Parameters with Multivariate Statistical Technique in Brown Swiss Cattle. *International journal of agriculture and biology*, 13, 2, 174–178.