



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN
CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

**PRUEBAS DE BONDAD DE AJUSTE DE BOOTSTRAP
PARA LA DISTRIBUCIÓN LOG-GAMMA
GENERALIZADA**

EDUARDO GUTIÉRREZ GONZÁLEZ

T E S I S

PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO

2010


La presente tesis titulada: **PRUEBAS DE BONDAD DE AJUSTE DE BOOTSTRAP PARA LA DISTRIBUCIÓN LOG-GAMMA GENERALIZADA**, realizada por el alumno: **EDUARDO GUTIÉRREZ GONZÁLEZ**, bajo la dirección del consejo particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

DOCTOR EN CIENCIAS

SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

CONSEJO PARTICULAR

CONSEJERO



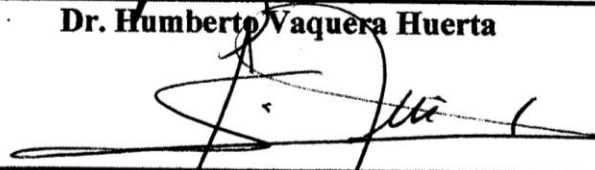
Dr. José A. Villaseñor Alva

ASESOR



Dr. Humberto Vaquera Huerta

ASESOR



Dr. Gilberto Rendón Sánchez

ASESOR



Dr. Filomena Ramírez Pérez

ASESOR



Dr. Barry C. Arnold

Montecillo, Texcoco, Estado de México, 21 de Enero de 2010

Pruebas de Bondad de Ajuste de Bootstrap para la Distribución Log-Gamma Generalizada

Eduardo Gutiérrez González, Dr.

Colegio de Postgraduados, 2010

Este trabajo trata sobre las pruebas de bondad de ajuste para la distribución log-gamma generalizada. El estudio se lleva a efecto siguiendo los pasos de las pruebas bootstrap de bondad de ajuste.

El trabajo se realiza en 4 capítulos, los dos primeros muestran un desarrollo analítico detallado sobre la distribución log-gamma y sus propiedades, mientras que en los dos últimos capítulos se hace el desarrollo por simulación del tamaño y potencia de la prueba. En el desarrollo analítico se formulan y demuestran varios resultados relacionados con la existencia y unicidad de tres estimadores para el parámetro de forma de la distribución log-gamma generalizada en su forma estándar. Además se realiza un estudio analítico detallado sobre el comportamiento asintótico de los estimadores propuestos, que debido a su complejidad de cálculo no se encuentran en la literatura existente actualmente.

Con respecto a la parte de simulación, se hace un estudio detallado con simulación Monte Carlo sobre las propiedades de los estimadores que fueron propuestos en el desarrollo analítico. Por medio de la técnica bootstrap y la invarianza con respecto a una transformación de los parámetros de localidad y escala, tanto de la distribución log-gamma generalizada (al fijar el valor del parámetro de forma) como del coeficiente de correlación, se propone una prueba de bondad de ajuste. Para la cual se calcula, el tamaño de la prueba y la potencia de la prueba para 20 distribuciones alternativas. La parte de simulación termina con una propuesta de una función en el proyecto R para realizar las pruebas de bondad de ajuste para la distribución log-gamma generalizada.

Finalmente, combinando algunos desarrollos analíticos y de simulación se realiza una aplicación a las cartas de control que se pueden utilizar en procesos de producción que tienen distribución log-gamma generalizada, o se pueden llevar a ésta por medio de la transformación, $Y = -X$ o $Y = \log(X)$, por ejemplo las distribuciones Weibull, Exponencial, Gumbel entre otras.

Palabras clave: Distribución log-gamma generalizada, parámetro de forma, pruebas de bondad de ajuste, pruebas de bootstrap, simulación Monte Carlo.

Goodness of Fit Tests of Bootstrap for generalized log-gamma distribution

Eduardo Gutiérrez González, Dr.

Colegio de Postgraduados, 2010

This paper deals with the goodness of fit tests for generalized log-gamma distribution. The study is carried out following the traditional steps for testing hypotheses.

Then the work is done in 4 chapters, the first two show a detailed analytical development log-gamma distribution and their property, while in the last two chapters is the development by simulating the size and power of the test.

In the analytical development are formulated and illustrated by several results related to the existence and uniqueness of three estimators for the shape parameter of generalized log-gamma standard distribution. Besides performing a detailed analytical study on the asymptotic behavior of estimators found, that due to its computational complexity not found in the literature today.

With respect to the simulation, a detailed study is done with Monte Carlo simulation on the properties of the estimators that were proposed in the analytical development. Using the bootstrap technique and taking advantage of invariance with respect to location and scale parameters of both the generalized log-gamma distribution and the correlation coefficient, we propose a goodness of fit test. Estimated for the size of the test and the power of the test for 20 alternative distributions. The simulation part ends by proposing a role in the project to undertake R testing goodness of fit for generalized log-gamma distribution.

Finally, some developments combining analytical and simulation is performed application site to the control charts that can be used in production processes that have distribution of generalized log-gamma, such as: Weibull, Exponential, Gumbel, etc..

Finally, some developments combining analytical and simulation is an application to the control charts that can be used in production processes with generalized log-gamma distribution, or it can be carried through the transformation $Y = -X$ o $Y = \log(X)$, for example the Weibull distribution, Exponential, Gumbel and others.

Key words: Generalized log-gamma standard distribution, shape parameter, goodness of fit tests, bootstrap test, Monte Carlo simulation.

Agradecimientos

Al consejo nacional de ciencia y Tecnología por el apoyo económico brindado para la realización de mis estudios de Doctor en Ciencias.

Al Colegio de Postgraduados por la oportunidad que me brindó.

Al Consejo particular integrado por:

- Dr. José A. Villaseñor Alva,
- Dr. Humberto Vaquera Huerta,
- Dr. Gilberto Rendón Sánchez,
- Dr. Filemón Ramírez Pérez,
- Dr. Barry C. Arnold,
- Dra. Elizabeth González Estrada,
- Dr. Paulino Pérez Rodríguez,

por el apoyo recibido y el tiempo dedicado a esta tesis.

Índice general

Índice de figuras	x
Índice de cuadros	xiii
Introducción	1
I.1. Descripción del problema	1
I.2. Objetivos	2
I.2.1. Objetivo general	2
I.2.2. Objetivos específicos	2
I.3. Justificación	2
I.4. Alcances y limitaciones del estudio	3
I.5. Resultados esperados	3
I.6. Metodología General	3
I.7. Reseña del estudio	4
1. Distribución Log-gamma generalizada y sus propiedades	5
1.1. Introducción	5
1.2. Función gamma	5
1.3. Familias gamma generalizada y log-gamma generalizada	7
1.4. Propiedades de la familia log-gamma generalizada	9
1.5. Propiedades de la familia gamma generalizada	17
1.6. Familias de localidad-escala	21
1.7. Coeficiente de asimetría para la log-gamma generalizada	22
2. Estimadores puntuales para el parámetro κ de la distribución $LGG(0, 1, \kappa)$	24

2.1.	Introducción	24
2.2.	Búsqueda de estimadores para la distribución $LGG(0, 1, \kappa)$	25
2.2.1.	EMV para los parámetros de la distribución $LGG(\xi, \sigma, \kappa)$	25
2.2.2.	EMV para el parámetro de forma de $LGG(0, 1, \kappa)$	27
2.2.3.	Estimador de momentos para el parámetro de forma de $LGG(0, 1, \kappa)$	32
2.2.4.	Estimador de momentos para la distribución $LGG(0, 1, \kappa)$ con transformación	33
2.3.	Comportamiento asintótico de los EMV	34
2.3.1.	¿Cómo determinar la distribución asintótica de $\hat{\kappa}_1$?	41
2.4.	Comportamiento asintótico del estimador de momentos	41
2.4.1.	¿Cómo determinar la distribución asintótica de $\hat{\kappa}_2$?	53
3.	Estudio de los estimadores de κ por Simulación	54
3.1.	Introducción	54
3.2.	La generación de números aleatorios	55
3.2.1.	Método de aceptación-rechazo, generación de la distribución $LGG(0, 1, \kappa)$	55
3.2.2.	Método de transformaciones, generación de la distribución $LGG(0, 1, \kappa)$	57
3.2.3.	Comentarios sobre la generación de muestras aleatorias de $LGG(0, 1, \kappa)$	57
3.3.	Métodos para calcular valores de una función inversa	58
3.3.1.	Método para calcular cuantiles de una distribución	59
3.3.2.	Método para calcular cuantiles de la distribución $LGG(0, 1, \kappa)$	60
3.4.	Estimadores para la simulación	61
3.4.1.	Estimador $\hat{\kappa}_1$ de máxima verosimilitud	61
3.4.2.	Estimador $\hat{\kappa}_2$ de momentos	62
3.4.3.	Estimador $\hat{\kappa}_3$ de momentos	63
3.5.	Propiedades asintóticas de los estimadores	63
3.5.1.	Comparación de los estimadores	63
3.5.2.	Comparación de los ECM	63
3.5.3.	ECM y varianza de los estimadores	65
3.6.	Propiedades de los estimadores en tamaños de muestra pequeños	68
3.6.1.	Comparación de los estimadores	68
3.6.2.	ECM y varianza de los estimadores	69
4.	Pruebas de bondad de ajuste	75
4.1.	Introducción	75

4.2. Una prueba para \mathcal{F}_{LGG}	76
4.3. Pruebas de bondad de ajuste bootstrap	77
4.3.1. Consistencia de las pruebas	78
4.4. Tamaño de la prueba	78
4.4.1. Metodología para determinar el tamaño de la prueba	79
4.4.2. Tamaños de las pruebas con base en el EMV $\hat{\kappa}_1$ y de momentos $\hat{\kappa}_2$	79
4.5. Potencias de las pruebas Bootstrap paramétrico	82
4.5.1. Metodología para determinar la potencia de las pruebas	82
4.5.2. Distribuciones alternativas para el estudio de la potencia de la prueba	83
4.5.3. Resultados de la Potencia de la prueba	84
4.6. Pruebas de bondad de ajuste para $LGG(\xi, \sigma, \kappa)$	88
4.6.1. Función en R para realizar la prueba de bondad de ajuste	88
4.7. Aplicaciones	90
4.7.1. Cartas de control \bar{X} y S bajo una política 3σ	90
4.7.2. Cartas de control \bar{X} y S para una distribución log-gamma generalizada	91
4.7.3. Metodología para las cartas de control \bar{X} y S , de una distribución $LGG(\xi, \sigma, \kappa)$	97
4.7.4. Problema	97
4.7.5. Ajuste de la curva	100
Conclusiones	101
C.1. Desarrollo analítico	101
C.2. Desarrollo por simulación	102
Bibliografía	103
B.1. Referencias	103
Anexos	107
A.1. Gráficas del Capítulo 1	107
A.2. Gráficas del Capítulo 2	109
A.3. Gráficas y cálculos por simulación del Capítulo 3	113
A.4. Gráficas y cálculos por simulación del Capítulo 4	116

Índice de figuras

1.1.	Función gama	6
1.2.	Función Log-gama generalizada estándar para $\kappa = 0.01, 0.02, \dots, 0.10$	10
1.3.	Función Log-gama generalizada estándar, anteriores y para $\kappa = 0.1, 0.2, \dots, 1$	10
1.4.	Función Log-gama generalizada estándar, anteriores y para $\kappa = 1, 10, 20 \dots, 100$	10
1.5.	Función Gama generalizada estándar para $\kappa = 0.1, 0.2, \dots, 1$ y $\kappa = 2, 3, \dots, 10$	19
1.6.	Función de asimetría $\gamma(x)$	23
2.1.	Con una muestra de tamaño $n = 1000$ y parámetro de forma $\kappa = 5$ y $\kappa = 0.5$, respectivamente: (a) Función log-verosimilitud, (b) Función de la derivada de la log-verosimilitud y (c) Función de la segunda derivada de la log-verosimilitud	31
2.2.	La figura izquierda para $\kappa = 0.1$ con $\hat{\kappa}_{1a} \sim N(0.1, 3.7788 \times 10^{-7})$ y $\hat{\kappa}_{1m} \sim N(0.1, 3.4408 \times 10^{-7})$. En la derecha para $\kappa = 0.5$ con $\hat{\kappa}_{1a} \sim N(0.5, 0.000161)$ y $\hat{\kappa}_{1m} \sim N(0.5, 0.000151)$, ambas aproximaciones se obtuvieron hasta el tercer momento.	39
2.3.	El número de la gráfica representa hasta que momento se usó para la aproximación de la distribución del EMV para $\kappa = 1$. La distribución $\hat{\kappa}_{1m} \sim N(1, 0.0010)$ y las aproximaciones: $\hat{\kappa}_{1a} \sim N(1, 0.012940)$, $N(1, 0.01194)$, $N(1, 0.00059)$, $N(1, 0.00049)$, $N(1, 0.00025)$ y $N(1, 0.00054)$, respectivamente.	40
2.4.	En la gráfica izquierda se utilizó hasta el sexto momento para $\kappa = 2$ con $\hat{\kappa}_{1a} \sim N(2, 0.0061)$ y $\hat{\kappa}_{1m} \sim N(2, 0.0075)$. En las otras dos gráficas representan la distribución para $\kappa = 5$ en la primera hasta el quinto momento con $\hat{\kappa}_{1a} \sim N(5, 0.1990)$ y $\hat{\kappa}_{1m} \sim N(5, 0.1189)$ en la gráfica de la derecha hasta el sexto momento con $\hat{\kappa}_{1a} \sim N(5, 0.0770)$ y $\hat{\kappa}_{1m} \sim N(5, 0.1199)$	40
2.5.	En la gráfica izquierda se utilizó hasta el sexto momento para $\kappa = 10$ con $\hat{\kappa}_{1a} \sim N(10, 1.945957)$ y $\hat{\kappa}_{1m} \sim N(10, 0.966055)$. En la gráfica de la derecha hasta el séptimo momento con $\hat{\kappa}_{1a} \sim N(10, 0.7297578)$ y $\hat{\kappa}_{1m} \sim N(10, 0.9814303)$	41

2.6.	Muestran las distribuciones aproximadas del parámetro κ con valores 0.1, 0.5, 1, 2, 5 y 10, obteniendo: $\hat{\kappa}_{2a} \sim N(0.0802, 0.000886)$ y $\hat{\kappa}_{2m} \sim N(0.0810, 0.000793)$; $\hat{\kappa}_{2a} \sim N(0.5273, 0.000696)$ y $\hat{\kappa}_{2m} \sim N(0.5277, 0.000571)$; $\hat{\kappa}_{2a} \sim N(0.9441, 0.000482)$ y $\hat{\kappa}_{2m} \sim N(0.94460.000407)$; $\hat{\kappa}_{2a} \sim N(1.9256, 0.005927)$ y $\hat{\kappa}_{2m} \sim N(1.9275, 0.003634)$; $\hat{\kappa}_{2a} \sim N(5.0016, 0.076203)$ y $\hat{\kappa}_{2m} \sim N(5.0095, 0.038852)$; $\hat{\kappa}_{2a} \sim N(9.9790, 0.559955)$ y $\hat{\kappa}_{2m} \sim N(10.0027, 0.262725)$, respectivamente.	52
3.1.	Gráficas de la función que se debe maximizar para que la función envolvente sea una $N(\mu, \sigma^2)$, para valores de $\kappa = 0.2, 0.3, \dots, 1.2$, no tienen un máximo absoluto para generar números aleatorios de la distribución $LGG(0, 1, \kappa)$	56
3.2.	Generación de números aleatorios $LGG(0, 1, \kappa)$ por el método de la transformación, $\kappa = 0.5$: (a) $n = 500$, (b) $n = 1000$ y (c) $n = 10,000$ repeticiones. La línea negra es el resultado de la generación de números aleatorios, la otra línea (roja) la función de densidad $LGG(0, 1, 0.5)$	57
3.3.	Gráficas de las funciones homótéticas para resolver la ecuación $2e^{-80000x} - e^{-300000x} - 1 = 0$	59
3.4.	Valores de los ECM para cada estimador con un tamaño de muestra $n = 10,000$; $m = 5,000$ repeticiones; error máximo de 0.00001	64
4.1.	La figura de la izquierda muestra el histograma de los datos transformados y la distribución log-gamma que mejor ajusta; a la derecha se muestran en lugar del histograma su función de densidad empírica de los datos transformados.	98
4.2.	La figura de la izquierda muestra las cartas de control \bar{X} para los datos transformados y la gráfica de la derecha para los datos originales.	99
4.3.	Carta de control S para los datos transformados u originales.	99

Índice de cuadros

2.1.	Porcentajes de valores negativos según el valor de κ	27
3.1.	Porcentaje en 100,000 corridas con promedio de valores positivos	58
3.2.	Valores de los estimadores con tamaño de muestra 10,000 tabla izquierda (sin repeticiones) y en la tabla derecha se hacen 10 repeticiones.	64
3.3.	Valores de los ECM para cada estimador con un tamaño de muestra $n = 10,000$ y $m = 5,000$ repeticiones; error máximo en las aproximaciones de 0.00001 (ver gráfica 3.4).	64
3.4.	Valores de las varianza y ECM para el estimador $\hat{\kappa}_1$ con un tamaño de muestra n en miles; $m = 5,000$ repeticiones; error máximo de 0.00001 (ver gráfica 3.7).	66
3.5.	Valores de las varianza y ECM para el estimador κ_2 con un tamaño de muestra n en miles; $m = 5,000$ repeticiones; error máximo de 0.00001 (ver gráfica 3.7).	66
3.6.	Valores de las varianzas y ECM para los estimadores κ_3 con un tamaño de muestra n en miles; $m = 5,000$ repeticiones; error máximo de 0.00001 (ver gráfica 3.7).	67
3.7.	Valores de los ECM para el estimador $\hat{\kappa}_1$, $\hat{\kappa}_2$ y $\hat{\kappa}_3$ en los 5 valores de $\kappa = 0.1, 0.5, 1, 5, 10$, se tomaron con un tamaño de muestra n dada en miles y $m = 5,000$ repeticiones	67
3.8.	Valores de proporcionalidad entre ECM, tamaño de la muestra n en miles y el valor del parámetro κ	68
3.9.	Valores de estimación para cada valor κ y tamaño de muestra n proporcionado para los estimadores $\hat{\kappa}_1$ (tabla izquierda) y $\hat{\kappa}_3$ (tabla derecha). En cada valor se realizarón 10,000 repeticiones y se tomó el promedio para la estimación.	69
3.10.	Valores de estimación para cada valor κ y tamaño de muestra n proporcionado para el estimador $\hat{\kappa}_2$. En cada valor se realizarón 10,000 repeticiones y se tomó el promedio para la estimación.	69

3.11. Valores de la varianza y error cuadrado medio para cada valor κ y tamaño de muestra n proporcionados para el estimador $\hat{\kappa}_1$. En cada valor se realizarón 10,000 repeticiones y se tomó el promedio para la estimación. Las tablas de la derecha, representan: La de arriba el resultado de multiplicar el ECM por el tamaño de la muestra n y dividir entre el valor de κ , mientras que la tabla de abajo el cuadrado del sesgo. 72

3.12. Valores de la varianza y error cuadrado medio para cada valor κ y tamaño de muestra n proporcionados para el estimador $\hat{\kappa}_2$. En cada valor se realizarón 10,000 repeticiones y se tomó el promedio para la estimación. Las tablas de la derecha, representan: La de arriba el resultado de multiplicar el ECM por el tamaño de la muestra n y dividir entre el valor de κ , mientras que la tabla de abajo el cuadrado del sesgo. 73

3.13. Valores de la varianza y error cuadrado medio para cada valor κ y tamaño de muestra n proporcionados para el estimador $\hat{\kappa}_3$. En cada valor se realizarón 10,000 repeticiones y se tomó el promedio para la estimación. Las tablas de la derecha, representan: La de arriba el resultado de multiplicar el ECM por el tamaño de la muestra n y dividir entre el valor de κ , mientras que la tabla de abajo el cuadrado del sesgo. 74

4.1. Tamaños de prueba para $\kappa_1 = 0.1, 0.5, 1, 2, 5$, usando el EMV y tamaño de muestra $n = 30, 40, \dots, 200$ realizando $m = 1000$ y $M = 500$ repeticiones con aproximaciones de los cuantiles en un error menor a 0.00001 con 1000 iteraciones. 80

4.2. Tamaños de prueba para $\kappa_2 = 0.1, 0.5, 1, 2, 5$, usando el estimador de momentos tamaño de muestra $n = 30, 40, \dots, 200$ realizando $m = 1000$ y $M = 500$ repeticiones con aproximaciones de los cuantiles en un error menor a 0.00001 con 1000 iteraciones. 81

4.3. Potencias para la distribución alternativa normal con $\mu = 0$ y $\sigma = 0.2$ 84

4.4. Potencias para la distribución alternativa t-Student con $\nu = 5$ 85

4.5. Potencias para la distribución alternativa logística con $\mu = 0$ y $\beta = 2$ 85

4.6. Potencias para la distribución alternativa Laplace con $\mu = 0$ y $\sigma = 2$ 85

4.7. Potencias para la distribución alternativa Beta con $\alpha = 4$ y $\beta = 10$ 85

4.8. Potencias para la distribución alternativa Cauchy con $\theta = 0$ y $\sigma = 2$ 86

4.9. Potencias para la distribución alternativa Gumbel con $\alpha = 0$ y $\beta = 2$ 86

4.10. Potencias para la distribución alternativa gamma con $\alpha = 4$ y $\beta = 5$ 86

4.11. Potencias para la distribución alternativa Ji-cuadrada con $\nu = 20$ 86

4.12. Potencias para la distribución alternativa Weibull con $\alpha = 2$ y $\beta = 2$ 86

4.13. Potencias para la distribución alternativa F con $\nu_1 = 4$ y $\nu_2 = 8$ 87

4.14. Potencias para la distribución alternativa Fréchet con $\mu = 0, \alpha = 2$ y $\beta = 5$ 87

4.15. Potencias para la distribución alternativa Pareto con $\theta = 4$ y $\beta = 20$ 87

4.16. Potencias para la distribución alternativa exponencial con $\lambda = 0.1$	87
4.17. Niveles de confianza para la $LGG(0, 1, \kappa)$ con 3σ y cantidad de defectuosos por cada 10,000 producidos.	91
4.18. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.1)$, $\kappa = 0.1$	92
4.19. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.2)$, $\kappa = 0.2$	92
4.20. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.3)$, $\kappa = 0.3$	93
4.21. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.4)$, $\kappa = 0.4$	93
4.22. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.5)$, $\kappa = 0.5$	94
4.23. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.75)$, $\kappa = 0.75$	94
4.24. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 1)$, $\kappa = 1$	95
4.25. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 5)$, $\kappa = 5$	95
4.26. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 10)$, $\kappa = 10$	96
4.27. Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 50)$, $\kappa = 50$	96
4.28. Valores de las 10 muestras para medir la resistencia a la ruptura de las fibras de carbono.	98
4.29. Valores de las 10 muestras transformadas, medias y desviaciones estándar de la tabla 4.28.	99

Introducción

1.1. Descripción del problema

La distribución gamma a sido ampliamente utilizada desde su aparición, con ella se han modelado problemas en el tiempo, por ejemplo los problemas de tráfico en líneas telefónicas, Erlang (1900); los Flujos máximos, Markovic (1965); la resistencia de componentes del concreto reforzado, Tichy Varlietk, (1965); entre muchas otras más.

Con esta distribución y en el afán de obtener familias de densidades que sean más flexibles se crea la distribución gamma generalizada, introducida por primera vez en el trabajo de L. Amoroso en 1925 y retomada por Stacy E.W. en 1962. Desde entonces y debido a la importancia de la distribución se le ve un uso potencial en el estudio de confiabilidad, ver por ejemplo los trabajos de Cantú, Villaseñor and Arnold (2001), en otras aplicaciones ver Mees and Gerard (1984) y Ryuichi Kaneko (2003). Además se han hecho estudios de todos sus parámetros, ver Boerrigter (1998), Efthymios (2005), Gomes, Combes and Dussauchoy (2008), Harter (1967), Rao, Kantam and Narasimham (1991), Sreekumar and Thomas (2007), Stacy and Mihran (1965), Taguchi (1980), Young and Bakir (1987), Wingo (1987), Wong (1993). Todos ellos con diferentes enfoques, por ejemplo los últimos trabajos presentan resultados desde un enfoque computacional, como son los trabajos de Efthymios (2005) y Gomes (2008), mientras que otros utilizan estadísticas de orden, ver Khan (1983) y Sreekumar (2007).

Además de los trabajos que han realizado varios investigadores para estimar los parámetros de la distribución gamma generalizada, también se ha trabajado la parte de inferencia de la distribución, ver por ejemplo DiCiccio (1987), Hager and Bain (1970), Lawless (1980), entre otros.

Una dificultad grande en este tipo de distribuciones la constituye la estimación del parámetro de forma y sobre el que se han hecho investigaciones para determinar sus valores, pero resulta bastante complicado, que de hecho en algunos trabajos únicamente se estudian los otros dos parámetros, ver por ejemplo Malik (1968a).

Uno de los principales problemas al estudiar la distribución gamma generalizada reside en la estimación del parámetro de forma. Así, mediante una transformación surge la distribución log-gamma generalizada, que conserva el parámetro de forma, pero su soporte se amplía a todo \mathbb{R} . Luego, surge la distribución log-gamma generalizada, siendo de interés el poder decidir si

los datos de una muestra tienen un comportamiento log-gamma generalizado.

El modelo log-gamma generalizado que será revisado en el trabajo está dado por:

$$f(x; \xi, \sigma, \kappa) := \frac{\kappa^{\kappa-1/2}}{\sigma\Gamma(\kappa)} \exp\left(\left(\frac{x-\xi}{\sigma}\right)\sqrt{\kappa} - \kappa e^{(x-\xi)/\sqrt{\kappa}}\right), \quad -\infty < x < \infty. \quad (\text{I.1.1})$$

en donde, ξ es el parámetro de localidad, σ el parámetro de escala y κ el parámetro de forma. En 1925 Amoroso planteó una familia de distribuciones más general, ver Gavin (2007), que tiene como caso particular esta distribución dada en su forma inicial, que bajo una transformación se llega a la representación (I.1.1), los detalles se pueden ver en la sección 1.3.

Como puede apreciarse para realizar inferencia sobre la distribución log-gamma generalizada un problema sigue siendo la estimación del parámetro de forma, luego un primer punto para realizar la prueba de bondad de ajuste para la distribución log-gamma generalizada será trabajar con los estimadores del parámetro κ .

1.2. Objetivos

El trabajo tiene diferentes metas a desarrollar.

1.2.1. Objetivo general

Caracterizar y desarrollar inferencia estadística para la distribución log-gamma generalizada.

1.2.2. Objetivos específicos

- Probar la existencia de estimadores de máxima verosimilitud y momentos para el parámetro de forma de la distribución $LGG(0, 1, \kappa)$.
- Conocer las distribuciones asintóticas de los estimadores EMV y momentos del parámetro de forma de la distribución $LGG(0, 1, \kappa)$.
- Comprobar el insesgamiento asintótico y consistencia en ECM de los estimadores propuestos para el parámetro de forma de la distribución log-gamma generalizada en su forma estándar.
- Construir pruebas de bondad de ajuste de bootstrap para la distribución $LGG(\xi, \sigma, \kappa)$ con base en los EMV y el de momentos.
- Proponer una función en el proyecto R, que bajo una muestra aleatoria determine si ésta proviene de una distribución $LGG(\xi, \sigma, \kappa)$ e indique los estimadores de los parámetros.
- Aplicar los resultados obtenidos para la construcción de cartas de control con distribución $LGG(\xi, \sigma, \kappa)$.

1.3. Justificación

El problema sobre la prueba de bondad de ajuste para una distribución log-gamma generalizada, no ha sido analizado, por esta razón el problema a resolver en este trabajo está encaminado a proponer pruebas de bondad de ajuste para esta distribución. Las pruebas serán

basadas en la estadística del coeficiente de correlación, que ya ha sido empleada en otros tipos de distribuciones del tipo loc-escala, ver por ejemplo Pérez (2008). Para esto se utiliza la invarianza de las familias loc-escala, misma que cumple el coeficiente de correlación muestral.

I.4. Alcances y limitaciones del estudio

Con las técnicas estadísticas de las pruebas de bondad de ajuste basadas en el coeficiente de correlación, la simulación Monte Carlo y el método de bootstrap, se plantean metodologías que ayudan a proponer pruebas de bondad de ajuste, en el caso de que las estadísticas de prueba no se puedan obtener explícitamente en forma analítica. En el trabajo se analiza la distribución log-gamma generalizada, para la que se obtiene un estudio detallado sobre las pruebas bootstrap basadas en los dos estimadores propuestos de forma analítica.

En el trabajo se obtienen de forma analítica los comportamientos asintóticos para los dos estimadores propuestos.

El trabajo aunque tiene desarrollos analíticos está limitado a la comprobación si las pruebas son UMP de tamaño α , pero son realizadas las comprobaciones analíticas sobre la consistencia en error cuadrado medio y el insesgamiento asintótico de los estimadores, posteriormente esto queda comprobado por simulación. Además se comprueba que las pruebas basadas en el coeficiente de correlación para cada estimador son consistentes.

En el trabajo con base en los desarrollos analíticos se establecen las cartas de control para procesos de producción con distribución del tipo log-gamma generalizada.

I.5. Resultados esperados

Con el trabajo se espera dar respuesta a los objetivos planteados. Es decir, desarrollar y obtener el comportamiento asintótico de los estimadores para el parámetro de forma de la distribución log-gamma generalizada en su forma estándar. Además de desarrollar las metodologías para la prueba bootstrap con cada uno de los estimadores κ propuestos.

Para la parte práctica se espera programar en el proyecto R una función que muestre si una muestra dada proviene de la distribución $LGG(\xi, \sigma, \kappa)$ e indique los valores de los estimadores puntuales para cada parámetro. La construcción de las cartas de control para procesos de producción con distribución log-gamma generalizada o distribuciones que se pueden transformar a ésta.

I.6. Metodología General

La solución al problema será desarrollado de la siguiente forma.

- Estudiar en detalle las propiedades, tanto de la función gamma como de la distribución log-gamma generalizada.
- Revisar las propiedades de las funciones loc-escala.
- Obtener estimadores para el parámetro de forma κ . Comprobar que los estimadores son asintóticamente insesgados y consistentes en error cuadrado medio.
- Obtener las distribuciones asintóticas de los estimadores.

- Desarrollar programas en el proyecto R para la comprobación de las propiedades estadísticas de los estimadores propuestos.
- Desarrollar programas en el proyecto R para el tamaño y potencia de la prueba.
- Desarrollar las cartas de control para procesos de producción con distribución log-gamma generalizada.

I.7. Reseña del estudio

El trabajo está dividido en 4 capítulos que se desarrollan de la siguiente forma.

Capítulo 1. En esta parte del trabajo se revisan y desarrollan las propiedades de la función gamma y desarrollan algunas propiedades para las distribuciones gamma generalizada y su correspondiente distribución log-gamma generalizada. Se analizan las familias del tipo loc-escala y la propiedad de invarianza para este tipo de familias. Se hace énfasis en el desarrollo de los momentos para la distribución log-gamma generalizada, porque el material aquí revisado juega un papel muy importante para el desarrollo de los capítulos subsecuentes, pero en particular para el capítulo 2.

Capítulo 2. En esta parte del trabajo se efectúan desarrollos para probar la existencia y unicidad de los estimadores del parámetro κ , en donde se obtienen tres estimadores. Con los estimadores encontrados se estudia el comportamiento asintótico de éstos, como son el insesgamiento y consistencia en error cuadrado medio. Para tales efectos se requiere estudiar los momentos de la distribución log-gamma generalizada, tarea que no es sencilla por lo complicado de los cálculos. Para simplificar esto último se demuestran desarrollos generales tanto en el caso del estimador de máxima verosimilitud como el de momentos.

Capítulo 3. Ahora inicia la parte de simulación, aquí debido a las distribuciones del tipo analizadas, se estudia desde la generación de números aleatorios, cálculo de cuantiles, hasta la comprobación de que tipo de estimadores son. Revisando la consistencia en error cuadrado medio de los estimadores y la propiedad de insesgamiento asintótico.

Capítulo 4. Después de haber revisado las propiedades de los estimadores se pasa a la revisión de las pruebas de bondad de ajuste para ambos parámetros, para esto se desarrolla primeramente el estadístico de prueba basado en el coeficiente de correlación, con el que se calculan los tamaños de prueba y la potencia de la misma. Comprobando que ambas pruebas son consistentes. Finalmente con la ayuda de los resultados obtenidos se construyen las cartas de control para los procesos de producción de artículos con distribución log-gamma generalizada.

Distribución Log-gamma generalizada y sus propiedades

1.1. Introducción

En el presente capítulo se revisarán las propiedades que se tienen con referencia a la familia de funciones de densidad que son de interés para la investigación, es decir la familia log-gamma generalizada y por su vínculo la familia gamma generalizada, así como las familias de localización y escala.

El estudio iniciará con la función gamma y sus propiedades. Continuando con un bosquejo sobre la historia de las distribuciones gamma y log-gamma generalizadas, sus parámetros de localización, escala y forma.

Posteriormente se revisarán las familias loc-escala y en particular la familia log-gamma generalizada, explicando la importancia de estas familias en la simplificación del estudio de los parámetros de localización y escala, quedando únicamente el parámetro de forma.

Finalmente serán revisadas algunas estimaciones sobre los parámetros, en donde se encontrará una función para cuantificar el coeficiente de asimetría, que será utilizado para estimar al parámetro de forma.

Durante el estudio se darán algunos desarrollos sobre las propiedades más importantes de la distribución log-gamma generalizada y la función gamma.

1.2. Función gamma

La función gamma ha jugado un papel muy importante en diferentes esferas de las ciencias desde su aparición cuando Leonard Euler (1707-1783) escribió una carta a Christian Goldbach (1690-1764) en el año 1729, en la que hacía referencia a la función meromorfa

$$\Gamma(x) = \lim_{n \rightarrow \infty} n! n^x \prod_{i=0}^n \frac{1}{x+i}. \quad (1.2.1)$$

Posteriormente, Adrian María Legendre (1752-1833) propuso en 1814 llamar a dicha función, Gamma, para ese tiempo se conocía la representación que actualmente se usa para la función gama y fue encontrada por el mismo Euler

$$\Gamma(x) = \int_0^{\infty} z^{x-1} e^{-z} dz. \quad (1.2.2)$$

Por el uso de esta función su estudio ha sido muy amplio, entre sus propiedades principales se tienen las siguientes:

- 1).- Las función meromorfa $\Gamma(x)$ tiene en los enteros negativos los únicos puntos de discontinuidad. Esto se puede apreciar en (1.2.1) y en la figura 1.1

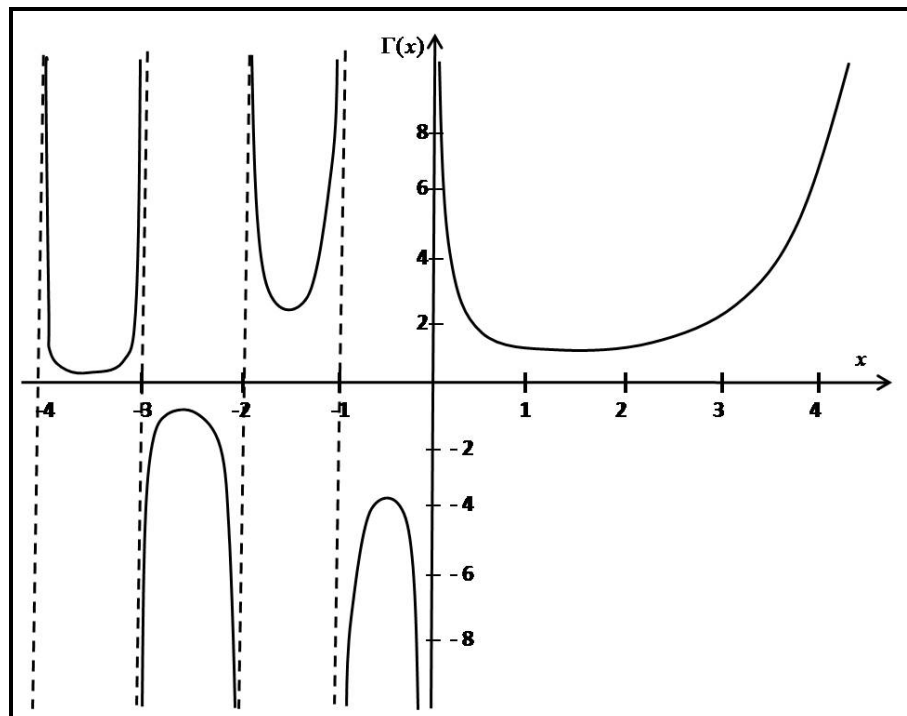


Figura 1.1: Función gama

- 2).- En general para $x > 0$, $\Gamma(x + 1) = x\Gamma(x)$.
 a).- En particular $\Gamma(1) = 1$.
 b).- En particular si $n \in \mathbb{N}$, entonces $\Gamma(n + 1) = n!$.
- 3).- Fórmula del complemento de Euler $\Gamma(x)\Gamma(1 - x) = \frac{\pi}{\sin(x\pi)}$ para toda $x \in \mathbb{R} - \mathbb{Z}$.
 a).- En particular $\Gamma(1/2) = \sqrt{\pi}$.
 b).- En particular $(\Gamma(x)\Gamma(1 - x))^2 = \left(\frac{\pi}{\sin(x\pi)}\right)^2 = \sum_{n \in \mathbb{Z}} \frac{1}{(x-n)^2}$ y $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.
- 4).- $\Gamma(x) = \frac{e^{-\gamma x}}{x} \prod_{m=1}^{\infty} \frac{e^{x/m}}{1+x/m}$, en donde $\gamma = \lim_{n \rightarrow \infty} \left(\sum_{m=1}^n \frac{1}{m} - \ln(n) \right) = 0.5772156649 \dots$, llamada constante de Euler.
- 5).- Comportamiento asintótico, $\Gamma(x) \sim \sqrt{2\pi} x^{x-1/2} e^{-x}$, se obtiene de la fórmula de Stirling.

6).- Para las proximidades de cero, se utiliza la representación: $\Gamma(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(x+n)} + \int_1^{\infty} z^{x-1} e^{-z} dz$.

Se obtiene de descomponer la integral de la función gamma en dos, una de cero a 1 y la otra de 1 a infinito. En la primera se sustituye la exponencial por su serie de Taylor y después se integra, la segunda no se toca.

a).- $\sum_{n=1}^{\infty} \frac{(-1)^n}{n!(x+n)}$ y $\int_1^{\infty} z^{x-1} e^{-z} dz$ son finitas alrededor del cero y convergen uniformemente, luego se pueden intercambiar el límite cuando $x \rightarrow 0^+$ y los operadores de suma y la integral, obteniendo $\lim_{x \rightarrow 0^+} \left\{ \sum_{n=1}^{\infty} \frac{(-1)^n}{n!(x+n)} + \int_1^{\infty} z^{x-1} e^{-z} dz \right\} = -\gamma$.

b).- Así, en las proximidades del cero $\Gamma(x) \doteq \frac{1}{x} - \gamma$.

c).- La constante de Euler también se obtiene de la integral $\int_{-\infty}^{\infty} x \exp(x - e^x) dx = -\gamma$.

7).- Las propiedades anteriores se cumplen si se cambian \mathbb{R} por \mathbb{C} .

8).- Relación entre las funciones gama y beta (fórmula de Dirichlet): $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, en

donde $B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt$.

9).- La función gamma incompleta, $\Gamma(\alpha) = \gamma(\alpha, x) + \Gamma(\alpha, x)$, para $x > 0$. En donde,

a).- $\Gamma(\alpha, x) = \int_x^{\infty} t^{\alpha-1} e^{-t} dt$ y cumple $\Gamma(\alpha + 1, x) = \alpha\Gamma(\alpha, x) + x^{\alpha} e^{-x}$.

b).- $\gamma(\alpha, x) = \int_0^x t^{\alpha-1} e^{-t} dt$ y cumple $\gamma(\alpha + 1, x) = \alpha\gamma(\alpha, x) - x^{\alpha} e^{-x}$.

10).- Derivadas del logaritmo natural de la función gama:

a).- La función Digamma $\psi(x) = \frac{d \log(\Gamma(x))}{dx} = -\gamma - \frac{1}{x} + \sum_{m=1}^{\infty} \frac{x}{m(x+m)}$. Se obtiene sustituyendo la representación en la propiedad (4) para la función gamma. Cuando la $x \approx 0$, entonces $\psi(x) \doteq -\frac{1}{x}$.

Cuando $x \rightarrow \infty$ entonces $\psi(x) \doteq -\gamma + \sum_{i=1}^{[x]} \frac{1}{i} \doteq \log(x)$ ($[x]$ parte entera de x).

b).- La función Trigamma $\psi'(x) = \frac{d^2 \log(\Gamma(x))}{dx^2} = \sum_{m=0}^{\infty} \frac{1}{(x+m)^2}$.

Cuando la $x \approx 0$, entonces $\psi'(x) \doteq \frac{1}{x^2}$.

c).- En general se llama n -polygamma $\psi^{(n-2)}(x) = \frac{d^{n-1} \log(\Gamma(x))}{dx^{n-1}} = \sum_{m=0}^{\infty} \frac{(-1)^{n-1} (n-2)!}{(x+m)^{n-1}}$,

para $n \geq 3$. Se obtiene al derivar sucesivamente la función digamma.

Cuando la $x \approx 0$, entonces $\psi^{(n-2)}(x) \doteq \frac{(-1)^{n-1} (n-2)!}{x^{n-1}}$.

1.3. Familias gamma generalizada y log-gamma generalizada

En el trabajo se revisará una familia de densidades llamada log-gamma generalizada con tres parámetros; localización $\xi \in \mathbb{R}$, escala $\sigma > 0$ y forma $\kappa > 0$, la familia está representada por la función de densidad dada en la expresión 1.3.1 y se denota \mathcal{F}_{LGG}

$$f(x; \xi, \sigma, \kappa) := \frac{\kappa^{\kappa-1/2}}{\sigma \Gamma(\kappa)} \exp \left(\left(\frac{x - \xi}{\sigma} \right) \sqrt{\kappa} - \kappa e^{\left(\frac{x - \xi}{\sigma} \right) / \sqrt{\kappa}} \right), \quad -\infty < x < \infty. \quad (1.3.1)$$

Mientras que su distribución correspondiente, gamma generalizada, estará dada por (1.3.2) y se denota \mathcal{F}_{GG} end donde α y β son parámetros de escala y κ el parámetro de forma, que coincide con el parámetro de forma de la distribución log-gamma generalizada

$$f(t; \alpha, \beta, \kappa) := \frac{\beta}{\alpha \Gamma(\kappa)} \kappa^{\kappa-1/2} \left(\frac{t}{\alpha}\right)^{\beta\sqrt{\kappa}-1} \exp\left(-\kappa \left(\frac{t}{\alpha}\right)^{\beta/\sqrt{\kappa}}\right), \quad t > 0, \alpha, \beta, \kappa > 0. \quad (1.3.2)$$

Note que (1.3.1) se puede obtener de (1.3.2) con la transformación $X = \log(T)$ y resulta que $\xi = \log(\alpha)$ y $\sigma = 1/\beta$.

Estas distribuciones se pueden obtener de diferentes formas, pero todas ellas a partir de la distribución gamma, $\Gamma(\kappa, 1)$, en donde κ sigue siendo el mismo parámetro de forma de (1.3.1) y (1.3.2). Sea Y una variable aleatoria con distribución $\Gamma(\kappa, 1)$, con $\kappa > 0$, es decir, $f_Y(y) = y^{\kappa-1}e^{-y}/\Gamma(\kappa)$, realizando la transformación $X = \sigma\sqrt{\kappa} \log\left(\frac{Y}{\kappa}\right) + \xi$, entonces $f_X(x) = f_Y(\phi(x))\phi'(x)$ en donde $y = \phi(x) = \kappa \exp\left((x - \xi)/\sigma\sqrt{\kappa}\right)$ se obtiene (1.3.1) con ξ y σ los parámetros de localidad y escala, respectivamente. Similarmente la distribución (1.3.2) se puede obtener directamente de la transformación $T = \alpha\left(\frac{Y}{\kappa}\right)^{\sqrt{\kappa}/\beta}$, entonces $f_T(t) = f_Y(\phi(t))\phi'(t)$ en donde $y = \phi(t) = \kappa\left(\frac{t}{\alpha}\right)^{\beta/\sqrt{\kappa}}$ se obtiene (1.3.2) con α y β parámetros de escala.

En lo referente a la historia de la distribución gamma generalizada, se puede encontrar que en el artículo de L. Amoroso titulado “Ricerche intorno alla curva dei redditi”, publicado en Ann. Mat. Pura Appl. en 1925, se tiene la distribución amoroso (1.3.3) que contiene a la distribución gamma generalizada como un caso particular cuando $\nu = 0$

$$f_T(t; \alpha, \kappa, \beta, \nu) = \frac{1}{\Gamma(\kappa)} \left|\frac{\beta}{\alpha}\right| \left(\frac{t - \nu}{\alpha}\right)^{\kappa\beta-1} \exp\left(-\left(\frac{t - \nu}{\alpha}\right)^\beta\right). \quad t \in \mathbb{R}, \nu, \alpha, \beta \in \mathbb{R}, \kappa > 0 \quad (1.3.3)$$

La distribución amoroso vuelve aparecer en el artículo de D’Addario titulado “Intorno alla curve dei redditi di Amoroso”, publicado en Econ. Fnanza en 1932.

Por medio de transformaciones hechas a partir de la distribución $\Gamma(\kappa, 1)$, de forma independiente a la distribución amoroso, algunos autores obtuvieron la distribución que llamaron gamma generalizada con función de densidad (1.3.4), obtenida con la transformación $T = \alpha\sqrt[\beta]{Y}$ en donde $Y \sim \Gamma(\kappa, 1)$ con $\alpha, \beta, \kappa > 0$

$$f_T(t) = \frac{1}{\Gamma(\kappa)} \left(\left(\frac{t}{\alpha}\right)^\beta\right)^{\kappa-1} e^{-(t/\alpha)^\beta} \left(\frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\right) = \frac{\beta}{\alpha} \frac{1}{\Gamma(\kappa)} \left(\frac{t}{\alpha}\right)^{\beta\kappa-1} e^{-(t/\alpha)^\beta}, \quad t > 0. \quad (1.3.4)$$

Definiendo de esta forma a la distribución gamma generalizada resulta que su correspondiente distribución log-gamma generalizada dada en (1.3.5) se obtiene de la transformación $X = \sigma \log\left(\frac{Y}{\kappa}\right) + \xi$ con $\sigma, \kappa > 0$ y $\xi \in \mathbb{R}$

$$f_X(x; \sigma, \xi, \kappa) = \frac{1}{\Gamma(\kappa)\sigma} \exp\left(\kappa\left(\frac{x - \xi}{\sigma}\right) - \exp\left(\frac{x - \xi}{\sigma}\right)\right), \quad x \in \mathbb{R}. \quad (1.3.5)$$

Note que (1.3.5) se puede obtener de (1.3.4) con la transformación $X = \log(T)$, resultando $\xi = \log(\alpha)$ y $\sigma = 1/\beta$.

El tipo de transformaciones que se han realizado fueron presentadas en el artículo de A.C. Olshen Titulado “Transformations of the Type III Distribution”, publicado en Annals of Mathematical Statistics en 1938. Otra aplicación de este tipo de transformaciones la podemos ver en la distribución que se aplica a $\Gamma(\kappa, \beta)$

$$f_T(t) = \frac{t^{\kappa-1}}{\Gamma(\kappa)\beta^\kappa} \exp\left(-\frac{t}{\beta}\right), \quad t > 0, \quad \beta > 0, \quad \kappa > 0. \quad (1.3.6)$$

Para obtener su correspondiente log-gamma distribución, por medio de la transformación $X = \log(Y) + \xi$ con $Y \sim \Gamma(\kappa, 1)$

$$f_X(x; \xi, \kappa) = \frac{\exp\left((x - \xi)\kappa - e^{x-\xi}\right)}{\Gamma(\kappa)}, \quad x \in \mathbb{R}, \quad \kappa > 0, \quad \xi \in \mathbb{R}. \quad (1.3.7)$$

Note que (1.3.7) se puede obtener de (1.3.6) con la transformación $X = \log(T)$ y se obtiene $\xi = \log(\beta)$.

Con las transformaciones realizadas a partir de la distribución $\Gamma(\kappa, 1)$ se puede observar que la inclusión de nuevos parámetros para hacer más flexible a la distribución y abarque una mayor gama de distribuciones, como son: gamma, ji-cuadrada, exponencial, Weibull, Gumbel, entre otras, acarrea problemas de identificación, por ejemplo la representación (1.3.4) es conocida como gamma generalizada y se atribuye a E.W.Stacy por su trabajo de 1962 en el que hizo una generalización de la distribución gamma, posteriormente en 1965 presentó otro trabajo para la estimación de los parámetros. Aunque como sugerencia del Dr. Barry C. Arnold se puede notar que la familia de distribuciones (1.3.1) coincide con la familia de distribuciones log-gamma (1.3.7) bajo algunos valores de los parámetros de localidad y escala.

Finalmente, el trabajo estará enfocado a la distribución que se ha llamado log-gamma generalizada en su forma estándar, es decir a (1.3.1) cuando $\xi = 0$ y $\sigma = 1$

$$f(z; \kappa) := \frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \exp\left(z\sqrt{\kappa} - \kappa e^{z/\sqrt{\kappa}}\right), \quad -\infty < z < \infty, \quad \kappa > 0. \quad (1.3.8)$$

1.4. Propiedades de la familia log-gamma generalizada

Algunas propiedades de la distribución $LGG(0, 1, \kappa)$ son las siguientes:

- 1).- Cuando $\kappa = 1$ coincide con la distribución Gumbel (ver subsección 4.5.2), bajo la transformación $Y = -X$. Se puede verificar sustituyendo el valor de $\kappa = 1$ en (1.3.8).
- 2).- La función de densidad log-gamma generalizada es sesgada a la izquierda $\forall \kappa > 0$, esto se puede apreciar en las gráficas 1.2, 1.3 y 1.4 y propiedades subsecuentes.

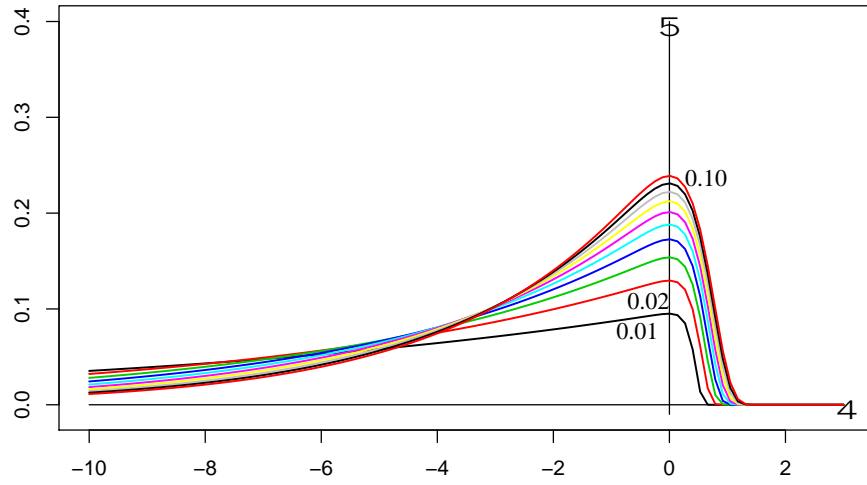


Figura 1.2: Función Log-gama generalizada estándar para $\kappa = 0.01, 0.02, \dots, 0.10$

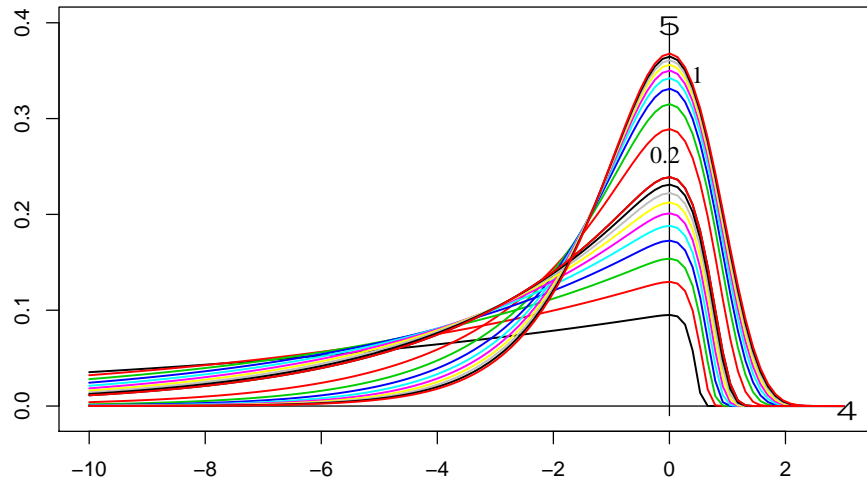


Figura 1.3: Función Log-gama generalizada estándar, anteriores y para $\kappa = 0.1, 0.2, \dots, 1$

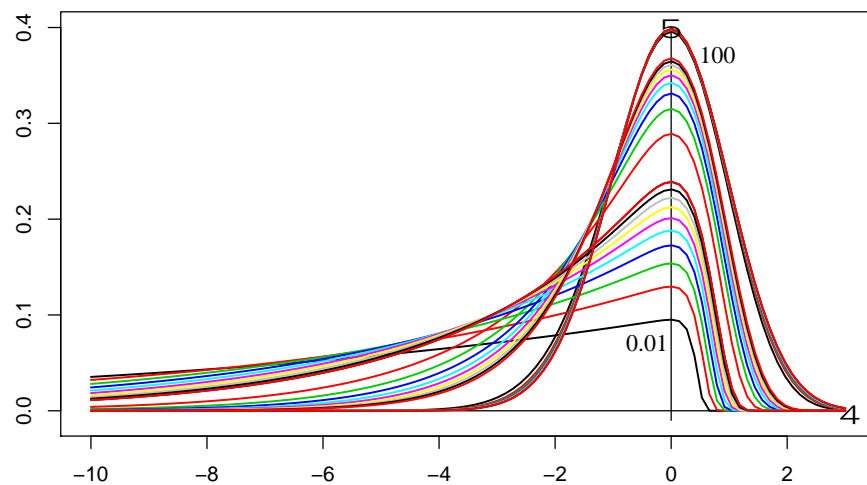


Figura 1.4: Función Log-gama generalizada estándar, anteriores y para $\kappa = 1, 10, 20, \dots, 100$

3).- Puntos máximo y de inflexión.

Proposición 1.1. *Sea Z una variable aleatoria con distribución $LGG(0, 1, \kappa)$, entonces en $z = 0$ se tiene su valor máximo para $f_Z(z)$, para todo $\kappa > 0$ y sus puntos de inflexión están dados por*

$$z_{1,2} = \sqrt{\kappa} \log \left(1 + \frac{1}{2\kappa} \pm \frac{1}{2\kappa} \sqrt{4\kappa + 1} \right).$$

Demostración. Derivando (1.3.8) con respecto a z e igualando a cero (nótese que se pueden quitar las constantes)

$$f'_Z(z) = \exp \left(z\sqrt{\kappa} - \kappa e^{z/\sqrt{\kappa}} \right) \left(\sqrt{\kappa} - \sqrt{\kappa} e^{z/\sqrt{\kappa}} \right) = 0.$$

De donde, $\sqrt{\kappa}(1 - e^{z/\sqrt{\kappa}}) = 0$, de aquí se obtiene el valor crítico $z = 0$. Para saber si se trata de un máximo se emplea el criterio de la segunda derivada.

$$f''_Z(z) = \exp \left(z\sqrt{\kappa} - \kappa e^{z/\sqrt{\kappa}} \right) \left(\kappa(1 - e^{z/\sqrt{\kappa}})^2 - e^{z/\sqrt{\kappa}} \right). \quad (1.4.1)$$

Sustituyendo $z = 0$, resulta que se trata de un máximo.

Para el punto de inflexión se utiliza la expresión (1.4.1), igualando a cero se obtiene $\kappa(1 - e^{z/\sqrt{\kappa}})^2 - e^{z/\sqrt{\kappa}} = 0$, realizando el cambio de variable $v = e^{z/\sqrt{\kappa}}$ y desarrollando el cuadrado resulta la ecuación $v^2 - 2v((2\kappa + 1)/(2\kappa)) + 1 = 0$, resolviendo la ecuación y regresando a la variable original resultan los puntos de inflexión

$$z_{1,2} = \sqrt{\kappa} \log \left(1 + \frac{1}{2\kappa} \pm \frac{1}{2\kappa} \sqrt{4\kappa + 1} \right). \quad (1.4.2)$$

El punto de inflexión depende del valor de κ . □

a).- Si $\kappa \rightarrow \infty$, entonces los puntos de inflexión se obtienen en ± 1 . En general en $\pm \sigma$.

b).- Si $\kappa \rightarrow 0$, entonces los puntos de inflexión se aproximan a cero.

c).- De la expresión (1.4.2) se deduce que los puntos de inflexión están en $(-1, 1)$.

4).- Cuando $\kappa \rightarrow \infty$, entonces $LGG(0, 1, \kappa)$ converge a la distribución normal estándar.

Proposición 1.2. *Sea Z una variable aleatoria con distribución $LGG(0, 1, \kappa)$, entonces $LGG(0, 1, \kappa) \xrightarrow{\kappa \rightarrow +\infty} N(0, 1)$.*

Demostración. Cuando κ crece puede utilizarse la propiedad 5 de la función gamma, obteniendo:

$$\frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \xrightarrow{\kappa \rightarrow +\infty} \frac{\kappa^{\kappa-1/2}}{\sqrt{2\pi}(\kappa)^{\kappa-1/2} e^{-\kappa}} = \frac{e^\kappa}{\sqrt{2\pi}}. \quad (1.4.3)$$

Sustituyendo el resultado asintótico (1.4.3) en la función de densidad (1.3.1), después de estandarizar con $z = \frac{x-\xi}{\sigma}$ se obtiene

$$f(z; 0, 1, \kappa) \xrightarrow{\kappa \rightarrow +\infty} \frac{1}{\sqrt{2\pi}} \exp \left(\kappa + z\sqrt{\kappa} - \kappa e^{\frac{z}{\sqrt{\kappa}}} \right), \quad -\infty < z < \infty.$$

Para z fija y κ creciendo, resulta $e^{\frac{z}{\sqrt{\kappa}}} \xrightarrow{\kappa \rightarrow +\infty} 1 + \frac{z}{\sqrt{\kappa}} + \left(\frac{z}{2\sqrt{\kappa}}\right)^2 + o\left(\frac{1}{\sqrt{\kappa}}\right)^3$, sustituyendo en la expresión anterior

$$\begin{aligned} f(z; 0, 1, \kappa) &\xrightarrow{\kappa \rightarrow +\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\kappa + z\sqrt{\kappa} - \kappa\left(1 + \frac{z}{\sqrt{\kappa}} + \frac{1}{2}\left(\frac{z}{\sqrt{\kappa}}\right)^2 + o\left(\frac{1}{\sqrt{\kappa}}\right)^3\right)\right) \\ &\xrightarrow{\kappa \rightarrow +\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \end{aligned}$$

La proposición queda demostrada. \square

5).- La función generatriz de momentos para $LGG(0, 1, \kappa)$ depende de κ .

Proposición 1.3. *Sea Z una variable aleatoria con distribución $LGG(0, 1, \kappa)$, entonces*

$$m_Z(t) = \frac{\Gamma(\kappa + t\sqrt{\kappa})}{\Gamma(\kappa)\kappa^{t\sqrt{\kappa}}}.$$

Demostración. Calculando la función generatriz de momentos

$$m_Z(t) = \frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \int_{-\infty}^{\infty} e^{tz} \exp\left(z\sqrt{\kappa} - \kappa e^{\left(\frac{z}{\sqrt{\kappa}}\right)}\right) dz.$$

Con el cambio de variable $w = \kappa e^{z/\sqrt{\kappa}} \Rightarrow z = \sqrt{\kappa} \log(w/\kappa) \Rightarrow dz = \sqrt{\kappa} \frac{dw}{w}$, luego

$$\begin{aligned} m_Z(t) &= \frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \int_0^{\infty} \exp\left(\sqrt{\kappa}\left(t + \sqrt{\kappa}\right) \log(w/\kappa)\right) e^{-w} \frac{\sqrt{\kappa} dw}{w} \\ &= \frac{1}{\Gamma(\kappa)\kappa^{t\sqrt{\kappa}}} \int_0^{\infty} w^{\sqrt{\kappa}(t+\sqrt{\kappa})-1} e^{-w} dw = \frac{\Gamma(\kappa + t\sqrt{\kappa})}{\Gamma(\kappa)\kappa^{t\sqrt{\kappa}}}. \end{aligned}$$

La proposición queda demostrada. \square

6).- El valor esperado de la distribución log-gamma generalizada depende del parámetro κ .

Proposición 1.4. *Sea Z una variable aleatoria con distribución $LGG(0, 1, \kappa)$, entonces*

$$E(Z) = \mu = \sqrt{\kappa} \left(-\gamma - \frac{1}{\kappa} - \log(\kappa) + \sum_{m=1}^{\infty} \frac{\kappa}{m(m + \kappa)} \right).$$

Demostración. Utilizando la función generatriz de momentos

$$\begin{aligned} E(Z) &= \left. \frac{dm_Z(t)}{dt} \right|_{t=0} = \left. \frac{d}{dt} \left(\frac{\Gamma(\kappa + t\sqrt{\kappa})}{\Gamma(\kappa)\kappa^{t\sqrt{\kappa}}} \right) \right|_{t=0} \\ &= \left. \frac{1}{\Gamma(\kappa)} \frac{\kappa^{t\sqrt{\kappa}} \frac{d}{dt} \Gamma(\kappa + t\sqrt{\kappa}) - \Gamma(\kappa + t\sqrt{\kappa}) \kappa^{t\sqrt{\kappa}} \sqrt{\kappa} \log(\kappa)}{(\kappa^{t\sqrt{\kappa}})^2} \right|_{t=0}. \end{aligned}$$

Sustituyendo en la expresión anterior $t = 0$, además del cambio de variable $v = \kappa + t\sqrt{\kappa}$

y $v = \kappa$ cuando $t = 0$ resulta, después de aplicar la regla de la cadena

$$E(Z) = \frac{1}{\Gamma(\kappa)} \left[\frac{d\Gamma(v)}{dv} \Big|_{v=\kappa} \sqrt{\kappa} - \Gamma(\kappa) \sqrt{\kappa} \log(\kappa) \right].$$

Utilizando $\Gamma(v) = \exp(\log \Gamma(v))$ y su derivada, se obtiene $\frac{d\Gamma(v)}{dv} = \Gamma(v)\psi(v)$ (ψ – función digamma), luego después de sustituir $v = \kappa$

$$E(Z) = \sqrt{\kappa}(\psi(\kappa) - \log(\kappa)). \quad (1.4.4)$$

Sustituyendo la propiedad 10a de la función gamma en (1.4.4) se concluye la proposición. \square

a).- En el caso particular de que $\kappa \in \mathbb{N}$, resulta $\mu = \sqrt{\kappa} \left(-\gamma - \log(\kappa) + \sum_{i=1}^{\kappa-1} \frac{1}{i} \right)$.

Demostración. Si $\kappa \in \mathbb{N}$, entonces

$$\sum_{m=1}^{\infty} \frac{\kappa}{m(m+\kappa)} = \sum_{m=1}^{\infty} \frac{1}{m} - \sum_{m=1}^{\infty} \frac{1}{m+\kappa} = \sum_{i=1}^{\kappa} \frac{1}{i}.$$

Los demás términos, después del κ -ésimo término se cancelan al realizar ambas sumas y los últimos κ -ésimos términos se eliminan por ser cantidades infinitesimales. \square

- i).- En particular, si $\kappa = 1$ se tiene $\mu = -\gamma$.
- ii).- En particular, si $\kappa = 2$ se tiene $\mu = \sqrt{2}(-\gamma - \log(2) + 1)$.
- iii).- En particular, si $\kappa = 3$ se tiene $\mu = \sqrt{3}(-\gamma - \log(3) + 3/2)$.

b).- Cuando $\kappa \rightarrow 0^+$, el valor esperado $\mu \sim \frac{1}{\sqrt{\kappa}}$.

Demostración. Cuando $\kappa \approx 0$, entonces $\sqrt{\kappa} \log(\kappa) \approx 0$, además de las propiedades 10a y 3b de la función gamma se obtiene $\psi(\kappa) = -\gamma - \frac{1}{\kappa} + \sum_{m=1}^{\infty} \frac{\kappa}{m(\kappa+m)}$ y cuando

$\kappa \approx 0$ resulta $\sum_{m=1}^{\infty} \frac{\kappa}{m(\kappa+m)} < \frac{\pi^2}{6}$. Luego, $\sqrt{\kappa} \left(-\gamma + \sum_{m=1}^{\infty} \frac{\kappa}{m(\kappa+m)} \right) \approx 0$, con esto $\sqrt{\kappa}\psi(\kappa) \approx \sqrt{\kappa}/\kappa = 1/\sqrt{\kappa}$, sustituyendo estos resultados en (1.4.4) se concluye $\mu \approx \frac{1}{\sqrt{\kappa}}$. \square

c).- Cuando $\kappa \rightarrow \infty$, el valor esperado $\mu \rightarrow 0$.

Demostración. Se concluye de la proposición 1.2. \square

7).- La n -ésima derivada de la función generatriz de momentos depende de κ .

Proposición 1.5. Sea Z una variable aleatoria con distribución $LGG(0, 1, \kappa)$, entonces

$$m_Z^{(n)}(t) = \frac{(\sqrt{\kappa})^n}{\Gamma(\kappa)\kappa^{t\sqrt{\kappa}}} \sum_{j=0}^n (-1)^j \binom{n}{j} \Gamma^{(n-j)}(v) (\log(\kappa))^j. \quad (1.4.5)$$

Evaluando en $t = 0$, resulta que la función generatriz de momentos es igual a:

$$m_Z^{(n)}(t)|_{t=0} = \frac{(\sqrt{\kappa})^n}{\Gamma(\kappa)} \sum_{j=0}^n (-1)^j \binom{n}{j} \Gamma^{(n-j)}(v)|_{v=\kappa} (\log(\kappa))^j. \quad (1.4.6)$$

En donde, $v = \kappa + t\sqrt{\kappa}$, $\Gamma^{(j)}(v) = \frac{d^j}{dv^j} \Gamma(v)$, $\Gamma^{(0)}(v) = \Gamma(v)$ y $(\sqrt{\kappa} \log(\kappa))^0 = 1$.

$$\Gamma^{(0)}(v)|_{v=\kappa} = \Gamma(\kappa) \quad (1.4.7)$$

$$\Gamma^{(1)}(v)|_{v=\kappa} = \Gamma(\kappa) \psi(\kappa) \quad (1.4.8)$$

$$\Gamma^{(2)}(v)|_{v=\kappa} = \Gamma(\kappa) (\psi^2(\kappa) + \psi'(\kappa)) \quad (1.4.9)$$

$$\Gamma^{(3)}(v)|_{v=\kappa} = \Gamma(\kappa) (\psi^3(\kappa) + 3\psi(\kappa)\psi'(\kappa) + \psi''(\kappa)) \quad (1.4.10)$$

$$\Gamma^{(4)}(v)|_{v=\kappa} = \Gamma(\kappa) (\psi^4(\kappa) + 6\psi^2(\kappa)\psi'(\kappa) + 4\psi(\kappa)\psi''(\kappa) + 3(\psi'(\kappa))^2 + \psi'''(\kappa)) \quad (1.4.11)$$

$$\Gamma^{(5)}(v)|_{v=\kappa} = \Gamma(\kappa) (\psi^5(\kappa) + 10\psi^3(\kappa)\psi'(\kappa) + 10\psi^2(\kappa)\psi''(\kappa) + 5\psi(\kappa)\psi'''(\kappa) + 15(\psi'(\kappa))^2\psi(\kappa) + 10\psi''(\kappa)\psi'(\kappa) + \psi^{(4)}(\kappa)) \quad (1.4.12)$$

$$\Gamma^{(6)}(v)|_{v=\kappa} = \Gamma(\kappa) (\psi^6(\kappa) + 15\psi^4(\kappa)\psi'(\kappa) + 20\psi^3(\kappa)\psi''(\kappa) + 15\psi^2(\kappa)\psi'''(\kappa) + 6\psi(\kappa)\psi^{(4)}(\kappa) + 45(\psi'(\kappa))^2\psi^2(\kappa) + 60\psi''(\kappa)\psi'(\kappa)\psi(\kappa) + 15(\psi'(\kappa))^3 + 15\psi'''(\kappa)\psi'(\kappa) + 10(\psi''(\kappa))^2 + \psi^{(5)}(\kappa)). \quad (1.4.13)$$

⋮

Demostración. Utilizando la función generatriz de momentos y la demostración de la proposición 1.4, en donde se tenía para $v = \kappa + t\sqrt{\kappa}$

$$\frac{dm_Z(t)}{dt} = \frac{\sqrt{\kappa}}{\Gamma(\kappa)} \left(\frac{\frac{d}{dv} \Gamma(v) - \Gamma(v) \log(\kappa)}{\kappa^{t\sqrt{\kappa}}} \right).$$

Derivando otra vez con respecto a t , se obtiene

$$\begin{aligned} \frac{d^2 m_Z(t)}{dt^2} &= \frac{\sqrt{\kappa}}{\Gamma(\kappa)} \left(\frac{\frac{d}{dt} \frac{d\Gamma(v)}{dv} - \frac{d\Gamma(v)}{dt} \log(\kappa) - \left(\frac{d}{dv} \Gamma(v) - \Gamma(v) \log(\kappa) \right) \sqrt{\kappa} \log(\kappa)}{\kappa^{t\sqrt{\kappa}}} \right) \\ &= \frac{(\sqrt{\kappa})^2}{\Gamma(\kappa)} \left(\frac{\frac{d^2}{dv^2} \Gamma(v) - 2 \log(\kappa) \frac{d}{dv} \Gamma(v) + (\log(\kappa))^2 \Gamma(v)}{\kappa^{t\sqrt{\kappa}}} \right). \end{aligned}$$

Nótese que al derivar en el numerador se está formando un triángulo de pascal, puesto que se están sumando los coeficientes de la expresión del numerador después de derivar y a partir del segundo término con ella misma sin derivar. Por ejemplo para la tercera

derivada con respecto a t , se obtiene

$$\begin{aligned} \frac{d^3 m_Z(t)}{dt^3} &= \frac{(\sqrt{\kappa})^2}{\Gamma(\kappa)} \left(\frac{\frac{d}{dt} \Gamma^{(2)}(v) - 2 \log(\kappa) \frac{d}{dt} \Gamma^{(1)}(v) + (\log(\kappa))^2 \frac{d}{dt} \Gamma(v)}{\kappa^{t\sqrt{\kappa}}} - \right. \\ &\quad \left. - \frac{\left(\Gamma^{(2)}(v) - 2 \log(\kappa) \Gamma^{(1)}(v) + (\log(\kappa))^2 \Gamma(v) \right) \sqrt{\kappa} \log(\kappa)}{\kappa^{t\sqrt{\kappa}}} \right) \\ &= \frac{(\sqrt{\kappa})^3}{\Gamma(\kappa)} \left(\frac{\Gamma^{(3)}(v) - 3 \log(\kappa) \Gamma^{(2)}(v) + 3 (\log(\kappa))^2 \Gamma^{(1)}(v) - (\log(\kappa))^3 \Gamma(v)}{\kappa^{t\sqrt{\kappa}}} \right). \end{aligned}$$

Derivando sucesivamente resulta la fórmula (1.4.5).

Para las derivadas de la función gamma, se utiliza el hecho que $\Gamma(v) = \exp(\log \Gamma(v))$ y las funciones polygamma. Por ejemplo

$$\begin{aligned} \Gamma^{(1)}(v) \Big|_{v=\kappa} &= \frac{d \exp(\log \Gamma(v))}{dv} \Big|_{v=\kappa} = \Gamma(v) \psi(v) \Big|_{v=\kappa} = \Gamma(\kappa) \psi(\kappa) \\ \Gamma^{(2)}(v) \Big|_{v=\kappa} &= \frac{d \Gamma(v) \psi(v)}{dv} \Big|_{v=\kappa} = \Gamma(v) (\psi^2(v) + \Gamma(v) \psi'(v)) \Big|_{v=\kappa} = \Gamma(\kappa) (\psi^2(\kappa) + \psi'(\kappa)) \\ \Gamma^{(3)}(v) \Big|_{v=\kappa} &= \frac{d \Gamma(v) (\psi^2(v) + \psi'(v))}{dv} \\ &= \Gamma(v) \psi(v) (\psi^2(v) + \psi'(v)) + \Gamma(v) (2\psi(v) \psi'(v) + \psi''(v)) \\ &= \Gamma(v) (\psi^3(v) + 3\psi(v) \psi'(v) + \psi''(v)) \Big|_{v=\kappa} \\ &= \Gamma(\kappa) (\psi^3(\kappa) + 3\psi(\kappa) \psi'(\kappa) + \psi''(\kappa)). \end{aligned}$$

En este desarrollo se puede apreciar que no se forma el triángulo de pascal, puesto que aparecen términos en varios factores, que al derivar rompen la regla del triángulo. \square

Note que la expresión para la derivada (1.4.6) se puede expresar como

$$m_Z^{(n)}(t) \Big|_{t=0} = \frac{(\sqrt{\kappa})^n}{\Gamma(\kappa)} \left\langle \Gamma^{(\cdot)}(v) \Big|_{v=\kappa} - \log(\kappa) \right\rangle^n \quad (1.4.14)$$

Entendiendo por $\Gamma^{(\cdot)}$ que la potencia del desarrollo del binomio, se considera como la derivada y no como la potencia.

8).- La varianza de la distribución log-gamma generalizada depende del parámetro κ .

Proposición 1.6. *Sea Z una variable aleatoria con distribución $LGG(0, 1, \kappa)$, entonces*

$$V(Z) = \sigma^2 = \kappa \psi'(\kappa).$$

Demostración. A partir de (1.4.6) para $n = 2$, resulta

$$\begin{aligned}
 m_Z^{(2)}(t)|_{t=0} &= \frac{(\sqrt{\kappa})^2}{\Gamma(\kappa)} \sum_{j=0}^2 (-1)^j \binom{2}{j} \Gamma^{(2-j)}(v)|_{v=\kappa} (\log(\kappa))^j \\
 &= \frac{\kappa}{\Gamma(\kappa)} \left(\Gamma^{(2)}(v) - 2\Gamma^{(1)}(v) \log(\kappa) + \Gamma(v) \log^2(\kappa) \right) \Big|_{v=\kappa} \\
 &= \frac{\kappa}{\Gamma(\kappa)} \left(\Gamma(\kappa)(\psi^2(\kappa) + \psi'(\kappa)) - 2(\Gamma(\kappa)\psi(\kappa)) \log(\kappa) + \Gamma(\kappa) \log^2(\kappa) \right) \\
 &= \kappa(\psi^2(\kappa) + \psi'(\kappa) - 2\psi(\kappa) \log(\kappa) + \log^2(\kappa)).
 \end{aligned}$$

Por otro lado,

$$\begin{aligned}
 \sigma^2 &= E(Z^2) - E^2(Z) = m_Z^{(2)}(t)|_{t=0} - \left(\sqrt{\kappa}(\psi(\kappa) - \log(\kappa)) \right)^2 \\
 &= \kappa(\psi^2(\kappa) + \psi'(\kappa) - 2\psi(\kappa) \log(\kappa) + \log^2(\kappa)) - \kappa(\psi^2(\kappa) - 2\psi(\kappa) \log(\kappa) + \log^2(\kappa)) \\
 &= \kappa\psi'(\kappa) = \sum_{m=0}^{\infty} \frac{\kappa}{(\kappa+m)^2}.
 \end{aligned}$$

La proposición queda demostrada. \square

a).- En particular cuando $\kappa \in \mathbb{N}$, la varianza $\sigma^2 = \kappa \left(\frac{\pi^2}{6} - \sum_{m=1}^{\kappa-1} \frac{1}{m^2} \right)$.

Demostración. Utilizando el resultado anterior y la propiedad 3b de la función gamma, se obtiene

$$\sigma^2 = \kappa \sum_{m=0}^{\infty} \frac{1}{(\kappa+m)^2} = \kappa \sum_{m=\kappa}^{\infty} \frac{1}{m^2} = \kappa \left(\frac{\pi^2}{6} - \sum_{m=1}^{\kappa-1} \frac{1}{m^2} \right).$$

\square

i).- Si $\kappa = 1$, entonces $\sigma^2 = \pi^2/6$.

ii).- Si $\kappa = 2$, entonces $\sigma^2 = \pi^2/3 - 2$

b).- Cuando $\kappa \rightarrow \infty$, entonces $\sigma^2 \rightarrow 1$. Se deduce de la propiedad 4 de las funciones log-gamma generalizada.

c).- Cuando $\kappa \rightarrow 0$, entonces $\sigma^2 \sim 1/\kappa$. Se tiene $\sum_{m=1}^{\infty} \frac{1}{(\kappa+m)^2} < \pi^2/6$, luego para κ pequeñas, $\kappa \sum_{m=1}^{\infty} \frac{1}{(\kappa+m)^2} \approx 0$, de donde

$$\sigma^2 = \kappa \sum_{m=0}^{\infty} \frac{1}{(\kappa+m)^2} = \frac{\kappa}{\kappa^2} + \kappa \sum_{m=1}^{\infty} \frac{1}{(\kappa+m)^2} \approx \frac{1}{\kappa}.$$

9).- El tercer momento central de la distribución log-gamma generalizada depende de κ .

Proposición 1.7. Sea Z una variable aleatoria con distribución $LGG(0, 1, \kappa)$, entonces

$$E((Z - \mu)^3) = (\sqrt{\kappa})^3 \psi''(\kappa).$$

Demostración. Siguiendo los mismos pasos que en el segundo momento central (varianza), y simplificando la notación por $\psi = \psi(\kappa)$ y también para sus derivadas, se tiene:

$$\begin{aligned} m_Z^{(3)}(t)|_{t=0} &= \frac{(\sqrt{\kappa})^3}{\Gamma(\kappa)} \sum_{j=0}^3 (-1)^j \binom{3}{j} \Gamma^{(3-j)}(v)|_{v=\kappa} (\log(\kappa))^j \\ &= \frac{(\sqrt{\kappa})^3}{\Gamma(\kappa)} \left(\Gamma^{(3)}(v) - 3\Gamma^{(2)}(v) \log(\kappa) + 3\Gamma^{(1)}(v) \log^2(\kappa) - \Gamma(v) \log^3(\kappa) \right) \Big|_{v=\kappa} \\ &= (\sqrt{\kappa})^3 \left(\psi^3 + 3\psi\psi' + \psi'' - 3(\psi^2 + \psi') \log(\kappa) + 3\psi \log^2(\kappa) - \log^3(\kappa) \right) \\ &= (\sqrt{\kappa})^3 \left(\psi^3 + 3\psi\psi' + \psi'' - 3\psi^2 \log(\kappa) - 3\psi' \log(\kappa) + 3\psi \log^2(\kappa) - \log^3(\kappa) \right). \end{aligned}$$

Por otro lado,

$$\begin{aligned} E[(Z - \mu)^3] &= E(Z^3) - 3\mu E(Z^2) + 3\mu^2 E(Z) - \mu^3 = m_Z^{(3)}(t)|_{t=0} - 3\mu m_Z^{(2)}(t)|_{t=0} + 2\mu^3 \\ &= (\sqrt{\kappa})^3 \left(\psi^3 + 3\psi\psi' + \psi'' - 3\psi^2 \log(\kappa) - 3\psi' \log(\kappa) + 3\psi \log^2(\kappa) - \log^3(\kappa) \right. \\ &\quad \left. - 3(\psi - \log(\kappa))(\psi^2 + \psi' - 2\psi \log(\kappa) + \log^2(\kappa)) + 2(\psi - \log(\kappa))^3 \right) \\ &= (\sqrt{\kappa})^3 \left(\psi^3 + 3\psi\psi' + \psi'' - 3\psi^2 \log(\kappa) - 3\psi' \log(\kappa) + 3\psi \log^2(\kappa) - \log^3(\kappa) \right. \\ &\quad \left. - 3\psi^3 - 3\psi\psi' + 9\psi^2 \log(\kappa) - 3\psi \log^2(\kappa) + 3\psi' \log(\kappa) - 6\psi \log^2(\kappa) + 3\log^3(\kappa) \right. \\ &\quad \left. + 2\psi^3 - 6\psi^2 \log(\kappa) + 6\psi \log^2(\kappa) - 2\log^3(\kappa) \right) \\ &= (\sqrt{\kappa})^3 \psi''(\kappa). \end{aligned}$$

La proposición queda demostrada. \square

10).- El cuarto momento central de la distribución log-gamma generalizada depende de κ .

Proposición 1.8. Sea Z una variable aleatoria con distribución $LGG(0, 1, \kappa)$, entonces

$$E((Z - \mu)^4) = \kappa^2 (\psi'''(\kappa) + 3(\psi'(\kappa))^2).$$

Demostración. Se realiza de la misma forma que los otros momentos centrales. \square

1.5. Propiedades de la familia gamma generalizada

La familia que está siendo revisada en el trabajo se refiere a la log-gamma generalizada, misma que está vinculada con la familia gamma generalizada

$$f(t; \alpha, \beta, \kappa) := \frac{\beta}{\alpha \Gamma(\kappa)} \kappa^{\kappa-1/2} \left(\frac{t}{\alpha} \right)^{\beta\sqrt{\kappa}-1} \exp \left(-\kappa \left(\frac{t}{\alpha} \right)^{\beta/\sqrt{\kappa}} \right), \quad t > 0, \alpha, \beta, \kappa > 0. \quad (1.5.1)$$

Propiedades de la distribución gamma generalizada:

- 1).- Cuando $\kappa = 1$ coincide con la distribución Weibull (log-Gumbel).
- 2).- Puntos máximos.

Proposición 1.9. Sea X una variable aleatoria $GG(\alpha, \beta, \kappa)$ entonces

- a).- Si $\beta\sqrt{\kappa} \leq 1$ la función es monótona decreciente y su valor máximo se obtiene cuando $x \rightarrow 0^+$.
- b).- Si $\beta\sqrt{\kappa} > 1$ su valor máximo se obtiene cuando $x = \alpha \left(\frac{\beta\sqrt{\kappa}-1}{\beta\sqrt{\kappa}} \right)^{\sqrt{\kappa}/\beta}$.

Demostración. Derivando la función de densidad (1.5.1) se obtiene

$$f'(x; \alpha, \beta, \kappa) = \frac{\beta\kappa^{\kappa-1/2}}{\alpha^2\Gamma(\kappa)} \left(\frac{x}{\alpha}\right)^{\beta\sqrt{\kappa}-2} \exp \left[-\kappa \left(\frac{x}{\alpha}\right)^{\beta/\sqrt{\kappa}} \right] \left[\beta\sqrt{\kappa} - 1 - \beta\sqrt{\kappa} \left(\frac{x}{\alpha}\right)^{\beta/\sqrt{\kappa}} \right].$$

De la expresión entre corchetes se observa que cuando $\beta\sqrt{\kappa} - 1 \leq 0$ la derivada es negativa, luego la función es monótona decreciente. Cuando $\beta\sqrt{\kappa} - 1 > 0$, se iguala a cero la derivada y resuelve la ecuación para $x > 0$ el único término que puede ser cero es el último corchete, de donde se obtiene el valor deseado. \square

NOTA 1.1. En el caso de que β sea un valor fijo y $\kappa \rightarrow \infty$ el máximo

$$\alpha \left(\frac{\beta\sqrt{\kappa}-1}{\beta\sqrt{\kappa}} \right)^{\sqrt{\kappa}/\beta} = \alpha \left(1 - \frac{1}{\beta\sqrt{\kappa}} \right)^{\sqrt{\kappa}/\beta} \xrightarrow{\kappa \rightarrow +\infty} \alpha e^{-1/\beta^2}$$

- 3).- Puntos de inflexión.

Proposición 1.10. Sea X una variable aleatoria $GG(\alpha, \beta, \kappa)$ entonces

- a).- Si $\beta\sqrt{\kappa} \leq 1$ la función es monótona decreciente y convexa.
- b).- Si $\beta\sqrt{\kappa} > 1$ su punto de inflexión se obtiene cuando

$$x = \begin{cases} \alpha \left[\frac{2\beta\kappa + \beta - 3\sqrt{\kappa} + \sqrt{\beta^2 + 4\beta^2\kappa - 6\beta\sqrt{\kappa} + \kappa}}{2\beta\kappa} \right]^{\sqrt{\kappa}/\beta}, & \beta\sqrt{\kappa} \in [1, 2]; \\ \alpha \left[\frac{2\beta\kappa + \beta - 3\sqrt{\kappa} \pm \sqrt{\beta^2 + 4\beta^2\kappa - 6\beta\sqrt{\kappa} + \kappa}}{2\beta\kappa} \right]^{\sqrt{\kappa}/\beta}, & \beta\sqrt{\kappa} \in [2, \infty). \end{cases}$$

Demostración. Derivando la función de densidad (1.5.1) dos veces se obtiene

$$f''(x; \alpha, \beta, \kappa) = \frac{\beta\kappa^{\kappa-1/2}}{\alpha^3\Gamma(\kappa)} \left(\frac{x}{\alpha}\right)^{\beta\sqrt{\kappa}-3} \exp \left[-\kappa \left(\frac{x}{\alpha}\right)^{\beta/\sqrt{\kappa}} \right] \left[(\beta\sqrt{\kappa} - 1)(\beta\sqrt{\kappa} - 2) - \right. \\ \left. (\beta\sqrt{\kappa} - 1)\beta\sqrt{\kappa} \left(\frac{x}{\alpha}\right)^{\beta/\sqrt{\kappa}} - \beta\sqrt{\kappa} \left(\beta\sqrt{\kappa} + \frac{\beta}{\sqrt{\kappa}} - 2 \right) \left(\frac{x}{\alpha}\right)^{\beta/\sqrt{\kappa}} + (\beta\sqrt{\kappa})^2 \left(\frac{x}{\alpha}\right)^{2\beta/\sqrt{\kappa}} \right].$$

Igualando a cero la segunda derivada el único término que puede ser cero es el último

corchete. Luego, denotando $m = \beta\sqrt{\kappa}$ y $z = \left(\frac{x}{\alpha}\right)^{\beta/\sqrt{\kappa}}$, resulta

$$\begin{aligned} (m-1)(m-2) - (m-1)mz - m(m+m/\kappa-2)z + m^2z^2 &= 0 \\ m^2z^2 - \left(2m^2 + \frac{m^2}{\kappa} - 3m\right)z + m^2 - 3m + 2 &= 0 \end{aligned}$$

De la última ecuación con la fórmula general y simplificando, resulta

$$z = \frac{2m + \frac{m}{\kappa} - 3 \pm \sqrt{\left(2m + \frac{m}{\kappa} - 3\right)^2 - 4(m^2 - 3m + 2)}}{2m} \quad (1.5.2)$$

Analizando la solución, se tiene que cuando $m^2 - 3m + 2 < 0$ la raíz es mayor a $2m + \frac{m}{\kappa} - 3$, luego no se puede restar el radical, porque $z < 0$, no puede ser. Entonces cuando $m^2 - 3m + 2 < 0$ únicamente se suma el radical, esto ocurre cuando $m \in [1, 2]$. Así, de (1.5.2) y simplificando la cantidad subradical se tiene

$$z = \frac{2m + \frac{m}{\kappa} - 3 + \sqrt{\frac{m^2}{\kappa^2} + \frac{4m^2 - 6m}{\kappa} + 1}}{2m}, \text{ para } m \in [1, 2] \quad (1.5.3)$$

Llevando a cabo un análisis similar, se obtiene que cuando $m \in (0, 1)$ el valor de $z < 0$, lo que no puede ser. Para el caso faltante, se tiene que la cantidad subradical $\frac{m^2}{\kappa^2} + \frac{4m^2 - 6m}{\kappa} + 1 > 0$, pero $2m + \frac{m}{\kappa} - 3 > \sqrt{\left(2m + \frac{m}{\kappa} - 3\right)^2 - 4(m^2 - 3m + 2)}$ cuando $m \in [2, \infty)$. Por lo tanto,

$$z = \frac{2m + \frac{m}{\kappa} - 3 \pm \sqrt{\frac{m^2}{\kappa^2} + \frac{4m^2 - 6m}{\kappa} + 1}}{2m}, \text{ para } m \in [2, \infty) \quad (1.5.4)$$

Regresando a la variable original x y los parámetros α , β y κ con la soluciones (1.5.3) y (1.5.4), se tienen los puntos de inflexión

$$x = \begin{cases} \alpha \left[\frac{2\beta\kappa + \beta - 3\sqrt{\kappa} + \sqrt{\beta^2 + 4\beta^2\kappa - 6\beta\sqrt{\kappa} + \kappa}}{2\beta\kappa} \right]^{\sqrt{\kappa}/\beta}, & \beta\sqrt{\kappa} \in [1, 2]; \\ \alpha \left[\frac{2\beta\kappa + \beta - 3\sqrt{\kappa} \pm \sqrt{\beta^2 + 4\beta^2\kappa - 6\beta\sqrt{\kappa} + \kappa}}{2\beta\kappa} \right]^{\sqrt{\kappa}/\beta}, & \beta\sqrt{\kappa} \in [2, \infty). \end{cases}$$

Se concluye la proposición. \square

- 4).- La función de densidad gamma generalizada es sesgada a la derecha para $\kappa > 0$, esto se puede apreciar en las gráficas 1.5 y propiedades subsecuentes.
- 5).- Comportamiento asintótico.

Proposición 1.11. Cuando $\kappa \rightarrow \infty$, entonces $GG(\alpha, \beta, \kappa)$ converge a la distribución log-normal con parámetros $\mu = \log(\alpha)$ y $\sigma = \beta^{-1}$.

Demostración. Cuando κ crece puede utilizarse la propiedad 5 de la función gamma,

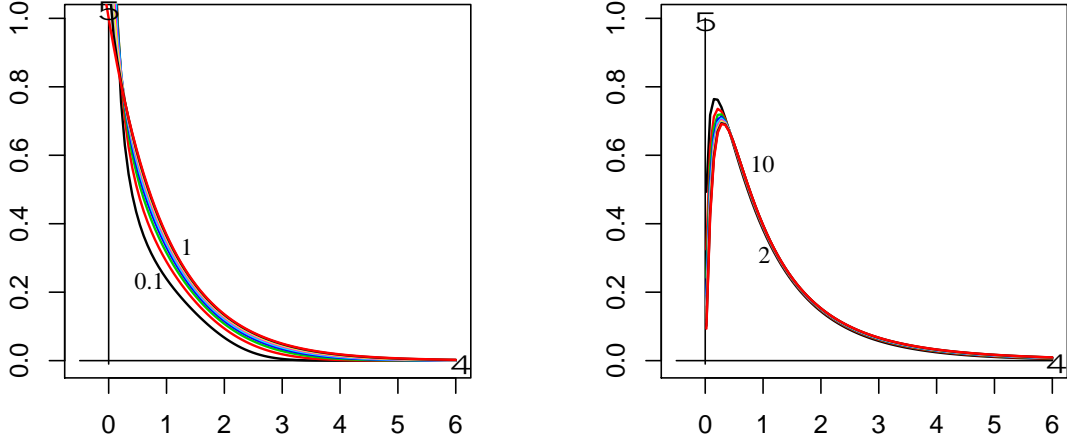


Figura 1.5: Función Gama generalizada estándar para $\kappa = 0.1, 0.2, \dots, 1$ y $\kappa = 2, 3, \dots, 10$

obteniendo:

$$\frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \xrightarrow{\kappa \rightarrow +\infty} \frac{\kappa^{\kappa-1/2}}{\sqrt{2\pi}(\kappa)^{\kappa-1/2}e^{-\kappa}} = \frac{e^{\kappa}}{\sqrt{2\pi}}. \quad (1.5.5)$$

Sustituyendo el resultado asintótico (1.5.5) en la función de densidad (1.5.1) se obtiene

$$f(t; \alpha, \beta, \kappa) \xrightarrow{\kappa \rightarrow +\infty} \frac{\beta e^{\kappa}}{\alpha \sqrt{2\pi}} \left(\frac{t}{\alpha}\right)^{\beta\sqrt{\kappa}-1} \exp\left(-\kappa \left(\frac{t}{\alpha}\right)^{\beta/\sqrt{\kappa}}\right).$$

Para z fija y κ creciendo, resulta

$$\begin{aligned} \left(\frac{t}{\alpha}\right)^{\beta/\sqrt{\kappa}} &= \exp\left(\frac{\beta}{\sqrt{\kappa}} \log\left(\frac{t}{\alpha}\right)\right) \\ &\xrightarrow{\kappa \rightarrow +\infty} 1 + \frac{\beta}{\sqrt{\kappa}} \log\left(\frac{t}{\alpha}\right) + \frac{1}{2} \left(\frac{\beta}{\sqrt{\kappa}} \log\left(\frac{t}{\alpha}\right)\right)^2 + o\left(\frac{\beta}{\sqrt{\kappa}} \log\left(\frac{t}{\alpha}\right)\right)^3, \end{aligned}$$

sustituyendo en la expresión anterior y $\left(\frac{t}{\alpha}\right)^{\beta\sqrt{\kappa}-1} = \exp\left(\beta\sqrt{\kappa} \log\left(\frac{t}{\alpha}\right)\right) \left(\frac{t}{\alpha}\right)^{-1}$

$$f(t; \alpha, \beta, \kappa) \xrightarrow{\kappa \rightarrow +\infty} \frac{\beta}{t\sqrt{2\pi}} \exp\left[\kappa + \beta\sqrt{\kappa} \log\left(\frac{t}{\alpha}\right) - \kappa \left(1 + \frac{\beta}{\sqrt{\kappa}} \log\left(\frac{t}{\alpha}\right) + \frac{1}{2} \left(\frac{\beta}{\sqrt{\kappa}} \log\left(\frac{t}{\alpha}\right)\right)^2 + o\left(\frac{1}{\sqrt{\kappa}}\right)^3\right)\right].$$

Simplificando los paréntesis $f(t; \alpha, \beta, \kappa) \xrightarrow{\kappa \rightarrow +\infty} \frac{\beta}{t\sqrt{2\pi}} \exp\left(-\frac{\beta^2}{2} (\log(t) - \log(\alpha))^2\right)$. \square

6).- El r -ésimo momento depende de κ .

Proposición 1.12. Sea T una variable aleatoria con distribución $GG(\alpha, \beta, \kappa)$, entonces

su r -ésimo momento está dado por:

$$E(T^r) = \left(\frac{\alpha}{\kappa\sqrt{\kappa}/\beta} \right)^r \frac{\Gamma(\kappa + r\sqrt{\kappa}/\beta)}{\Gamma(\kappa)}. \quad (1.5.6)$$

Demostración. Por definición del r -ésimo momento

$$E(T^r) = \int_{-\infty}^{\infty} t^r f(t; \alpha, \beta, \kappa) dt = \frac{\beta\kappa^{\kappa-1/2}}{\alpha\Gamma(\kappa)} \int_0^{\infty} t^r \left(\frac{t}{\alpha} \right)^{\beta\sqrt{\kappa}-1} \exp\left(-\kappa \left(\frac{t}{\alpha} \right)^{\beta/\sqrt{\kappa}}\right) dt.$$

Realizando el cambio de variable $x = \kappa \left(\frac{t}{\alpha} \right)^{\beta/\sqrt{\kappa}} \Rightarrow \frac{t}{\alpha} = \left(\frac{x}{\kappa} \right)^{\sqrt{\kappa}/\beta} \Rightarrow dt = \frac{\alpha}{\beta\sqrt{\kappa}} \left(\frac{x}{\kappa} \right)^{\sqrt{\kappa}/\beta-1} dx$

$$\begin{aligned} E(T^r) &= \frac{\beta\kappa^{\kappa-1/2}}{\alpha\Gamma(\kappa)} \int_0^{\infty} \left(\alpha \left(\frac{x}{\kappa} \right)^{\sqrt{\kappa}/\beta} \right)^r \left(\left(\frac{x}{\kappa} \right)^{\sqrt{\kappa}/\beta} \right)^{\beta\sqrt{\kappa}-1} e^{-x} \left(\frac{\alpha}{\beta\sqrt{\kappa}} \left(\frac{x}{\kappa} \right)^{\sqrt{\kappa}/\beta-1} \right) dx \\ &= \frac{\alpha^r \kappa^{\kappa-1}}{\Gamma(\kappa)} \int_0^{\infty} \left(\frac{x}{\kappa} \right)^{\kappa+r\sqrt{\kappa}/\beta-1} e^{-x} dx \\ &= \frac{\alpha^r \kappa^{\kappa-1-\kappa-r\sqrt{\kappa}/\beta+1}}{\Gamma(\kappa)} \int_0^{\infty} x^{\kappa+r\sqrt{\kappa}/\beta-1} e^{-x} dx \\ &= \left(\frac{\alpha}{\kappa\sqrt{\kappa}/\beta} \right)^r \frac{\Gamma(\kappa + r\sqrt{\kappa}/\beta)}{\Gamma(\kappa)}. \end{aligned}$$

Con esto se verifica la proposición. \square

a).- El valor esperado de la distribución gamma generalizada depende del parámetro κ . Sustituyendo $r = 1$ en (1.5.6)

$$E(T) = \frac{\alpha}{\kappa\sqrt{\kappa}/\beta} \frac{\Gamma(\kappa + \sqrt{\kappa}/\beta)}{\Gamma(\kappa)}.$$

b).- La varianza de la distribución gamma generalizada depende del parámetro κ . Sustituyendo $r = 1$ en (1.5.6)

$$E(T^2) = \left(\frac{\alpha}{\kappa\sqrt{\kappa}/\beta} \right)^2 \frac{\Gamma(\kappa + 2\sqrt{\kappa}/\beta)}{\Gamma(\kappa)}.$$

La varianza se obtiene de

$$V(T) = E(T^2) - (E(T))^2 = \left(\frac{\alpha}{\kappa\sqrt{\kappa}/\beta} \right)^2 \left[\frac{\Gamma(\kappa + 2\sqrt{\kappa}/\beta)}{\Gamma(\kappa)} - \left(\frac{\Gamma(\kappa + \sqrt{\kappa}/\beta)}{\Gamma(\kappa)} \right)^2 \right].$$

1.6. Familias de localidad-escala

En el estudio de las distribuciones puede observarse que existen variables aleatorias X , cuya distribución depende de los parámetros de localidad y escala, denotados por μ y σ , respectivamente, pero la variable resultante de la transformación $Y = \frac{X-\mu}{\sigma}$, no depende de los parámetros. Por ejemplo la distribución normal $X \sim N(\mu, \sigma^2)$, mientras que $Y = \frac{X-\mu}{\sigma} \sim N(0, 1)$. A este tipo de distribuciones se les llama *de localidad y escala*, o en forma simple **loc-escala**. En estas condiciones se dice que la variable es invariante con respecto a localidad y escala.

Sea Y una variable aleatoria en $(-\infty, \infty)$ que proviene de un modelo loc-escala, entonces su función de densidad de probabilidades está dada por

$$f(y) = \frac{1}{b} f_0\left(\frac{y-u}{b}\right), \quad -\infty < y < \infty. \quad (1.6.1)$$

Donde u - parámetro de localidad y $b > 0$ el parámetro de escala, además $f_0(z)$ con $z = \frac{y-u}{b}$ una función de densidad específica en $(-\infty, \infty)$.

1.7. Coeficiente de asimetría para la log-gamma generalizada

En las secciones previas se revisaron las propiedades de la familia log-gamma generalizada y se mostró que se trata de una familia loc-escala de distribuciones, por tal manera es de mayor interés trabajar buscando los estimadores para el parámetro de forma, κ . La estimación del parámetro de forma para la distribución log-gamma generalizada es crucial para la realización de inferencia de estas distribuciones, por tales motivos se buscan diferentes estimadores.

El parámetro de forma es posible que sea estimado con el coeficiente de asimetría, por tales razones es lógico buscar la forma de calcular este coeficiente.

Teorema 1.13. *Sea Z una variable aleatoria $LGG(0, 1, \kappa)$, entonces*

$$\gamma(\kappa) = \frac{E((Z - \mu)^3)}{[E((Z - \mu)^2)]^{3/2}} = \frac{\psi''(\kappa)}{(\psi'(\kappa))^{3/2}} = -\frac{2 \sum_{m=0}^{\infty} \frac{1}{(\kappa+m)^3}}{\left(\sum_{m=0}^{\infty} \frac{1}{(\kappa+m)^2}\right)^{3/2}}. \quad (1.7.1)$$

En particular:

1).- Si $\kappa \rightarrow 0^+$, entonces $\gamma \rightarrow -2^+$.

2).- Si $\kappa \rightarrow \infty$, entonces $\gamma \rightarrow 0^-$.

3).- Si $\kappa \in \mathbb{N}$, entonces $\gamma(\kappa) = -\frac{2 \left(1.202056903\dots - \sum_{m=1}^{\kappa-1} \frac{1}{m^3}\right)}{\left(\frac{\pi^2}{6} - \sum_{m=1}^{\kappa-1} \frac{1}{m^2}\right)^{3/2}}$.

Demostración. La demostración de (1.7.1) se concluye de las propiedades 8 y 9 de una distribución log-gamma generalizada y de las propiedades 10b y 10c de la función gamma.

Para los casos particulares:

1).- Cuando $\kappa \rightarrow 0^+$, se pueden descomponer las series del coeficiente de asimetría

$$\sum_{m=0}^{\infty} \frac{1}{(\kappa + m)^3} = \frac{1}{\kappa^3} + \sum_{m=1}^{\infty} \frac{1}{(\kappa + m)^3} \approx \frac{1}{\kappa^3} + 1.202057$$

$$\sum_{m=0}^{\infty} \frac{1}{(\kappa + m)^2} = \frac{1}{\kappa^2} + \sum_{m=1}^{\infty} \frac{1}{(\kappa + m)^2} \approx \frac{1}{\kappa^2} + \frac{\pi^2}{6}.$$

Sustituyendo en (1.7.1), entonces $\gamma \xrightarrow[k \rightarrow +0]{} -2^+$.

2).- Se concluye de la proposición 1.2.

3).- Si $\kappa \in \mathbb{N}$, entonces

$$\sum_{m=0}^{\infty} \frac{1}{(\kappa + m)^3} = \sum_{m=1}^{\infty} \frac{1}{m^3} - \sum_{m=1}^{\kappa-1} \frac{1}{m^3}$$

$$\sum_{m=0}^{\infty} \frac{1}{(\kappa + m)^2} = \sum_{m=1}^{\infty} \frac{1}{m^2} - \sum_{m=1}^{\kappa-1} \frac{1}{m^2}.$$

Con estas igualdades, la propiedad 3b de una función gamma y $\sum_{m=1}^{\infty} \frac{1}{m^3} \doteq 1.202056903\dots$, sustituyendo en (1.7.1). El teorema queda demostrado. \square

Finalmente de la demostración, nótese que el coeficiente de asimetría siempre es negativo como se muestra en la figura 1.6.

Para el estudio de las propiedades del estimador κ por medio del coeficiente de asimetría en el capítulo 2 se demostrarán algunas propiedades y en el capítulo 3 se hará por medio del proyecto R y una simulación con los momentos muestrales centrales.

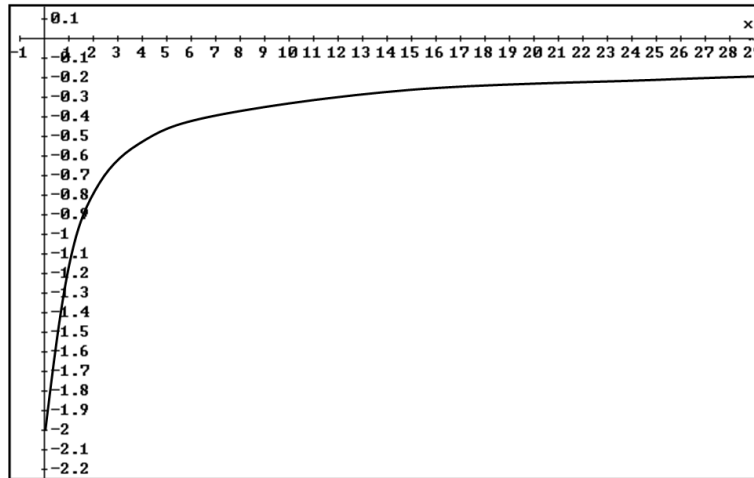


Figura 1.6: Función de asimetría $\gamma(x)$

Nótese que a partir del valor de κ y del hecho que la función del coeficiente de asimetría es estrictamente monótona creciente tiene inversa. por lo tanto se puede aproximar una función para determinar el valor de κ a partir del coeficiente de asimetría.

Estimadores puntuales para el parámetro κ de la distribución $LGG(0, 1, \kappa)$

2.1. Introducción

En este capítulo serán revisados los estimadores para la distribución $LGG(0, 1, \kappa)$. El problema inicia con un planteamiento general para la distribución $LGG(\xi, \sigma, \kappa)$ y la problemática para estimar los tres parámetros. Posteriormente con base en la invarianza de la distribución con respecto a los estimadores de localidad y escala, y que la estadística que será utilizada en los capítulos subsecuentes, coeficiente de correlación, también es invariante con respecto a dichos parámetros, el problema se reduce en buscar estimadores para el parámetro de forma κ .

Durante el capítulo se mostrarán tres estimadores para el parámetro κ , el primero de ellos por el método clásico de máxima verosimilitud y los otros dos estimadores por el método de momentos. En cada uno de ellos se demuestra su existencia.

En el caso del estimador de máxima verosimilitud, éste es único debido a la propiedad de invarianza de los EMV con respecto a transformaciones.

Con respecto a los estimadores de momentos, éstos se obtienen de forma indirecta, uno de ellos es resultado de una transformación que regresa a la distribución gama, en el otro caso se utiliza el coeficiente de asimetría, función monótona creciente con respecto al parámetro κ , posteriormente por medio de su inversa o una aproximación de ésta se puede estimar al parámetro de forma κ .

También en una sección por separado se estudia el comportamiento asintótico de cada uno de los dos primeros estimadores encontrados. En esta sección se demuestra que los estimadores de máxima verosimilitud y el estimador de momentos, basado en el coeficiente de asimetría, tienen una distribución asintótica normal.

En el caso del estimador de máxima verosimilitud la demostración se basa en la comprobación del cumplimiento de las restricciones del teorema respectivo, para el caso particular

de la familia log-gamma generalizada. El resultado queda expresado de dos forma, bajo una serie de los momentos no centrales y la otra bajo los momentos centrales. Para esto último se demuestra un resultado para calcular los momentos centrales que ahorra mucho tiempo de cálculo.

Finalmente, para el estimador de momentos se demuestra que tiene una distribución asintótica normal, para esto se obtiene un resultado general para la distribución del cociente de dos momentos centrales de la suma de variables aleatorias. Con el resultado encontrado se verifica el caso particular para la distribución asintótica del estimador de momentos para una distribución log-gamma generalizada. En este punto la obtención de los momentos centrales para la familia log-gamma generalizada en su forma estándar es muy importante, porque al igual que en el estimador de máxima verosimilitud la varianza de la distribución asintótica está dada en función los momentos centrales, mismos que no se conocen y su cálculo es extremadamente laborioso, pero con el resultado demostrado dicho cálculo se reduce considerablemente.

2.2. Búsqueda de estimadores para la distribución $LGG(0, 1, \kappa)$

Un paso importante para obtener buenas pruebas de bondad de ajuste, radica en la elección de los estimadores de los parámetros. Por tales razones en esta sección se revisarán algunos estimadores para el parámetro de la distribución $LGG(0, 1, \kappa)$.

Primeramente se revisará el problema general $LGG(\xi, \sigma, \kappa)$, luego con base en el capítulo anterior en donde se revisaron las propiedades de la familia log-gamma generalizada, resultó que se trata de una familia loc-escala de distribuciones. Por tales razones será de mayor interés trabajar buscando los estimadores únicamente para el parámetro de forma, κ , cuando la distribución está dada en su forma estándar.

2.2.1. EMV para los parámetros de la distribución $LGG(\xi, \sigma, \kappa)$

En caso de buscar los EMV cuando se trata de la distribución log-gamma generalizada es necesario resolver un sistema de ecuaciones no lineal que se obtiene de la función de verosimilitud, en el caso de la distribución log-gamma generalizada dicha función está dada por:

$$L(\kappa, \xi, \sigma; \mathbf{x}) = \left(\frac{\kappa^{\kappa-1/2}}{\sigma \Gamma(\kappa)} \right)^n \exp \left(\sqrt{\kappa} \sum_{i=1}^n \frac{x_i - \xi}{\sigma} - \kappa \sum_{i=1}^n e^{(x_i - \xi)/(\sigma \sqrt{\kappa})} \right).$$

Como se sabe generalmente se trabaja con la función log-verosimilitud, de tal forma que calculando el logaritmo natural de la expresión anterior se tiene:

$$\ell(\kappa, \xi, \sigma; \mathbf{x}) = n \left((\kappa - 1/2) \log(\kappa) - \log(\sigma) - \log \Gamma(\kappa) \right) + \sqrt{\kappa} \sum_{i=1}^n \frac{x_i - \xi}{\sigma} - \kappa \sum_{i=1}^n e^{(x_i - \xi)/(\sigma \sqrt{\kappa})}. \quad (2.2.1)$$

Para obtener su máximo, se deriva parcialmente con respecto a cada uno de los parámetros y se iguala a cero cada ecuación obteniendo un sistema no lineal con tres ecuaciones.

Derivando con respecto a ξ

$$\begin{aligned} \frac{\partial \ell(\kappa, \xi, \sigma; \mathbf{x})}{\partial \xi} &= -\frac{\sqrt{\kappa}n}{\sigma} - \kappa \sum_{i=1}^n \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) \left(-\frac{1}{\sigma\sqrt{\kappa}}\right) = 0 \\ &= -n + \sum_{i=1}^n \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) = 0 \end{aligned} \quad (2.2.2)$$

Derivando con respecto a σ y sustituyendo $n = \sum_{i=1}^n \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right)$ de (2.2.2), se tiene

$$\begin{aligned} \frac{\partial \ell(\kappa, \xi, \sigma; \mathbf{x})}{\partial \sigma} &= -\frac{n}{\sigma} - \sqrt{\kappa} \sum_{i=1}^n \frac{x_i - \xi}{\sigma^2} - \kappa \sum_{i=1}^n \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) \left(-\frac{x_i - \xi}{\sigma^2\sqrt{\kappa}}\right) = 0 \\ &= -n\sigma - n\sqrt{\kappa}(\bar{x} - \xi) + \sqrt{\kappa} \sum_{i=1}^n (x_i - \xi) \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) = 0 \\ &= -n\sigma - n\sqrt{\kappa}\bar{x} + \sqrt{\kappa} \sum_{i=1}^n x_i \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) = 0 \end{aligned} \quad (2.2.3)$$

Derivando con respecto a κ y posteriormente sustituyendo $n = \sum_{i=1}^n \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right)$ de (2.2.2) y $\sum_{i=1}^n (x_i - \xi) \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) = \frac{n\sigma + n\sqrt{\kappa}(\bar{x} - \xi)}{\sqrt{\kappa}}$ de (2.2.3), se tiene

$$\begin{aligned} \frac{\partial \ell(\kappa, \xi, \sigma; \mathbf{x})}{\partial \kappa} &= n \left(\log(\kappa) + \frac{\kappa - 1/2}{\kappa} - \psi(\kappa) \right) + \frac{1}{2\sqrt{\kappa}} \sum_{i=1}^n \left(\frac{x_i - \xi}{\sigma} \right) - \\ &\quad \sum_{i=1}^n \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) - \kappa \sum_{i=1}^n \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) \left(\frac{x_i - \xi}{\sigma} \right) \left(-\frac{1}{2\kappa^{3/2}} \right) = 0 \\ &= n \left(\log(\kappa) - \frac{1}{2\kappa} - \psi(\kappa) \right) + \frac{n}{2\sigma\sqrt{\kappa}} (\bar{x} - \xi) + \frac{n\sigma + n\sqrt{\kappa}(\bar{x} - \xi)}{2\kappa\sigma} = 0 \\ &= \log(\kappa) - \psi(\kappa) + \frac{\bar{x} - \xi}{\sigma\sqrt{\kappa}} = 0 \end{aligned} \quad (2.2.4)$$

De las ecuaciones (2.2.2), (2.2.3) y (2.2.4) se forma el sistema de ecuaciones (2.2.5)

$$\begin{aligned} \log(\kappa) - \psi(\kappa) + \frac{\bar{x} - \xi}{\sigma\sqrt{\kappa}} &= 0 \\ \sum_{i=1}^n \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) - n &= 0 \\ \sqrt{\kappa} \sum_{i=1}^n x_i \exp\left(\frac{x_i - \xi}{\sigma\sqrt{\kappa}}\right) - n\sigma - n\sqrt{\kappa}\bar{x} &= 0 \end{aligned} \quad (2.2.5)$$

El sistema de ecuaciones (2.2.5) no se puede resolver analíticamente, pero puede simplificarse su solución considerando el valor del parámetro κ conocido, denotándolo por κ_0 y

sustituyendo en este sistema y despejando ξ de la ecuación 1 del sistema (2.2.5)

$$\xi = \bar{x} + \sigma\sqrt{\kappa_0}(\log(\kappa_0) - \psi(\kappa_0)). \quad (2.2.6)$$

Sustituyendo el valor (2.2.6) en cualquiera de la otras dos ecuaciones del sistema (2.2.5), por ejemplo en la segunda ecuación, se tiene la ecuación (2.2.7) con respecto a σ

$$\begin{aligned} \sum_{i=1}^n \exp\left(\frac{x_i - \bar{x} - \sigma\sqrt{\kappa_0}(\log(\kappa_0) - \psi(\kappa_0))}{\sigma\sqrt{\kappa_0}}\right) - n &= 0 \\ \sum_{i=1}^n \exp\left(\frac{x_i - \bar{x}}{\sigma\kappa_0}\right) - n\kappa_0 e^{-\psi(\kappa_0)} &= 0. \end{aligned} \quad (2.2.7)$$

2.2.2. EMV para el parámetro de forma de $LGG(0, 1, \kappa)$

La distribución log-gamma generalizada es de tipo loc-escala, por esta razón se analiza la distribución en su forma estándar, ver (1.3.4), en donde se busca el EMV del parámetro de forma κ . La función de verosimilitud que se debe maximizar está dada por:

$$\begin{aligned} L(\kappa; \mathbf{z}) &= \prod_{i=1}^n \frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \exp\left(z_i\sqrt{\kappa} - \kappa e^{z_i/\sqrt{\kappa}}\right) \\ &= \left(\frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)}\right)^n \exp\left(\sqrt{\kappa} \sum_{i=1}^n z_i - \kappa \sum_{i=1}^n e^{z_i/\sqrt{\kappa}}\right). \end{aligned} \quad (2.2.8)$$

Teorema 2.1. *Sea Z una variable aleatoria $LGG(0, 1, \kappa)$, entonces para toda $\kappa > 0$ y una realización z_1, \dots, z_n de una muestra aleatoria de tamaño n con $n > 20$ y que se conserven en forma aproximada los porcentajes de la tabla 2.1 el EMV de κ siempre existe y es único.*

NOTA 2.1. Los porcentajes de valores negativos de la distribución $LGG(0, 1, \kappa)$ están dados en la siguiente tabla para algunos valores de κ . Los valores de la tabla se obtienen calculando las probabilidades de que la variable aleatoria $X \in LGG(0, 1, \kappa)$ sea negativa.

κ	0.1	0.5	0.8	1	2	3	4	5	$\rightarrow +\infty$
% de negativos	82.8 %	68.3 %	64.7 %	63.2 %	59.4 %	57.7 %	56.7 %	56 %	$\rightarrow 50$ %

Cuadro 2.1: Porcentajes de valores negativos según el valor de κ

Demostración. Se llevará a cabo en tres partes, que se aplicarán a la función log-verosimilitud

$$\ell(\kappa; \mathbf{z}) = n \left(\left(\kappa - \frac{1}{2} \right) \log(\kappa) + \bar{z}\sqrt{\kappa} - \log(\Gamma(\kappa)) \right) - \kappa \sum_{i=1}^n e^{z_i/\sqrt{\kappa}}. \quad (2.2.9)$$

Parte 1. Se mostrará que la derivada de la log-verosimilitud cambia de signo para $\kappa > 0$.

Derivando la función (2.2.9)

$$\ell'(\kappa; \mathbf{z}) = n \left(\log(\kappa) + 1 - \frac{1}{2\kappa} + \frac{\bar{z}}{2\sqrt{\kappa}} - \psi(\kappa) - \frac{1}{n} \sum_{i=1}^n e^{z_i/\sqrt{\kappa}} + \frac{1}{2\sqrt{\kappa}} \sum_{i=1}^n \frac{z_i}{n} e^{z_i/\sqrt{\kappa}} \right). \quad (2.2.10)$$

Analizando los términos fuera de las sumatorias para valores pequeños de κ . En este caso $\psi(\kappa) \approx -\frac{1}{\kappa}$, de donde $-\frac{1}{2\kappa} - \psi(\kappa) \approx \frac{1}{2\kappa}$. Luego se tienen los términos positivos $1 + \frac{1}{2\kappa}$ y negativos a $\log(\kappa) + \frac{\bar{z}}{2\sqrt{\kappa}}$ (como κ es pequeño el mayor porcentaje de las $z_i < 0$, $n > 20$), por consiguiente $\bar{z} < 0$. Por orden de crecimiento para κ pequeñas es más fuerte $\frac{1}{2\kappa}$ que la logarítmica y $\frac{1}{\sqrt{\kappa}}$, de donde la suma de términos fuera de las sumatorias es positivo de orden $\frac{1}{2\kappa}$. Para las sumatorias resulta lo siguiente, por ser κ pequeña existe un gran porcentaje de valores z_i negativos, luego es posible dividir las sumatorias en dos, una de términos $z_i < 0$ y los restantes. Denotando a tales sumatorias por Σ^- y Σ^+ , respectivamente, resulta

$$-\frac{1}{n} \sum_{i=1}^n e^{z_i/\sqrt{\kappa}} + \frac{1}{2\sqrt{\kappa}} \sum_{i=1}^n \frac{z_i}{n} e^{z_i/\sqrt{\kappa}} = -(\Sigma_1^- + \Sigma_1^+) + (\Sigma_2^- + \Sigma_2^+) = (-\Sigma_1^- + \Sigma_2^-) + (-\Sigma_1^+ + \Sigma_2^+).$$

Note que los términos en $e^{z_i/\sqrt{\kappa}}$ y $z_i e^{z_i/\sqrt{\kappa}}$ para $z_i < 0$ son pequeños, luego $(-\Sigma_1^- + \Sigma_2^-)$ es despreciable. Cuando los valores z_i no son negativos, resulta que $\frac{z_i}{\sqrt{\kappa}} > 1$, para κ pequeñas y por lo tanto $\Sigma_2^+ > \Sigma_1^+$, de donde $(-\Sigma_1^+ + \Sigma_2^+) > 0$. Luego, para valores pequeños de κ la derivada de la función log-verosimilitud es positiva grande de orden $O(\frac{1}{\kappa})$.

Revisando para valores grandes de κ , en este caso $\psi(\kappa) \approx \log \kappa - \frac{1}{2\kappa}$, ver propiedades 5 y 10 de la función gamma, de donde $\log(\kappa) - \frac{1}{2\kappa} - \psi(\kappa) \approx 0$. Luego fuera de las sumatorias quedará $1 + \frac{\bar{z}}{2\sqrt{\kappa}}$. Para las sumatorias se tiene $e^{z_i/\sqrt{\kappa}} = 1 - z_i/\sqrt{\kappa} + o(1/\sqrt{\kappa})$, de esta forma

$$\begin{aligned} -\frac{1}{n} \sum_{i=1}^n e^{z_i/\sqrt{\kappa}} &\doteq -\frac{1}{n} \sum_{i=1}^n \left(1 - \frac{z_i}{\sqrt{\kappa}}\right) = -1 + \frac{\bar{z}}{\sqrt{\kappa}} \\ \frac{1}{2\sqrt{\kappa}} \sum_{i=1}^n \frac{z_i}{n} e^{z_i/\sqrt{\kappa}} &\doteq \frac{1}{2\sqrt{\kappa}} \sum_{i=1}^n \frac{z_i}{n} \left(1 - \frac{z_i}{\sqrt{\kappa}}\right) = \frac{1}{2\sqrt{\kappa}} \bar{z} - \frac{1}{2\kappa} \bar{z}^2. \end{aligned}$$

Esto es $1 + \frac{\bar{z}}{2\sqrt{\kappa}} - 1 + \frac{\bar{z}}{\sqrt{\kappa}} + \frac{1}{2\sqrt{\kappa}} \bar{z} - \frac{1}{2\kappa} \bar{z}^2 = \frac{2}{\sqrt{\kappa}} \bar{z} - \frac{1}{2\kappa} \bar{z}^2$, ahora utilizando el hecho de que en general hay un mayor porcentaje de $z_i < 0$, esto es $\bar{z} < 0$, se concluye que cuando κ es grande la derivada de la función log-verosimilitud es negativa pequeña. Luego, queda concluida la primera parte.

NOTA 2.2. Con esta primera parte y tomando en consideración que la derivada de la función log-verosimilitud es continua, como resultado de operaciones elementales entre funciones continuas diferentes de cero en un mismo dominio $(0, \infty)$, se asegura que la función (2.2.10) tiene un número impar de raíces reales. Esto es equivalente a decir, que existe un número impar de valores máximos y/o mínimos relativos de la función log-verosimilitud.

Parte 2. Se mostrará que la función log-verosimilitud es cóncava para $\kappa > 0$, esto se realiza probando que su segunda derivada es negativa. Derivando la función (2.2.10) resulta

$$\ell''(\kappa; \mathbf{z}) = n \left(\frac{1}{\kappa} + \frac{1}{2\kappa^2} - \frac{\bar{z}}{4\kappa^{3/2}} - \psi'(\kappa) \right) + \frac{1}{4\kappa^{3/2}} \sum_{i=1}^n z_i e^{z_i/\sqrt{\kappa}} - \frac{1}{4\kappa^2} \sum_{i=1}^n z_i^2 e^{z_i/\sqrt{\kappa}}. \quad (2.2.11)$$

Analizando las partes por separado, la que no está en las sumatorias, se tiene que $\psi'(\kappa) \doteq \frac{1}{\kappa^2}$, luego $\frac{1}{\kappa} + \frac{1}{2\kappa^2} - \frac{\bar{z}}{4\kappa^{3/2}} - \psi'(\kappa) \doteq \frac{1}{\kappa} - \frac{\bar{z}}{4\kappa^{3/2}} - \frac{1}{2\kappa^2}$, cuando κ es pequeña el orden que domina es $\frac{1}{2\kappa^2}$, luego la parte fuera de las sumatorias es negativa. Para las sumatorias resulta más sencillo porque por orden de crecimiento para κ pequeñas domina $\frac{1}{\kappa^2}$ sobre $\frac{1}{\kappa^{3/2}}$, además este término es negativo. Luego, de ambos resultados se tiene que la segunda derivada de la función log-verosimilitud es negativa y muy grande en valor absoluto.

Para κ grandes se analizan la parte de la función (2.2.11) fuera de las sumatorias, nótese que $\psi'(\kappa) \doteq \frac{1}{\kappa} + \frac{1}{2\kappa^2}$ cuando κ es grande, de donde quedará únicamente $-\frac{\bar{z}}{4\kappa^{3/2}}$. Analizando las sumatorias y considerando que cuando κ es grande $e^{z_i/\sqrt{\kappa}} = 1 - z_i/\sqrt{\kappa} + o(1/\sqrt{\kappa})$, de esta forma al dividir entre n

$$\begin{aligned} \frac{1}{4\kappa^{3/2}} \sum_{i=1}^n \frac{z_i}{n} e^{z_i/\sqrt{\kappa}} &= \frac{1}{4\kappa^{3/2}} \left(\bar{z} - \frac{\bar{z}^2}{\sqrt{\kappa}} \right) = \frac{\bar{z}}{4\kappa^{3/2}} - \frac{\bar{z}^2}{4\kappa^2} \\ -\frac{1}{4\kappa^2} \sum_{i=1}^n z_i^2 e^{z_i/\sqrt{\kappa}} &= -\frac{1}{4\kappa^2} \left(\bar{z}^2 - \frac{\bar{z}^3}{\sqrt{\kappa}} \right) = -\frac{\bar{z}^2}{4\kappa^2} + \frac{\bar{z}^3}{4\kappa^{5/2}} \end{aligned}$$

Sumando todos los términos $-\frac{\bar{z}}{4\kappa^{3/2}} + \frac{\bar{z}}{4\kappa^{3/2}} - \frac{\bar{z}^2}{4\kappa^2} - \frac{\bar{z}^2}{4\kappa^2} + \frac{\bar{z}^3}{4\kappa^{5/2}} = -\frac{\bar{z}^2}{2\kappa^2} + \frac{\bar{z}^3}{4\kappa^{5/2}}$. Esta cantidad es pequeña pero negativa, porque para el signo el orden de decrecimiento que domina es el más pequeño, luego $\frac{1}{\kappa^2}$ domina a $\frac{1}{\kappa^{5/2}}$, además $\bar{z}^3 < 0$. Por lo tanto, cuando κ es grande la segunda derivada de la función log-verosimilitud vuelve a ser negativa aunque pequeña en valor, con esto se concluye que la función es cóncava. Falta por demostrar que la función no cambia su cóncavidad.

Parte 3. Falta mostrar que la función (2.2.11) es creciente para toda $\kappa > 0$, para esto se calcula su derivada

$$\begin{aligned} \ell'''(\kappa; \mathbf{z}) &= n \left(-\frac{1}{\kappa^2} - \frac{1}{\kappa^3} + \frac{3\bar{z}}{8\kappa^{5/2}} - \psi''(\kappa) \right) - \frac{3}{8\kappa^{5/2}} \sum_{i=1}^n z_i e^{z_i/\sqrt{\kappa}} + \frac{3}{8\kappa^3} \sum_{i=1}^n z_i^2 e^{z_i/\sqrt{\kappa}} + \\ &\quad \frac{1}{8\kappa^{7/2}} \sum_{i=1}^n z_i^3 e^{z_i/\sqrt{\kappa}}. \end{aligned}$$

Cuando κ pequeña, de la expresión entre paréntesis, resulta que $\psi''(\kappa) \approx -\frac{2}{\kappa^3}$, luego $-\frac{1}{\kappa^2} - \psi''(\kappa) \approx \frac{1}{\kappa^3}$. Por otro lado, $\bar{z} < 0$, así $\frac{1}{\kappa^2} - \frac{3\bar{z}}{8\kappa^{5/2}} \leq \frac{3\bar{z}}{4\kappa^{5/2}}$, esto implica que $-\frac{1}{\kappa^2} + \frac{3\bar{z}}{8\kappa^{5/2}} \geq -\frac{3\bar{z}}{4\kappa^{5/2}}$. Finalmente $n \left(-\frac{1}{\kappa^2} - \frac{1}{\kappa^3} + \frac{3\bar{z}}{8\kappa^{5/2}} - \psi''(\kappa) \right) \geq n \left(-\frac{3\bar{z}}{4\kappa^{5/2}} + \frac{1}{\kappa^3} \right) > 0$.

La parte de las sumatorias es más sencilla, la κ es pequeña y luego un gran porcentaje de las z_i son negativas, de tal forma $z_i/\sqrt{\kappa}$ será muy grande negativo y $e^{z_i/\sqrt{\kappa}} \approx 0$, es decreciente

a cero más rápido que la potencia a infinito. Por lo tanto si en la tercera sumatoria se separan las partes de z_i negativas de las positivas, se tendrá que en estos términos $z_i^j e^{z_i/\sqrt{\kappa}} \approx 0$ para $j = 1, 2, 3$. Denotando las sumas de partes positivas y negativas para $\sum_{i=1}^n \frac{z_i^3}{n} e^{z_i/\sqrt{\kappa}} = \Sigma_3^- + \Sigma_3^+$, luego $\Sigma_3^- \approx 0$ y la parte $\Sigma_3^+ > 0$ será bastante grande. La segunda sumatoria siempre es positiva, y en la primera sumatoria se aplica el mismo criterio. Con esto se concluye que la derivada de la función de concavidad es positiva con un valor muy grande.

Cuando κ es grande el porcentaje de z_i negativas se aproxima al 50%. En la expresión entre paréntesis se puede utilizar el hecho de que $\psi''(\kappa) \approx -\frac{1}{\kappa^2} - \frac{1}{\kappa^3}$ (ver la propiedad 5 de la función gamma), luego $-\frac{1}{\kappa^2} - \frac{1}{\kappa^3} - \psi''(\kappa) \approx 0$ quedando únicamente entre paréntesis $\frac{3\bar{z}}{8\kappa^{5/2}}$. Por otro lado en la exponencial $z_i/\sqrt{\kappa} \approx 0^-$, luego $e^{z_i/\sqrt{\kappa}} = 1 - z_i/\sqrt{\kappa} + o(1/\kappa)$ y las sumatorias se pueden realizar de la siguiente forma, después de dividir entre n :

$$\begin{aligned} -\frac{3}{8\kappa^{5/2}} \sum_{i=1}^n \frac{z_i}{n} e^{z_i/\sqrt{\kappa}} &\doteq -\frac{3}{8\kappa^{5/2}} \left(\bar{z} - \frac{1}{\sqrt{\kappa}} \bar{z}^2 \right) = -\frac{3}{8\kappa^{5/2}} \bar{z} + \frac{3}{8\kappa^3} \bar{z}^2 \\ \frac{3}{8\kappa^3} \sum_{i=1}^n \frac{z_i^2}{n} e^{z_i/\sqrt{\kappa}} &\doteq \frac{3}{8\kappa^3} \left(\bar{z}^2 - \frac{1}{\sqrt{\kappa}} \bar{z}^3 \right) = \frac{3}{8\kappa^3} \bar{z}^2 - \frac{3}{8\kappa^{7/2}} \bar{z}^3 \\ \frac{1}{8\kappa^{7/2}} \sum_{i=1}^n \frac{z_i^3}{n} e^{z_i/\sqrt{\kappa}} &\doteq \frac{1}{8\kappa^{7/2}} \left(\bar{z}^3 - \frac{1}{\sqrt{\kappa}} \bar{z}^4 \right) = \frac{1}{8\kappa^{7/2}} \bar{z}^3 - \frac{1}{8\kappa^4} \bar{z}^4. \end{aligned}$$

Sumando estos términos y considerando el término sobrante de los paréntesis, resulta

$$\frac{3}{4\kappa^3} \bar{z}^2 - \frac{1}{4\kappa^{7/2}} \bar{z}^3 - \frac{1}{8\kappa^4} \bar{z}^4.$$

Cuando κ es grande el primer término se aproxima a cero más lento que los demás y es positivo, también por el sesgo izquierdo de la distribución $\bar{z}^3 < 0$, luego el segundo término también es positivo, por lo tanto el único término negativo es el tercero, pero se aproxima a cero más rápido que los otros dos términos, con esto se concluye que también es positiva, aunque pequeña, la derivada de la función de convexidad. Así se ha demostrado que la función para la concavidad es creciente, el teorema queda demostrado.

NOTA 2.3. Con las partes segunda y tercera y la continuidad de las funciones, se concluye que la función log-verosimilitud tiene un único EMV, para $\kappa > 0$.

□

EJEMPLO 2.1

Supóngase que se tiene una realización z_1, \dots, z_{1000} de una distribución $LGG(0, 1, \kappa)$, para los valores de $\kappa = 5$ y $\kappa = 0.5$, se desea encontrar los EMV para κ , en cada caso.

Con apoyo de la función `Optimize` del proyecto R se realiza una búsqueda del parámetro κ , con la función (2.2.9), para cada una de las realizaciones correspondientes, obteniendo para la primer situación un estimador con valor $\hat{\kappa} = 4.832019\dots$ y en el segundo caso $\hat{\kappa} = 0.4941234\dots$, como se puede apreciar en las figuras 2.1.

Interpretación de las figuras 2.1, para estudiar el estimador de máxima verosimilitud. En las figuras (a) en ambos casos se tiene a la función log-verosimilitud y como se puede apreciar es cóncava y el valor de su máximo está en los valores indicados arriba. Por otro lado, los valores de la función son negativos, muy grandes en valor, para disminuir este valor se aconseja no multiplicar por el tamaño de muestra la función (2.2.9). En las figuras (b) se tienen las derivadas en ambos casos y como es de esperarse según la demostración del teorema, el valor de la derivada decrece de valores positivos muy grandes, hasta valores negativos pequeños próximos a cero, justo en donde tiene su valor máximo la función log-verosimilitud. Finalmente las figuras (c) muestran las segundas derivadas de la función log-verosimilitud y como es de esperarse de la demostración es creciente, pero siempre toma valores negativos.

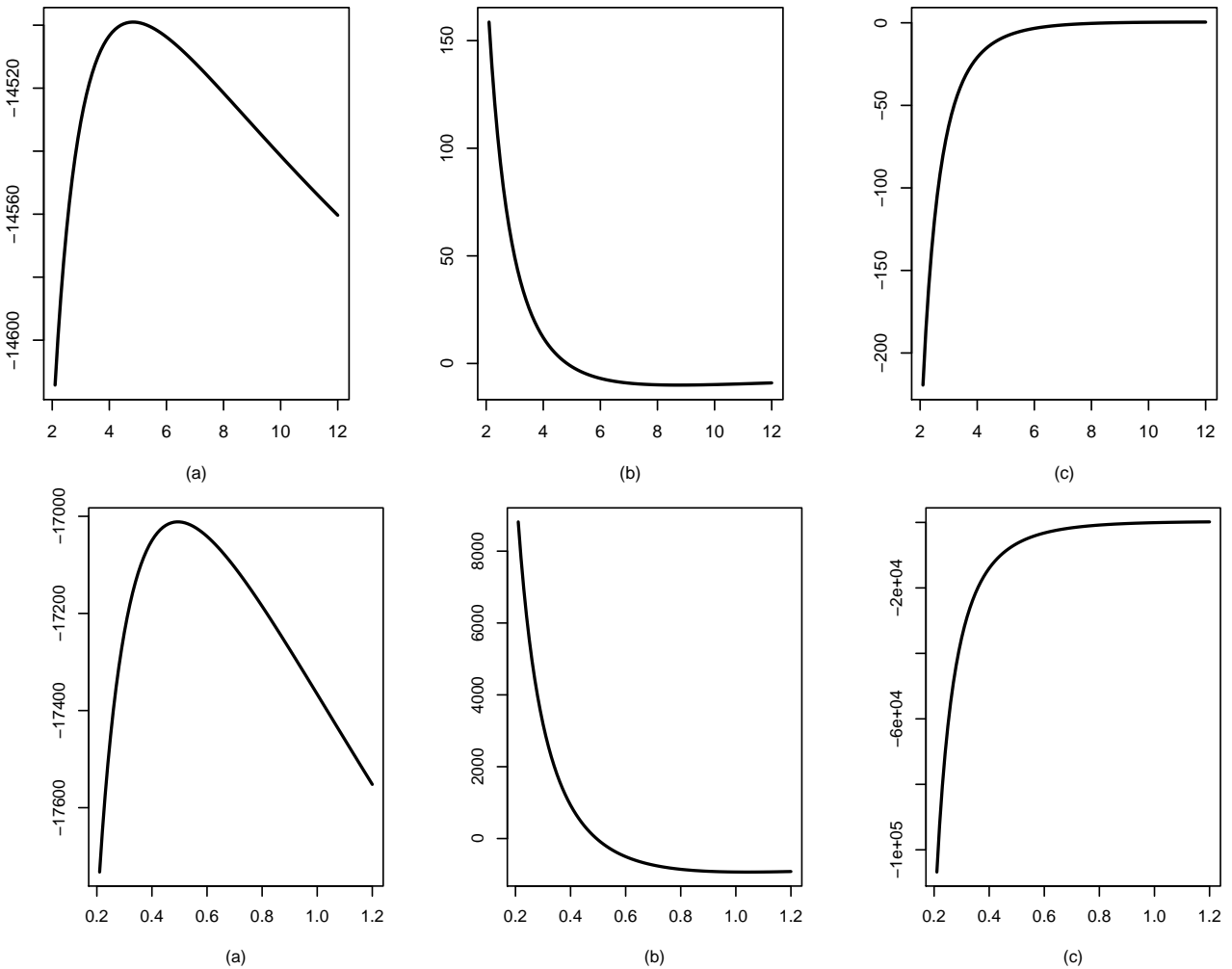


Figura 2.1: Con una muestra de tamaño $n = 1000$ y parámetro de forma $\kappa = 5$ y $\kappa = 0.5$, respectivamente: (a) Función log-verosimilitud, (b) Función de la derivada de la log-verosimilitud y (c) Función de la segunda derivada de la log-verosimilitud

2.2.3. Estimador de momentos para el parámetro de forma de $LGG(0, 1, \kappa)$

Siguiendo con la búsqueda de estimadores para el parámetro de forma de la distribución $LGG(0, 1, \kappa)$, se puede pensar en el coeficiente de asimetría de una variable aleatoria, mismo que se puede estimar por sus estimadores de momentos.

Teorema 2.2. *Sea Z una variable aleatoria $LGG(0, 1, \kappa)$, entonces para toda $\kappa > 0$ y una realización z_1, \dots, z_n se puede establecer un estimador de momentos con base en su coeficiente de asimetría muestral el cual es único para cada realización.*

Demostración. Utilizando la definición de coeficiente de asimetría (2.2.12) poblacional y las proposiciones 1.6 y 1.7 de la distribución $LGG(0, 1, \kappa)$, resulta

$$\gamma = \frac{E((Z - \mu_Z)^3)}{\left[E((Z - \mu_Z)^2)\right]^{3/2}} = \frac{\psi''(\kappa)}{(\psi'(\kappa))^{3/2}}. \quad (2.2.12)$$

Por otro lado, de las propiedades de la función polygamma, ver propiedad 10c de la función gamma, se tiene que $\psi'(\kappa)$ es una función estricta decreciente positiva en κ , $\psi'(\kappa) = \sum_{m=0}^{\infty} (1/(m + \kappa)^2)$, mientras que $\psi''(\kappa)$ es una función estrictamente creciente negativa en κ , $\psi''(\kappa) = -2 \sum_{m=0}^{\infty} (1/(m + \kappa)^3)$. Luego, el cociente que representa el coeficiente de asimetría en (2.2.12) es una función negativa estricta creciente para todo valor $\kappa > 0$.

Es decir, la función es estrictamente monótona creciente y continua (2.2.12) asigna a cada valor $\kappa > 0$ un coeficiente de asimetría entre $[-2, 0)$. Luego, por la monotonía y continuidad existe su función inversa que asigna a cada coeficiente de asimetría un valor único de κ .

La demostración finaliza asignando los estimadores de momentos para $E((Z - \mu_Z)^3)$ y $E((Z - \mu_Z)^2)$, para cada realización z_1, \dots, z_n . \square

NOTA 2.4. A partir de una realización z_1, \dots, z_n de la distribución log-gamma generalizada, se calcula su coeficiente de asimetría, sea éste

$$\hat{\gamma} = \frac{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^3}{\left(\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2\right)^{3/2}}$$

Ahora por medio de la función continua creciente $G(\kappa) = \frac{\psi''(\kappa)}{(\psi'(\kappa))^{3/2}}$ y el valor $\hat{\gamma}$, se encuentra que el estimador del parámetro κ , estará dado por $\hat{\kappa} = G^{-1}(\hat{\gamma})$. Como la función $G(\kappa) = -2 \frac{\sum_{m=0}^{\infty} \frac{1}{(\kappa+m)^3}}{\left(\sum_{m=0}^{\infty} \frac{1}{(\kappa+m)^2}\right)^{3/2}}$, su inversa se calcula de la forma que se explica en el capítulo siguiente o con un ajuste de curvas.

En el último punto de la nota 2.4 la fórmula recursiva para aproximar a $\hat{\kappa}$ se puede obtener por medio de la fórmula de Newton-Raphson que será analizada en el siguiente capítulo y para

el caso que se está analizado la fórmula recursiva queda dada por:

$$\kappa_{m+1} = \kappa_m - \frac{2\psi'(\kappa_m) \left(\psi''(\kappa_m) - \hat{\gamma}(\psi'(\kappa_m))^{3/2} \right)}{2\psi'(\kappa_m)\psi'''(\kappa) - 3(\psi''(\kappa_m))^2}. \quad (2.2.13)$$

La aproximación del valor κ_{m+1} queda determinado por la cantidad de iteraciones deseadas o por el error deseado de aproximación.

2.2.4. Estimador de momentos para la distribución $LGG(0, 1, \kappa)$ con transformación

En el desarrollo del capítulo previo se demostró que la distribución log-gamma generalizada se obtiene a partir de la distribución gamma, aumentando parámetros para hacerla más flexible, por otra lado, se requiere estimar el parámetro de forma, que desde luego no se altera (en la transformación de una variable a otra se alteran sus parámetros de escala y localización, pero no así el de forma).

Luego, supóngase que $Z \in LGG(0, 1, \kappa)$, es decir, $f_Z(z, \kappa) = \kappa^{\kappa-1/2} \exp(z\sqrt{\kappa} - \kappa \exp(z/\sqrt{\kappa}))$, por lo tanto $Y = \kappa \exp(Z/\sqrt{\kappa})$ tiene una distribución $\Gamma(\kappa, 1)$. Por otro lado, cuando se tiene una variable aleatoria $Y \in \Gamma(\kappa, 1)$, entonces $Z = \sqrt{\kappa} \log(Y/\kappa)$ tiene una distribución $LGG(0, 1, \kappa)$.

Puesto que los parámetros de forma no cambian con la transformación, del párrafo anterior se puede establecer otro estimador de momentos para el parámetro κ .

Teorema 2.3. *Sea Z una variable aleatoria $LGG(0, 1, \kappa)$, entonces para $\kappa > 0$ y una realización z_1, \dots, z_n , existe un estimador de momentos, $\hat{\kappa}$, raíz de la ecuación dado por*

$$\frac{1}{n} \sum_{i=1}^n \exp\left(\frac{z_i}{\sqrt{\hat{\kappa}}}\right) = 1. \quad (2.2.14)$$

Demostración. A través de la transformación $Y = \kappa \exp(Z/\sqrt{\kappa})$, se obtiene que $Y \in \Gamma(\kappa, 1)$, luego un estimador de κ está dado por $\kappa = \bar{Y}$, pero en cada Y_i se encuentra el parámetro, luego regresando a las variables Z_i , se tiene $\kappa = \sum_{i=1}^n \kappa \exp(z_i/\sqrt{\kappa})$. De donde se concluye el teorema. \square

NOTA 2.5. Para encontrar un valor de este estimador se utiliza la fórmula recursiva:

$$\kappa_{m+1} = \kappa_m + \frac{\kappa_m^{3/2} \left(\frac{1}{n} \sum_{i=1}^n \exp(z_i/\sqrt{\kappa_m}) - 1 \right)}{\frac{1}{n} \sum_{i=1}^n z_i \exp(z_i/\sqrt{\kappa_m})}. \quad (2.2.15)$$

De la expresión (2.2.14) se puede observar que el cálculo del estimador tendrá serios problemas en el caso de muestras pequeñas, puesto que es fácil notar que en dicha ecuación juega un papel fundamental la distribución de los valores de la variable. Puesto que si existe mayoría en algún signo la ecuación (2.2.14) no tendrá solución.

2.3. Comportamiento asintótico de los EMV

Como se sabe el método de máxima verosimilitud fue introducido por Fisher en el año de 1922, con esta técnica se pueden encontrar buenos estimadores para los parámetros de una distribución, aunque tienen la complejidad que en muchas ocasiones se dificulta encontrar la distribución de estos estimadores, debido a que la ecuación para determinar el estimador no se resuelve analíticamente y en tal caso el estimador se puede obtener únicamente de manera puntual para cada realización dada.

Una de las motivaciones de estudiar los estimadores reside en poder obtener su distribución y estudiar sus propiedades, como el insesgamiento y la consistencia en probabilidad y error cuadrado medio, con lo que se puede determinar que tan bueno es el estimador. Pero en caso de no poder obtener el estimador de máxima verosimilitud en forma explícita hace imposible determinar su distribución.

En el estudio de la distribución log-gamma generalizada en su forma estándar se propuso un estimador de máxima verosimilitud, $\hat{\kappa}_1$, para el parámetro de forma κ . Este estimador no se pudo encontrar en su forma analítica, por esta razón se estudiarán los resultados para el caso asintótico de los EMV.

En el caso de muestras grandes se conoce el método score-Fisher, con el que bajo ciertas condiciones se puede dar respuesta a la distribución asintótica del estimador de máxima verosimilitud.

Teorema 2.4. Sean X_1, \dots, X_n variables aleatorias iid con función de densidad de probabilidad $f(x; \theta)$, $\theta \in \Theta \subset \mathbb{R}$ y satisfacen las siguientes condiciones:

1).- $\frac{\partial}{\partial \theta} f(x; \theta)$ y $\frac{\partial^2}{\partial \theta^2} f(x; \theta)$ existen casi en cualquier parte, y son tales que

$$\left| \frac{\partial}{\partial \theta} f(x; \theta) \right| \leq H_1(x) \text{ y } \left| \frac{\partial^2}{\partial \theta^2} f(x; \theta) \right| \leq H_2(x)$$

en donde $\int_{\mathbb{R}} H_j(x) dx < \infty$, para $j = 1, 2$;

2).- $\frac{\partial}{\partial \theta} \log(f(x; \theta))$ y $\frac{\partial^2}{\partial \theta^2} \log(f(x; \theta))$ existen casi en cualquier parte, y son tales que

a).- X tiene la información de Fisher finita, es decir,

$$0 < I(\theta) = E \left(\left[\frac{\partial}{\partial \theta} \log(f(X; \theta)) \right]^2 \right) < \infty.$$

b).- Cuando $\delta \rightarrow 0$,

$$E \left(\sup_{\{h: |h| \leq \delta\}} \left| \frac{\partial^2}{\partial \theta^2} \log(f(X; \theta + h)) - \frac{\partial^2}{\partial \theta^2} \log(f(X; \theta)) \right| \right) = \psi_\delta \rightarrow 0.$$

Entonces el estimador de máxima verosimilitud $\hat{\theta}_n$ de θ es tal que $\sqrt{n}(\hat{\theta}_n - \theta) \rightarrow N(0, I^{-1}(\theta))$ en distribución.

La demostración se puede ver en Sen y Singer (1993) páginas 205-207. En el caso de la distribución log-gamma generalizada se obtuvo el siguiente resultado.

Teorema 2.5. Sean X_1, \dots, X_n variables aleatorias iid con distribución LGG(0, 1, κ) y $\hat{\kappa}_1$ el EMV para κ dado en (2.2.8) del teorema 2.1, entonces para $\delta > 0$ y $\kappa \in [2\delta, \infty)$ se tiene que $\sqrt{n}(\hat{\kappa}_1 - \kappa) \xrightarrow{D} N(0, I^{-1}(\kappa))$ con

$$I(\kappa) \doteq (c(\kappa) - 1)^2 + \frac{\mu_2}{4} + \sum_{j=0}^m \left(\frac{1 + 2^{j-1}}{2 \cdot j!} - \frac{2^{j+1} + 1 - c(\kappa)}{(j+1)!} + \frac{2^{j+2} - 2c(\kappa)}{(j+2)!} \right) \mu_{j+2}.$$

En donde, $c(\kappa) = \frac{\mu}{\sqrt{\kappa}} - \frac{1}{2\kappa} + 1$ y los momentos $\mu_j = E(X^j)/\kappa^{j/2}$ la igualdad se cumple cuando $m \rightarrow \infty$. Los momentos $E(X^j)$ se pueden calcular con la fgm dada en (1.4.5).

Demostración. Para probar la implicación se tienen que comprobar los supuestos del teorema 2.4, para esto se parte de la función de densidad $f(x; \kappa) = \frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \exp(x\sqrt{\kappa} - \kappa e^{x/\sqrt{\kappa}})$, y considerando que $\kappa^{\kappa-1/2} = \exp((\kappa-1/2) \log(\kappa))$ y $\Gamma(\kappa) = \exp(\log \Gamma(\kappa))$. Luego, sus primeras dos derivadas con respecto al parámetro están dadas por:

$$\begin{aligned} \frac{\partial f(x; \kappa)}{\partial \kappa} &= \frac{\partial}{\partial \kappa} \exp \left[\left(\kappa - \frac{1}{2} \right) \log(\kappa) - \log(\Gamma(\kappa)) + x\sqrt{\kappa} - \kappa e^{x/\sqrt{\kappa}} \right] \\ &= f(x; \kappa) \left[\log(\kappa) - \psi(\kappa) - \frac{1}{2\kappa} + 1 + \frac{x}{2\sqrt{\kappa}} - e^{x/\sqrt{\kappa}} + \frac{x}{2\sqrt{\kappa}} e^{x/\sqrt{\kappa}} \right] \\ \frac{\partial^2 f(x; \kappa)}{\partial \kappa^2} &= f(x; \kappa) \left(\left[\log(\kappa) - \psi(\kappa) - \frac{1}{2\kappa} + 1 + \frac{x}{2\sqrt{\kappa}} - e^{x/\sqrt{\kappa}} + \frac{x}{2\sqrt{\kappa}} e^{x/\sqrt{\kappa}} \right]^2 + \right. \\ &\quad \left. \left[\frac{1}{\kappa} - \psi'(\kappa) + \frac{1}{2\kappa^2} - \frac{x}{4\kappa^{3/2}} + \frac{x}{2\kappa^{3/2}} e^{x/\sqrt{\kappa}} - \frac{x}{4\kappa^{3/2}} e^{x/\sqrt{\kappa}} - \frac{x^2}{4\kappa^2} e^{x/\sqrt{\kappa}} \right] \right). \end{aligned}$$

Ahora se comprobará la condición 1 del teorema 2.4. En el caso de la primera derivada la parte entre corchetes es una función continua y se analizan los dos extremos del dominio de κ . Para el caso en que $\kappa \rightarrow \infty$ los términos que pueden tener problemas en la acotación son $\log(\kappa) - \psi$, pero de la propiedad 10a se tiene que $\log(\kappa) - \psi \rightarrow 0$. En el caso de que κ sea pequeña, no hay problemas porque es acotada por abajo (el intervalo es cerrado). Luego,

$$\left| \frac{\partial f(x; \kappa)}{\partial \kappa} \right| = c_1 |f(x; \kappa)| \Rightarrow H_1(x) = c_1 f(x; 1), \quad c_1 - \text{constante.}$$

En el caso de la segunda derivada se analiza de forma similar, debido a que la expresión al cuadrado es la misma que se analizó en la primera derivada, quedaría la expresión del segundo corchete, pero cuando $\kappa \rightarrow \infty$ no existen problemas, luego

$$\left| \frac{\partial^2 f(x; \kappa)}{\partial \kappa^2} \right| = c_2 |f(x; \kappa)| \Rightarrow H_2(x) = c_2 f(x; 1), \quad c_2 - \text{constante.}$$

La condición 2b del teorema 2.4 es directa porque

$$\log(f(x; \kappa)) = \left(\kappa - \frac{1}{2} \right) \log(\kappa) - \log(\Gamma(\kappa)) + x\sqrt{\kappa} - \kappa e^{x/\sqrt{\kappa}}.$$

Su segunda derivada está dada por:

$$\frac{1}{\kappa} - \psi'(\kappa) + \frac{1}{2\kappa^2} - \frac{x}{4\kappa^{3/2}} + \frac{x}{2\kappa^{3/2}}e^{x/\sqrt{\kappa}} - \frac{x}{4\kappa^{3/2}}e^{x/\sqrt{\kappa}} - \frac{x^2}{4\kappa^2}e^{x/\sqrt{\kappa}}$$

es una función continua en un intervalo cerrado en el que la función está definida y acotada.

Para la información de Fisher, se tiene

$$\left[\frac{\partial}{\partial \theta} \log (f(x; \theta)) \right]^2 = \left[\log(\kappa) - \psi(\kappa) - \frac{1}{2\kappa} + 1 + \frac{x}{2\sqrt{\kappa}} - e^{x/\sqrt{\kappa}} + \frac{x}{2\sqrt{\kappa}}e^{x/\sqrt{\kappa}} \right]^2.$$

Para elevar al cuadrado se denota $c = c(\kappa) = \log(\kappa) - \psi(\kappa) - \frac{1}{2\kappa} + 1$, para simplificar los cálculos, luego

$$\begin{aligned} \left[\frac{\partial}{\partial \theta} \log (f(x; \theta)) \right]^2 &= \left[c + \frac{x}{2\sqrt{\kappa}} - e^{x/\sqrt{\kappa}} + \frac{x}{2\sqrt{\kappa}}e^{x/\sqrt{\kappa}} \right]^2 \\ &= c^2 + \frac{x^2}{4\kappa} + e^{2x/\sqrt{\kappa}} + \frac{x^2}{4\kappa}e^{2x/\sqrt{\kappa}} + \frac{cx}{\sqrt{\kappa}} - 2ce^{x/\sqrt{\kappa}} + \frac{cx}{\sqrt{\kappa}}e^{x/\sqrt{\kappa}} - \\ &\quad - \frac{x}{\sqrt{\kappa}}e^{x/\sqrt{\kappa}} + \frac{x^2}{2\kappa}e^{x/\sqrt{\kappa}} - \frac{x}{\sqrt{\kappa}}e^{2x/\sqrt{\kappa}} \\ &= c^2 + \frac{cx}{\sqrt{\kappa}} + \frac{x^2}{4\kappa} + \left(-2c + \frac{(c-1)x}{\sqrt{\kappa}} + \frac{x^2}{2\kappa} \right) e^{\frac{x}{\sqrt{\kappa}}} + \left(1 - \frac{x}{\sqrt{\kappa}} + \frac{x^2}{4\kappa} \right) e^{\frac{2x}{\sqrt{\kappa}}}. \end{aligned}$$

Para calcular el valor esperado se puede utilizar una aproximación de las exponenciales por series de potencias con centro en cero, denotando $y = \frac{x}{\sqrt{\kappa}}$ resultan.

$$\left[\frac{\partial}{\partial \theta} \log (f(x; \theta)) \right]^2 = c^2 + cy + \frac{y^2}{4} + \left(-2c + (c-1)y + \frac{y^2}{2} \right) e^y + \left(1 - y + \frac{y^2}{4} \right) e^{2y}.$$

Desarrollando la serie $e^y = 1 + \frac{1}{1!}y + \frac{1}{2!}y^2 + \dots$, y sustituyendo en $\left(-2c + (c-1)y + \frac{y^2}{2} \right) e^y$, se realiza el producto y agrupan sus términos semejantes en y resulta:

- Potencia 0: $-2c$.
- Potencia 1: $c - 1 + (-2c) = -c - 1$.
- Potencia 2: $\frac{1}{2} + \frac{1}{1!}(c-1) + \frac{1}{2!}(-2c)$.
- Potencia 3: $\frac{1}{1!}\frac{1}{2} + \frac{1}{2!}(c-1) + \frac{1}{3!}(-2c)$.

Luego, la fórmula que se obtiene para los términos de potencias mayores a 1

$$-2c - (c+1)y + \sum_{j=0}^{\infty} \left(\frac{1}{j!} \frac{1}{2} + \frac{(c-1)}{(j+1)!} - \frac{2c}{(j+2)!} \right) y^{j+2}. \quad (2.3.1)$$

Similarmente para el otro producto $(1 - y + \frac{y^2}{4})(1 + \frac{1}{1!}2y + \frac{1}{2!}2^2y^2 + \dots)$

- Potencia 0: 1.
- Potencia 1: $-1 + 2 = 1$.
- Potencia 2: $\frac{1}{4} + \frac{1}{1!}(-2) + \frac{1}{2!}(2^2)$.
- Potencia 3: $\frac{1}{1!}\frac{1}{4} + \frac{1}{2!}(-2^2) + \frac{1}{3!}(2^3)$.

Resultando la suma:

$$1 + y + \sum_{j=0}^{\infty} \left(\frac{1}{j!} \frac{1}{4} - \frac{2}{(j+1)!} + \frac{4}{(j+2)!} \right) 2^j y^{j+2}. \quad (2.3.2)$$

Sumando (2.3.1), (2.3.2) con $c^2 + cy + \frac{y^2}{4}$, tomando una suma parcial de la serie y aplicando el valor esperado, el teorema queda demostrado. \square

Del teorema 2.5 se deduce el siguiente corolario.

Corolario 2.6. *Bajo las condiciones del teorema 2.5 se concluye:*

- a) $\hat{\kappa}_1$ es un estimador asintóticamente insesgado de κ .
- b) $\hat{\kappa}_1$ es consistente en ECM para κ .

Analizando la descomposición de la serie de Taylor para la exponencial con centro en cero se tiene que su aproximación resulta lenta, una forma de agilizar la convergencia consiste en utilizar como centro el valor que tenga mayor coincidencia, de tal forma que si se utiliza $E(X) = \mu$ la aproximación al valor de la información de Fisher resulta más rápida.

Corolario 2.7. *Bajo las condiciones del teorema 2.5 el valor de la información de Fisher está dado*

$$I(\kappa) = A_1 + \frac{\mu'_2}{4\kappa} + A_2 \sum_{j=0}^{\infty} \left(\frac{1}{2j!} - \frac{1}{2\kappa(j+1)!} + \frac{A_3}{(j+2)!} \right) \frac{\mu'_{j+2}}{\kappa^{\frac{j+2}{2}}} + A_2^2 \sum_{j=0}^{\infty} \left(\frac{1}{4j!} + \frac{2A_4}{(j+1)!} + \frac{4A_4^2}{(j+2)!} \right) 2^j \frac{\mu'_{j+2}}{\kappa^{\frac{j+2}{2}}}.$$

En donde, $\mu = E(X)$, $\mu'_j = E((X - \mu)^j)$ son los momentos centrales, $A_4 = \frac{\mu}{2\sqrt{\kappa}} - 1$, $A_3 = \frac{\mu^2}{2\kappa} + \frac{(c-1)\mu}{\sqrt{\kappa}} - 2c$, $A_2 = e^{\psi(\kappa)}/\kappa$, y $A_1 = \left(\frac{\mu}{2\sqrt{\kappa}} + c\right)^2 + A_2A_3 + A_2^2A_4^2$.

Demostración. En la demostración se conservará la notación establecida en la formulación del corolario. Partiendo de la relación

$$\left[\frac{\partial}{\partial \theta} \log(f(x; \theta)) \right]^2 = c^2 + \frac{cx}{\sqrt{\kappa}} + \frac{x^2}{4\kappa} + \left(-2c + \frac{(c-1)x}{\sqrt{\kappa}} + \frac{x^2}{2\kappa} \right) e^{\frac{x}{\sqrt{\kappa}}} + \left(1 - \frac{x}{\sqrt{\kappa}} + \frac{x^2}{4\kappa} \right) e^{\frac{2x}{\sqrt{\kappa}}}. \quad (2.3.3)$$

Sustituyendo en la ecuación (2.3.3) $x = x - \mu + \mu = y + \mu$ y $x^2 = y^2 + 2\mu y + \mu^2$ en donde $y = x - \mu$ y $A_5 := c/\sqrt{\kappa} + \mu/2\kappa$, resulta

$$\begin{aligned} \left[\frac{\partial}{\partial \theta} \log(f(x; \theta)) \right]^2 &= \left(\frac{\mu}{2\sqrt{\kappa}} + c \right)^2 + A_5 y + \frac{y^2}{4\kappa} + e^{\frac{\mu}{\sqrt{\kappa}}} \left(A_3 + \left(\frac{c-1}{\sqrt{\kappa}} + \frac{\mu}{\kappa} \right) y + \frac{1}{2\kappa} y^2 \right) e^{\frac{y}{\sqrt{\kappa}}} \\ &\quad + e^{\frac{2\mu}{\sqrt{\kappa}}} \left(\left(\frac{\mu}{2\sqrt{\kappa}} - 1 \right)^2 + \left(\frac{\mu}{2\sqrt{\kappa}} - 1 \right) \frac{y}{\sqrt{\kappa}} + \frac{1}{4\kappa} y^2 \right) e^{\frac{2y}{\sqrt{\kappa}}} \end{aligned}$$

Considerando que la descomposición en series de Taylor con $y = 0$, se tiene

$$\begin{aligned} e^{\frac{y}{\sqrt{\kappa}}} &= 1 + \frac{1}{1!} \left(\frac{1}{\sqrt{\kappa}} \right) y + \frac{1}{2!} \left(\frac{1}{\sqrt{\kappa}} \right)^2 y^2 + \frac{1}{3!} \left(\frac{1}{\sqrt{\kappa}} \right)^3 y^3 + \dots \\ e^{\frac{2y}{\sqrt{\kappa}}} &= 1 + \frac{1}{1!} \left(\frac{2}{\sqrt{\kappa}} \right) y + \frac{1}{2!} \left(\frac{2}{\sqrt{\kappa}} \right)^2 y^2 + \frac{1}{3!} \left(\frac{2}{\sqrt{\kappa}} \right)^3 y^3 + \dots \end{aligned}$$

Luego, realizando el producto por cada una de las series. Para la primer serie se tiene

$$\left(A_3 + \left(\frac{c-1}{\sqrt{\kappa}} + \frac{\mu}{\kappa} \right) y + \frac{1}{2\kappa} y^2 \right) \left(1 + \frac{1}{1!} \left(\frac{1}{\sqrt{\kappa}} \right) y + \frac{1}{2!} \left(\frac{1}{\sqrt{\kappa}} \right)^2 y^2 + \frac{1}{3!} \left(\frac{1}{\sqrt{\kappa}} \right)^3 y^3 + \dots \right).$$

Al realizar el producto se van agrupando los términos por la potencia de y , obteniendo

- Potencia 0: A_3 .
- Potencia 1: $\left(\frac{c-1}{\sqrt{\kappa}} + \frac{\mu}{\kappa} \right) + \frac{1}{1!} \left(\frac{1}{\sqrt{\kappa}} \right) A_3$ en donde $\left(\frac{c-1}{\sqrt{\kappa}} + \frac{\mu}{\kappa} \right) = \frac{1}{\sqrt{\kappa}} \left(c - 1 + \frac{\mu}{\sqrt{\kappa}} \right) = -\frac{1}{2\kappa\sqrt{\kappa}}$.
- Potencia 2: $\frac{1}{2\kappa} + \frac{1}{1!} \left(\frac{1}{\sqrt{\kappa}} \right) \left(-\frac{1}{2\kappa\sqrt{\kappa}} \right) + \frac{1}{2!} \left(\frac{1}{\sqrt{\kappa}} \right)^2 A_3$
- Potencia 3: $\frac{1}{1!} \left(\frac{1}{\sqrt{\kappa}} \right) \frac{1}{2\kappa} - \frac{1}{2!} \left(\frac{1}{\sqrt{\kappa}} \right)^2 \frac{1}{2\kappa\sqrt{\kappa}} + \frac{1}{3!} \left(\frac{1}{\sqrt{\kappa}} \right)^3 A_3$

Luego, de las potencias 2 y 3 puede verse la regla para el producto

$$\begin{aligned} A_3 - \frac{1}{2\kappa\sqrt{\kappa}} y + \sum_{j=0}^{\infty} \left[\frac{1}{j!} \left(\frac{1}{\sqrt{\kappa}} \right)^j \frac{1}{2\kappa} - \frac{1}{(j+1)!} \left(\frac{1}{\sqrt{\kappa}} \right)^{j+1} \frac{1}{2\kappa\sqrt{\kappa}} + \frac{1}{(j+2)!} \left(\frac{1}{\sqrt{\kappa}} \right)^{j+2} A_3 \right] y^{j+2} \\ A_3 - \frac{1}{2\kappa\sqrt{\kappa}} y + \sum_{j=0}^{\infty} \left[\frac{1}{2j!} - \frac{1}{2\kappa(j+1)!} + \frac{1}{(j+2)!} A_3 \right] \left(\frac{y}{\sqrt{\kappa}} \right)^{j+2}. \end{aligned} \quad (2.3.4)$$

Para la otra serie se hace lo mismo, denotando por $A_4 = \frac{\mu}{2\sqrt{\kappa}} - 1$, el producto será

$$\left(A_4^2 + y \frac{A_4}{\sqrt{\kappa}} + \frac{1}{4\kappa} y^2 \right) \left(1 + \frac{1}{1!} \left(\frac{2}{\sqrt{\kappa}} \right) y + \frac{1}{2!} \left(\frac{2}{\sqrt{\kappa}} \right)^2 y^2 + \frac{1}{3!} \left(\frac{2}{\sqrt{\kappa}} \right)^3 y^3 + \dots \right)$$

- Potencia 0: A_4^2 .

- Potencia 1: $A_6 := A_4/\sqrt{\kappa} + 2A_4^2/\sqrt{\kappa}$.
- Potencia 2: $\frac{1}{4\kappa} + \frac{1}{1!} \left(\frac{2}{\sqrt{\kappa}}\right) \frac{A_4}{\sqrt{\kappa}} + \frac{1}{2!} \left(\frac{2}{\sqrt{\kappa}}\right)^2 A_4^2$
- Potencia 3: $\frac{1}{1!} \left(\frac{2}{\sqrt{\kappa}}\right) \frac{1}{4\kappa} + \frac{1}{2!} \left(\frac{2}{\sqrt{\kappa}}\right)^2 \frac{A_4}{\sqrt{\kappa}} + \frac{1}{3!} \left(\frac{2}{\sqrt{\kappa}}\right)^3 A_4^2$

Luego, de las potencias 2 y 3 puede verse la regla para el producto

$$A_4^2 + A_6y + \sum_{j=0}^{\infty} \left[\frac{1}{j!} \left(\frac{2}{\sqrt{\kappa}}\right)^j \frac{1}{4\kappa} + \frac{1}{(j+1)!} \left(\frac{2}{\sqrt{\kappa}}\right)^{j+1} \frac{A_4}{\sqrt{\kappa}} + \frac{1}{(j+2)!} \left(\frac{2}{\sqrt{\kappa}}\right)^{j+2} A_4^2 \right] y^{j+2}$$

$$A_4^2 + A_6y + \sum_{j=0}^{\infty} \left[\frac{1}{4j!} + \frac{2A_4}{(j+1)!} + \frac{4A_4^2}{(j+2)!} \right] 2^j \left(\frac{y}{\sqrt{\kappa}}\right)^{j+2}. \quad (2.3.5)$$

Ahora se multiplica la expresión (2.3.4) por A_2 y la expresión (2.3.5) por A_2^2 y se suman con A_1 . Finalmente se calcula su valor esperado, para esto se considera que $E(Y) = 0$, por lo tanto los momentos centrales están dados por $E((X - \mu)^j) = E(Y^j) = \mu'_j$. \square

En las siguientes gráficas (2.2)-(2.5) se muestran diferentes formas en que se aproximan los momentos a la distribución asintótica del estimador de máxima verosimilitud para el parámetro de forma de la distribución $LGG(0, 1, \kappa)$. En general, en todas las gráficas que se presenten para ilustrar la aproximación con los momentos centrales de la información de Fisher se tiene:

- 1).- La curva en negro representa la distribución de los estimadores de κ , obtenida con 10,000 realizaciones y $n = 10,000$,
- 2).- la curva en azul ($\hat{\kappa}_{1m}$) corresponde a la distribución normal con los valores de las realizaciones de los estimadores y
- 3).- la gráfica en rojo ($\hat{\kappa}_{1a}$) corresponde a la aproximación de la distribución normal, obtenida con los parámetros de la información de fisher.

Para $\kappa = 0.1$ y $\kappa = 0.5$ se obtuvieron, buenas aproximaciones con únicamente tres momentos centrales, ver gráficas 2.2.

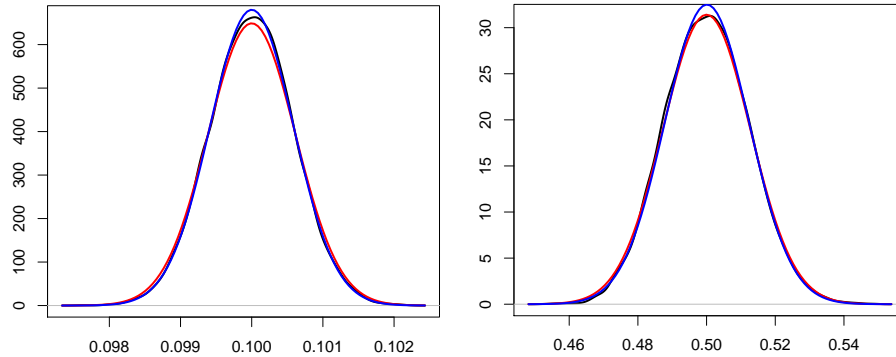


Figura 2.2: La figura izquierda para $\kappa = 0.1$ con $\hat{\kappa}_{1a} \sim N(0.1, 3.7788 \times 10^{-7})$ y $\hat{\kappa}_{1m} \sim N(0.1, 3.4408 \times 10^{-7})$. En la derecha para $\kappa = 0.5$ con $\hat{\kappa}_{1a} \sim N(0.5, 0.000161)$ y $\hat{\kappa}_{1m} \sim N(0.5, 0.000151)$, ambas aproximaciones se obtuvieron hasta el tercer momento.

En las gráficas 2.3 se muestra cómo evoluciona la aproximación para $\kappa = 1$ con momentos centrales desde 2 hasta 7.

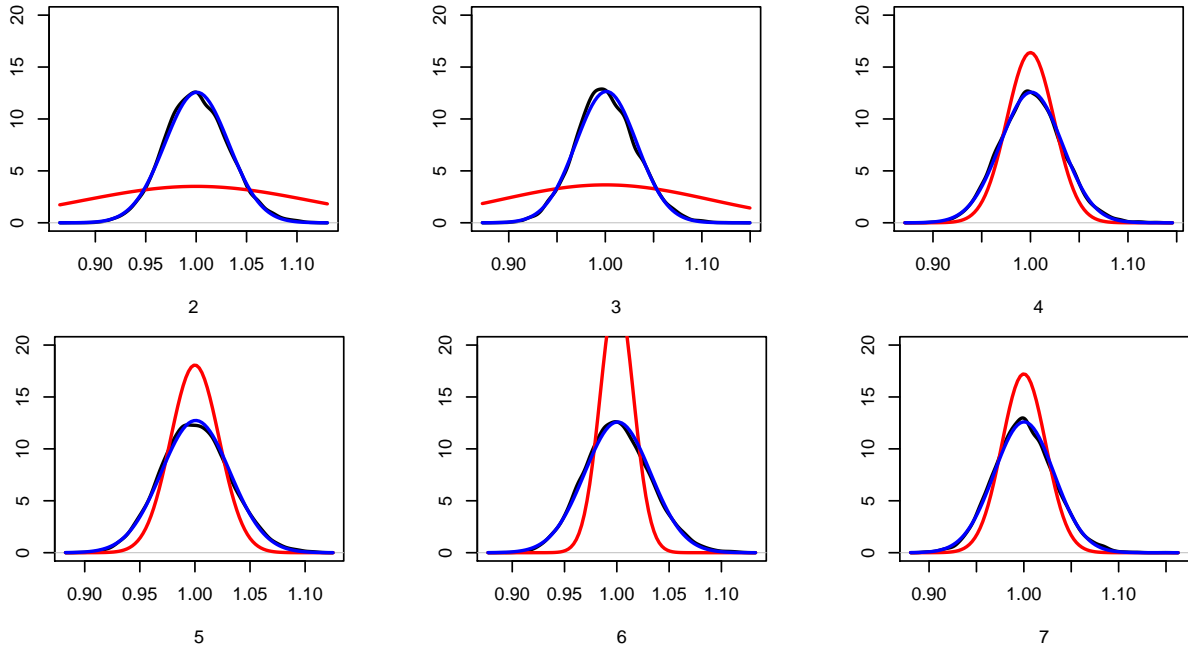


Figura 2.3: El número de la gráfica representa hasta que momento se usó para la aproximación de la distribución del EMV para $\kappa = 1$. La distribución $\hat{\kappa}_{1m} \sim N(1, 0.0010)$ y las aproximaciones: $\hat{\kappa}_{1a} \sim N(1, 0.012940)$, $N(1, 0.01194)$, $N(1, 0.00059)$, $N(1, 0.00049)$, $N(1, 0.00025)$ y $N(1, 0.00054)$, respectivamente.

Las aproximaciones con los momentos pueden ser de otros tipos, en algunos casos la distribución puede quedar con colas más pesadas y al aumentar momentos las colas pueden pasar a ser más ligeras y viceversa, como se muestra en la figuras 2.4.

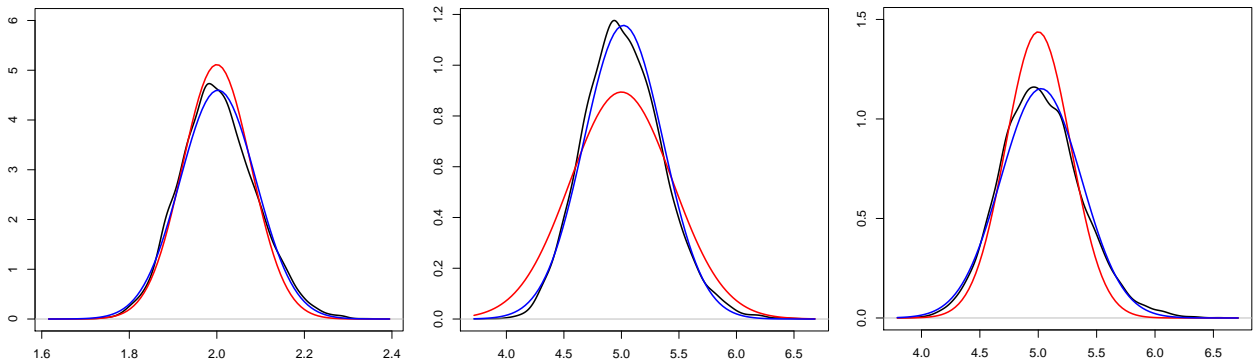


Figura 2.4: En la gráfica izquierda se utilizó hasta el sexto momento para $\kappa = 2$ con $\hat{\kappa}_{1a} \sim N(2, 0.0061)$ y $\hat{\kappa}_{1m} \sim N(2, 0.0075)$. En las otras dos gráficas representan la distribución para $\kappa = 5$ en la primera hasta el quinto momento con $\hat{\kappa}_{1a} \sim N(5, 0.1990)$ y $\hat{\kappa}_{1m} \sim N(5, 0.1189)$ en la gráfica de la derecha hasta el sexto momento con $\hat{\kappa}_{1a} \sim N(5, 0.0770)$ y $\hat{\kappa}_{1m} \sim N(5, 0.1199)$.

El último caso que será revisado corresponde a $\kappa = 10$, para aproximaciones de sexto y séptimo momentos, donde se muestra una mejoría considerable con el séptimo momento, ver figuras 2.5.

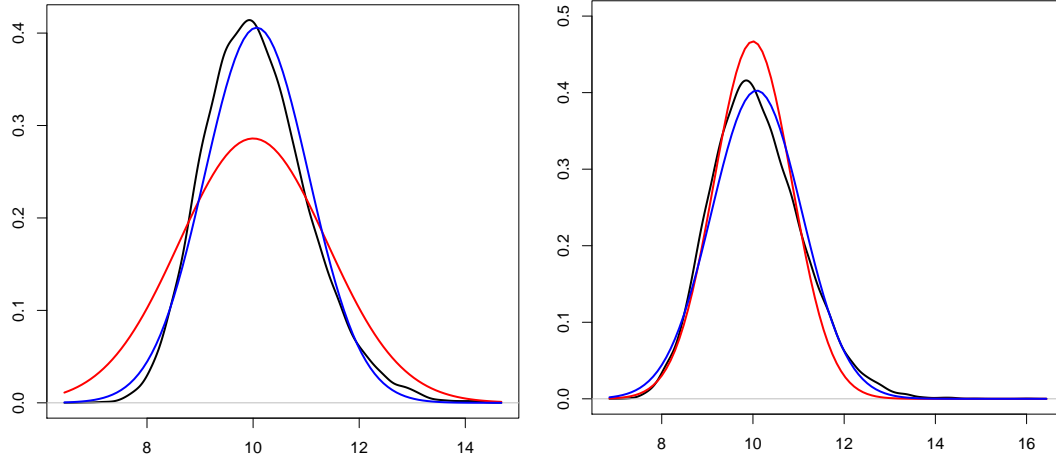


Figura 2.5: En la gráfica izquierda se utilizó hasta el sexto momento para $\kappa = 10$ con $\hat{\kappa}_{1a} \sim N(10, 1.945957)$ y $\hat{\kappa}_{1m} \sim N(10, 0.966055)$. En la gráfica de la derecha hasta el séptimo momento con $\hat{\kappa}_{1a} \sim N(10, 0.7297578)$ y $\hat{\kappa}_{1m} \sim N(10, 0.9814303)$.

2.3.1. ¿Cómo determinar la distribución asintótica de $\hat{\kappa}_1$?

La distribución asintótica de $\hat{\kappa}_1$ está dada en el teorema 2.5

$$\hat{\kappa}_1 \xrightarrow[n \rightarrow +\infty]{} N\left(k, \frac{1}{n} I^{-1}(k)\right). \quad (2.3.6)$$

Se recomienda calcular la matriz de información de Fisher con el resultado del corolario 2.7 y no con la del teorema 2.5. Para esto se puede establecer la siguiente metodología.

- 1).- Calcular el estimador de máxima verosimilitud del parámetro κ , sea éste $\hat{\kappa}_0$.
- 2).- El valor $\hat{\kappa}_0$ sustituirlo en la expresión del corolario 2.7, para esto se utilizan las fórmulas de los momentos centrales dadas en 2.7.
- 3).- El valor de la varianza para la distribución asintótica se obtiene de forma aproximada y depende de la cantidad de momentos centrales elegidos.

Con esta metodología se trazaron las gráficas (2.2)-(2.5).

Por ejemplo, si $n = 1000$ y resulta $\hat{\kappa}_0 = 4$, entonces hasta el quinto momento central la distribución de $\hat{\kappa}_1 \sim N(4, 0.7981)$. Para los cálculos se utilizó la función en el proyecto R dada en el anexo correspondiente al capítulo 2.

2.4. Comportamiento asintótico del estimador de momentos

En el estudio de los estimadores para el parámetro de forma, κ , de la distribución log-gamma generalizada se propuso el estimador de momentos del coeficiente de asimetría $\hat{\gamma}$, y se probó que este estimador es único para cada valor de κ

$$\hat{\gamma} = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^3}{\left(\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2\right)^{3/2}}, \text{ donde } m_j = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^j \text{ para } j = 1, 2, \dots$$

Además se demostró que su correspondiente función de momentos centrales

$$G(\kappa) = \frac{\mu'_3}{(\mu'_2)^{3/2}} = \frac{\psi''(\kappa)}{(\psi'(\kappa))^{3/2}}, \text{ donde } \mu'_j = E((Z - \mu)^j) \text{ para } j = 1, 2, \dots$$

es monótona creciente, por consiguiente tiene inversa, luego

$$G(\kappa) = \gamma \Rightarrow \kappa = G^{-1}(\gamma).$$

Se encontró la función G en forma analítica, pero depende de las funciones polygama (trigamma y cuatrigamma), por tal razón no se tiene una expresión analítica satisfactoria para la función inversa de G . Pero se puede proceder a buscar por mínimos cuadrados la curva que mejor ajuste. Así, el problema de transformar $(0, \infty) \xrightarrow{G} (-2, 0)$ por medio de la función G y su inversa $(-2, 0) \xrightarrow{G^{-1}} (0, \infty)$, se resuelve ajustando la curva.

Para el ajuste se tomaron valores de κ , de tal forma que el coeficiente de asimetría vario sus valores desde -1.998 hasta -0.002 con incrementos de 2 milésimas, de esta forma al realizar el ajuste se encontró la curva de mejor ajuste

$$\kappa = G^{-1}(\gamma) \doteq \frac{1.0517}{(-\gamma)^{2.033}}, \text{ con } r^2 = 0.9785. \quad (2.4.1)$$

Con esta función los mejores ajustes se obtienen cuando $\kappa \in (0.26, 50)$ o $\gamma \in (-1.8, -0.14)$, los errores relativos son menores a 0.2 y en la mayoría de ellos menores a 0.05.

Note que el ajuste (2.4.1) es próximo a $1/\gamma^2$, pero esta función ajusta bien únicamente cuando $\kappa \in (30, 62500)$ o $\gamma \in (-0.184, 0)$.

NOTA 2.6. Al realizar el ajuste de curvas se encontró la curva que mejor ajusta y está dada en (2.4.1) con un coeficiente de correlación $r^2 = 0.978$, pero realizando un ajuste con una función seccionada, se tienen mejores ajustes. Así, la curva seccionada que mejor ajusta está dada por:

$$\kappa = \begin{cases} 0.9367 \left(-\frac{\mu'_3}{\mu'_2} \right) + 1.9411, & \text{si } \gamma < -1.500, \quad r^2 = 0.9963 \\ 1.3520 \left(-\frac{\mu'_2}{\mu'_3} \right)^{2.231}, & \text{si } -1.500 \leq \gamma < -1.102, \quad r^2 = 0.9978 \\ 1.2101 \left(-\frac{\mu'_2}{\mu'_3} \right)^{1.872}, & \text{si } -1.102 \leq \gamma < -0.315, \quad r^2 = 0.9999 \\ 1.0078 \left(-\frac{\mu'_2}{\mu'_3} \right)^{1.998}, & \text{si } -0.315 \leq \gamma < 0, \quad r^2 \doteq 1 \end{cases} \quad (2.4.2)$$

Con respecto al estimador de momentos se tiene que sustituir en (2.4.1) γ por su esti-

mador, obteniendo:

$$\hat{\kappa}_2 = 1.0517 \frac{(m_2)^{3.0495}}{(-m_3)^{2.033}}. \quad (2.4.3)$$

En este momento falta obtener la distribución asintótica del estimador (2.4.3), para esto son necesarios los siguientes resultados, que se aplican para la convergencia en probabilidad y distribución.

Teorema 2.8 (Cramér-Wold). Sean $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \dots$ vectores aleatorios en \mathbb{R}^p ; entonces $\mathbf{X}_n \xrightarrow{D} \mathbf{X}$ si y solamente si, para cualquier vector fijo $\boldsymbol{\lambda} \in \mathbb{R}^p$ se tiene $\boldsymbol{\lambda}^t \mathbf{X}_n \xrightarrow{D} \boldsymbol{\lambda}^t \mathbf{X}$.

La demostración se puede ver en Sen y Singer (1993) página 106.

Teorema 2.9 (Slutsky). Sean $\{X_n\}$ y $\{Y_n\}$ sucesiones de variables variables aleatorias tales que $X_n \xrightarrow{D} X$ y $Y_n \xrightarrow{p} c$ donde c es una constante. Entonces se cumple lo siguiente:

- 1).- $X_n + Y_n \xrightarrow{D} X + c$.
- 2).- $X_n Y_n \xrightarrow{D} cX$.
- 3).- $X_n/Y_n \xrightarrow{D} X/c$, si $c \neq 0$.

La demostración se puede ver en Sen y Singer (1993) página 127.

Ahora en el caso cuando la estadística puede ser expresada como una suma de variables aleatorias independientes, como son los momentos muestrales, se obtiene una estadística asintóticamente normal.

Teorema 2.10. Considérese la estadística $T_n = T(X_1, \dots, X_n)$ donde las X_i 's son variables aleatorias iid. Sea $T_n = G_n + R_n$ donde $G_n = \sum_{i=1}^n g(X_i)$ y $n^{-1/2} R_n \xrightarrow{p} 0$. Además, sea $E(g(X_i)) = \xi$ y $Var(g(X_i)) = \nu^2 < \infty$ y supóngase que $(G_n - n\xi)/\sqrt{n\nu} \xrightarrow{D} N(0, 1)$. Entonces $(T_n - n\xi)/\sqrt{n\nu} \xrightarrow{D} N(0, 1)$.

La demostración se puede ver en Sen y Singer (1993) página 126.

En un caso más general para la distribución asintótica se encuentra el **método Delta**, que da respuesta cuando la estadística de interés está afectada por una función continua.

Teorema 2.11. Suponga que $\sqrt{n}(T_n - \theta)/\sigma \xrightarrow{D} N(0, 1)$ y sea g una función continua tal que $g'(\theta)$ existe y $g'(\theta) \neq 0$. Entonces se sigue que

$$\sqrt{n} \frac{(g(T_n) - g(\theta))}{\sigma g'(\theta)} \xrightarrow{D} N(0, 1).$$

La demostración se puede ver en Sen y Singer (1993) página 131.

Una generalización de este teorema al caso multivariado tiene muchas aplicaciones, cuando se trata de estadísticas que están relacionadas entre si.

Teorema 2.12. Sea $\{\mathbf{T}_n\}$ una sucesión de p -vectores aleatorios, tales que $\sqrt{n}(\mathbf{T}_n - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma})$ y considere una función valuada real $g(\mathbf{T}_n)$ tal que $\dot{g}(\boldsymbol{\theta}) = \left. \frac{\partial g(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\boldsymbol{\theta}}$ no es cero y continua en una vecindad de $\boldsymbol{\theta}$. Entonces

$$\sqrt{n}(g(\mathbf{T}_n) - g(\boldsymbol{\theta})) \xrightarrow{D} N(0, \tau^2) \text{ con } \tau^2 = (\dot{g}(\boldsymbol{\theta}))^t \boldsymbol{\Sigma} (\dot{g}(\boldsymbol{\theta})).$$

La demostración se puede ver en Sen y Singer (1993) página 133.

Ahora se formulará y demostrará un resultado para el desarrollo del trabajo.

Teorema 2.13. Sea $\{X_n\}$ una sucesión de variables aleatorias iid con media μ y varianza σ^2 , además $t_1, t_2 \in \mathbb{N}$ con $t_1 > t_2$ y $E(X_1^{2t_1}) < \infty$. Denótese al vector $\mathbf{T}_n = (m_{t_1} - \mu'_{t_1}, m_{t_2} - \mu'_{t_2})^t$ con m_t y μ'_t los t -ésimos momentos centrales muestrales y poblacionales, respectivamente. Entonces la distribución asintótica de

$$\sqrt{n}\mathbf{T}_n \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma}), \text{ con } \boldsymbol{\Sigma} = \begin{pmatrix} \mu'_{2t_1} - \mu'_{t_1}{}^2 & \mu'_{t_1+t_2} - \mu'_{t_1}\mu'_{t_2} \\ \mu'_{t_1+t_2} - \mu'_{t_1}\mu'_{t_2} & \mu'_{2t_2} - \mu'_{t_2}{}^2 \end{pmatrix}.$$

Demostración. Sea la diferencia de los j -ésimos momentos centrales muestrales y poblacionales

$$\begin{aligned} m_j - \mu'_j &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^j - \mu'_j = \frac{1}{n} \sum_{i=1}^n ((X_i - \mu) - (\bar{X} - \mu))^j - \mu'_j \\ &= \frac{1}{n} \sum_{i=1}^n ((X_i - \mu)^j - \mu'_j) + \frac{1}{n} \sum_{i=1}^n \sum_{s=0}^{j-1} (-1)^{s+1} \binom{j}{s} (X_i - \mu)^s (\bar{X} - \mu)^{j-s}. \end{aligned}$$

Sustituyendo j por t_1 y t_2 y denotando por

$$\begin{aligned} G_{n,t_m} &= \frac{1}{n} \sum_{i=1}^n ((X_i - \mu)^{t_m} - \mu'_{t_m}), \text{ y } Y_{i,t_m} = (X_i - \mu)^{t_m} - \mu'_{t_m} \text{ con } m = 1, 2 \\ R_{n,t_m} &= \frac{1}{n} \sum_{i=1}^n \sum_{s=0}^{t_m-1} (-1)^{s+1} \binom{t_m}{s} (X_i - \mu)^s (\bar{X} - \mu)^{t_m-s}, \text{ con } m = 1, 2 \end{aligned}$$

Ahora se calcula el valor esperado y la varianza de Y_{i,t_m} , obteniendo:

$$\begin{aligned} E(Y_{i,t_m}) &= E((X_i - \mu)^{t_m} - \mu'_{t_m}) = 0 \\ \text{Var}(Y_{i,t_m}) &= \text{Var}((X_i - \mu)^{t_m} - \mu'_{t_m}) = \text{Var}((X_i - \mu)^{t_m}) \\ &= E((X_i - \mu)^{2t_m}) - [E((X_i - \mu)^{t_m})]^2 \\ &= \mu'_{2t_m} - \mu'_{t_m}{}^2. \end{aligned}$$

Por otro lado, si se desarrolla el producto que se encuentra en R_{n,t_m} , y se calcula su valor esperado, se tiene que $E(X_i^{2t_m}) \leq E(X_1^{2t_1}) < \infty$, luego $E(\sqrt{n}R_{n,t_m}) = o(1)$.

Fijando dos valores arbitrarios $\lambda_1, \lambda_2 \in \mathbb{R}$, se tiene

$$\begin{aligned} Z_n &:= \sqrt{n}(\lambda_1(m_{t_1} - \mu'_{t_1}) + \lambda_2(m_{t_2} - \mu'_{t_2})) \\ &= \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n (\lambda_1 Y_{i,t_1} + \lambda_2 Y_{i,t_2}) \right) + \sqrt{n}(\lambda_1 R_{n,t_1} + \lambda_2 R_{n,t_2}) \\ &= \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n W_i \right) + \sqrt{n}(\lambda_1 R_{n,t_1} + \lambda_2 R_{n,t_2}) \end{aligned}$$

En donde $W_i = \lambda_1 Y_{i,t_1} + \lambda_2 Y_{i,t_2}$, para $i = 1, 2, \dots, n$ son variables aleatorias independientes, tales que

$$\begin{aligned} E(W_i) &= \lambda_1 E(Y_{i,t_1}) + \lambda_2 E(Y_{i,t_2}) = 0 \\ Var(W_i) &= \lambda_1^2 Var(Y_{i,t_1}) + \lambda_2^2 Var(Y_{i,t_2}) + 2\lambda_1 \lambda_2 Cov(Y_{i,t_1}, Y_{i,t_2}) \end{aligned}$$

Para la covarianza

$$\begin{aligned} Cov(Y_{i,t_1}, Y_{i,t_2}) &= E((X_i - \mu)^{t_1} - \mu'_{t_1})((X_i - \mu)^{t_2} - \mu'_{t_2}) \\ &= E((X_i - \mu)^{t_1+t_2}) - \mu'_{t_1} \mu'_{t_2} \\ &= \mu'_{t_1+t_2} - \mu'_{t_1} \mu'_{t_2} \end{aligned}$$

Así, $E(W_i) = 0$ y $Var(W_i) = \lambda_1^2(\mu'_{2t_1} - \mu'_{t_1}{}^2) + \lambda_2^2(\mu'_{2t_2} - \mu'_{t_2}{}^2) + 2\lambda_1 \lambda_2(\mu'_{t_1+t_2} - \mu'_{t_1} \mu'_{t_2})$, con el teorema clásico del límite central y el teorema de Slutsky y (2.9) para la convergencia en distribución se concluye que Z_n tiene una distribución normal asintótica. Pero λ_1 y λ_2 son arbitrarios, es decir se tiene una forma cuadrática y por consiguiente del teorema de Cramér-Wold (2.8) se concluye que el vector \mathbf{T}_n tiene una distribución normal asintótica dada por

$$\sqrt{n} \begin{pmatrix} m_{t_1} - \mu'_{t_1} \\ m_{t_2} - \mu'_{t_2} \end{pmatrix} = \sqrt{n} \mathbf{T}_n \xrightarrow{D} N(\mathbf{0}, \Sigma), \text{ con } \Sigma = \begin{pmatrix} \mu'_{2t_1} - \mu'_{t_1}{}^2 & \mu'_{t_1+t_2} - \mu'_{t_1} \mu'_{t_2} \\ \mu'_{t_1+t_2} - \mu'_{t_1} \mu'_{t_2} & \mu'_{2t_2} - \mu'_{t_2}{}^2 \end{pmatrix}.$$

El teorema queda demostrado □

Del teorema anterior se requieren los momentos centrales para determinar la matriz de varianzas y covarianzas de la distribución asintótica, para el caso de la distribución log-gamma generalizada estándar se proporciona un método para calcularlos.

Teorema 2.14. *Sea $T \in LGG(0, 1, \kappa)$ con media μ , entonces su n -ésimo momento central se obtiene de $\frac{(\sqrt{\kappa})^n}{\Gamma(\kappa)} \Gamma^{(n)}(v) \Big|_{v=\kappa}$*

$$E((T - \mu)^n) = \frac{(\sqrt{\kappa})^n}{\Gamma(\kappa)} \times \text{Términos de } \{\Gamma^{(n)}(v) \Big|_{v=\kappa}\} \text{ que no contengan a } \psi^j(\kappa) \text{ para } j \geq 1.$$

Demostración. El n -ésimo momento central $\mu'_n = E((T - \mu)^n)$ se obtiene desarrollando el

binomio y sustituyendo $E(T^i)$ por la expresión (1.4.5) y $\mu = \sqrt{\kappa}(\psi(\kappa) - \log(\kappa))$

$$\begin{aligned}
 \mu'_n &= E\left(\sum_{i=0}^n (-1)^{n-i} \binom{n}{i} T^i \mu^{n-i}\right) = \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} \mu^{n-i} E(T^i) \\
 &= \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} [\sqrt{\kappa}(\psi(\kappa) - \log(\kappa))]^{n-i} \frac{(\sqrt{\kappa})^i}{\Gamma(\kappa)} \sum_{j=0}^i (-1)^j \binom{i}{j} \Gamma^{(i-j)}(v) \Big|_{v=\kappa} \log^j(\kappa) \\
 &= \sum_{i=0}^n \sum_{j=0}^i (-1)^{n-i+j} \binom{n}{i} \binom{i}{j} \frac{(\sqrt{\kappa})^n}{\Gamma(\kappa)} \Gamma^{(i-j)}(v) \Big|_{v=\kappa} \log^j(\kappa) \sum_{s=0}^{n-i} (-1)^s \binom{n-i}{s} \log^s(\kappa) \psi^{n-i-s}(\kappa) \\
 &= \frac{(\sqrt{\kappa})^n}{\Gamma(\kappa)} \sum_{i=0}^n \sum_{j=0}^i \sum_{s=0}^{n-i} (-1)^{n+i+j+s} \binom{n}{i} \binom{i}{j} \binom{n-i}{s} \psi^{n-i-s}(\kappa) \log^{j+s}(\kappa) \Gamma^{(i-j)}(v) \Big|_{v=\kappa}, \text{ para } i \geq j, s.
 \end{aligned}$$

Ahora se probará que todos los términos en donde aparece $\log^{j+s}(\kappa)$ con $j + s \geq 1$ son cero.

Cuando $j + s = 1$, se tienen dos situaciones

- Si $j = 0$ y $s = 1 \Rightarrow i \leq n - 1$, entonces

$$\sum_{i=0}^{n-1} (-1)^{n+i+1} \binom{n}{i} \binom{i}{0} \binom{n-i}{1} \psi^{n-i-1}(\kappa) \log(\kappa) \Gamma^{(i)}(v) \Big|_{v=\kappa}.$$

- Si $s = 0$ y $j = 1 \Rightarrow i \geq 1$, entonces

$$\sum_{i=1}^n (-1)^{n+i+1} \binom{n}{i} \binom{i}{1} \binom{n-i}{0} \psi^{n-i}(\kappa) \log(\kappa) \Gamma^{(i-1)}(v) \Big|_{v=\kappa}.$$

Realizando el cambio $i - 1 = t$

$$\sum_{t=0}^{n-1} (-1)^{n+t} \binom{n}{t+1} \binom{t+1}{1} \binom{n-t}{0} \psi^{n-t-1}(\kappa) \log(\kappa) \Gamma^{(t)}(v) \Big|_{v=\kappa}.$$

Sumando ambos casos se obtiene cero, porque son de signo cambiado, $\binom{n}{i} \binom{n-i}{1} = \binom{n}{i+1} \binom{i+1}{1}$ y $\binom{i}{0} = \binom{n-i}{0}$.

Cuando $j + s = 2$, se tienen tres situaciones

- Si $j = 0$ y $s = 2 \Rightarrow i \leq n - 2$, entonces

$$\sum_{i=0}^{n-2} (-1)^{n+i} \binom{n}{i} \binom{i}{0} \binom{n-i}{2} \psi^{n-i-2}(\kappa) \log^2(\kappa) \Gamma^{(i)}(v) \Big|_{v=\kappa}.$$

- Si $s = 0$ y $j = 2 \Rightarrow i \geq 2$, entonces

$$\sum_{i=2}^n (-1)^{n+i} \binom{n}{i} \binom{i}{2} \binom{n-i}{0} \psi^{n-i}(\kappa) \log^2(\kappa) \Gamma^{(i-2)}(v) \Big|_{v=\kappa}.$$

Realizando el cambio $i - 2 = t$

$$\sum_{t=0}^{n-2} (-1)^{n+t} \binom{n}{t+2} \binom{t+2}{2} \binom{n-t-2}{0} \psi^{n-t-2}(\kappa) \log^2(\kappa) \Gamma^{(t)}(v) \Big|_{v=\kappa}.$$

- Si $j = 1 \Rightarrow i \geq 1$ y $s = 1 \Rightarrow i \leq n - 1$, entonces

$$\sum_{i=1}^{n-1} (-1)^{n+i} \binom{n}{i} \binom{i}{1} \binom{n-i}{1} \psi^{n-i-1}(\kappa) \log^2(\kappa) \Gamma^{(i-1)}(v) \Big|_{v=\kappa}.$$

Realizando el cambio $i - 1 = t$

$$\sum_{t=0}^{n-2} (-1)^{n+t+1} \binom{n}{t+1} \binom{t+1}{1} \binom{n-t-1}{1} \psi^{n-t-2}(\kappa) \log^2(\kappa) \Gamma^{(t)}(v) \Big|_{v=\kappa}.$$

Sumando los tres casos se obtiene cero, porque $\binom{n}{i} \binom{i}{0} \binom{n-i}{2} = \binom{n}{i+2} \binom{i+2}{2} \binom{n-i-2}{0}$ ambos términos tienen el mismo signo y $2 \binom{n}{i} \binom{i}{0} \binom{n-i}{2} = \binom{n}{i+1} \binom{i+1}{1} \binom{n-i-1}{1}$ con los términos de signo cambiado.

Finalmente se analiza el caso general, cuando $j + s = m \leq n$. Partiendo del desarrollo de las dos situaciones anteriores se puede observar que después de establecer los $m + 1$ casos $j = 0, s = m; j = 1, s = m - 1; \dots; j = m, s = 0$ y realizado los cambios de variables en los índices de sumatorias se llega a tener que verificar, si $\sum_{j=0}^m (-1)^j \binom{n}{i+j} \binom{i+j}{j} \binom{n-i-j}{m-j} = 0$ para esto se desarrollan las combinatorias

$$\begin{aligned} \sum_{j=0}^m (-1)^j \binom{n}{i+j} \binom{i+j}{j} \binom{n-i-j}{m-j} &= \sum_{j=0}^m (-1)^j \frac{n!}{(i+j)!(n-i-j)!} \frac{(i+j)!}{(i+j-j)!} \frac{(n-i-j)!}{(n-i-j-m+j)!} \\ &= \sum_{j=0}^m \frac{(-1)^j n!}{j! i! (m-j)! (n-m-i)!} = \frac{n!}{i! (n-m-i)! m!} \sum_{j=0}^m \frac{(-1)^j m!}{j! (m-j)!} \\ &= \frac{n!}{i! (n-m-i)! m!} (1-1)^m = 0. \end{aligned}$$

De donde se concluye que lo únicos términos que quedan en los momentos centrales, son aquellos que no contienen al factor $\log(\kappa)$. Es decir,

$$E((T - \mu)^n) = (-1)^n \frac{(\sqrt{\kappa})^n}{\Gamma(\kappa)} \sum_{i=0}^n (-1)^i \binom{n}{i} \psi^{n-i}(\kappa) \Gamma^{(i)}(v) \Big|_{v=\kappa} \quad (2.4.4)$$

Ahora se mostrará que en la expresión (2.4.4) después de desarrollar la derivada de $\Gamma^{(i)}(v)$, como se muestra en las expresiones (1.4.7)-(1.4.13), no aparecerán los términos que contengan a $\psi^j(\kappa)$ para $j \geq 1$.

- Para los términos $\psi^n(\kappa)$. Estos términos aparecen al desarrollar $\Gamma^{(i)}(v)$, que contendrá a elegir $\psi(\kappa)$ con su mayor potencia i y coeficiente 1, luego al sumar todos estos términos se tiene la suma de sus coeficientes

$$\sum_{i=0}^n (-1)^i \binom{n}{i} = (1 - 1)^n = 0.$$

- Para los términos $\psi^{n-2}(\kappa)$, no existen términos con potencia $n - 1$. Similarmente estos términos aparecen al desarrollar $\Gamma^{(i)}(v)$, eligiendo $\psi(\kappa)$ con potencia $i - 2$, de donde $i \geq 2$ su coeficiente es $\binom{i}{2}$, luego al sumar todos estos términos se tiene la suma de sus coeficientes

$$\sum_{i=2}^n (-1)^i \binom{n}{i} \binom{i}{2} = \sum_{i=0}^{n-2} (-1)^i \binom{n}{i+2} \binom{i+2}{2} = \binom{n}{2} \sum_{i=0}^{n-2} (-1)^i \binom{n-2}{i} = \binom{n}{2} (1 - 1)^{n-2} = 0.$$

- En general para los términos $\psi^{n-m}(\kappa)$, con $m \leq n$. Similarmente estos términos aparecen al desarrollar $\Gamma^{(i)}(v)$ con $\psi(\kappa)$ a la potencia $i - m$, de donde $i \geq m$ que tendrá coeficiente $\binom{i}{m}$, luego al sumar todos estos términos

$$\begin{aligned} \sum_{i=m}^n (-1)^i \binom{n}{i} \binom{i}{m} &= (-1)^m \sum_{i=0}^{n-m} (-1)^i \binom{n}{i+m} \binom{i+m}{m} = (-1)^m \binom{n}{m} \sum_{i=0}^{n-m} (-1)^i \binom{n-m}{i} \\ &= (-1)^m \binom{n}{m} (1 - 1)^{n-m} = 0. \end{aligned}$$

El teorema queda demostrado. \square

Con estos resultados se puede formular un teorema para la distribución del estimador $\hat{\kappa}_2$, del parámetro para una distribución $LGG(0, 1, \kappa)$.

Teorema 2.15. Sean X_1, \dots, X_n variables aleatorias iid con distribución $LGG(0, 1, \kappa)$, tales que $E(X_1^6) < \infty$, además $\hat{\kappa}_2$ el estimador de κ dado en (2.4.1), en forma general $\hat{\kappa}_2 = c_1 \frac{m_2^{c_2}}{(-m_3)^{c_3}}$ con c_1, c_2, c_3 constantes reales y m_2 y m_3 segundo y tercer momento central muestral, entonces $\sqrt{n}(\hat{\kappa}_2 - \kappa_2) \xrightarrow{D} N(0, \tau^2)$ con

$$\begin{aligned} \kappa_2 &= c_1 \frac{\mu_2'^{c_2}}{(-\mu_3')^{c_3}} \\ \tau^2 &= (\kappa_2)^2 \left(\frac{c_3^2 \mu_6'}{\mu_3'^2} + \frac{c_2^2 \mu_4'}{\mu_2'^2} - 2 \frac{c_2 c_3 \mu_5'}{\mu_2' \mu_3'} - (c_2 - c_3)^2 \right). \end{aligned}$$

En donde, los momentos centrales μ_j' para $j = 2, 3, 4, 5, 6$ están dados en la nota 2.7.

Demostración. Sea $\mathbf{T}_n = \begin{pmatrix} m_3 \\ m_2 \end{pmatrix}$, $\boldsymbol{\theta} = \begin{pmatrix} \mu'_3 \\ \mu'_2 \end{pmatrix}$ y el supuesto $E(X_1^6) < \infty$, luego del teorema 2.13 con $t_1 = 3$ y $t_2 = 2$, se tiene

$$\sqrt{n}(\mathbf{T}_n - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma}) \text{ con } \boldsymbol{\Sigma} = \begin{pmatrix} \mu'_6 - \mu'^2_3 & \mu'_5 - \mu'_3\mu'_2 \\ \mu'_5 - \mu'_3\mu'_2 & \mu'_4 - \mu'^2_2 \end{pmatrix}.$$

Por lo tanto, del teorema 2.12 para la función $g(\boldsymbol{\theta}) = c_1 \frac{(\mu'_2)^{c_2}}{(-\mu'_3)^{c_3}}$ y $g(\mathbf{T}_n) = c_1 \frac{(m_2)^{c_2}}{(-m_3)^{c_3}}$ se cumple

$$\sqrt{n}(g(\mathbf{T}_n) - g(\boldsymbol{\theta})) \xrightarrow{D} N(0, \tau^2) \text{ con } \tau^2 = (\dot{g}(\boldsymbol{\theta}))^t \boldsymbol{\Sigma} (\dot{g}(\boldsymbol{\theta})).$$

Ahora se calcula la derivada

$$\begin{aligned} g'_{\mu'_3} &= c_1 c_3 \frac{(\mu'_2)^{c_2}}{(-\mu'_3)^{c_3+1}} = \kappa_2 \frac{c_3}{-\mu'_3} \\ g'_{\mu'_2} &= c_1 c_2 \frac{(\mu'_2)^{c_2-1}}{(-\mu'_3)^{c_3}} = \kappa_2 \frac{c_2}{\mu'_2} \end{aligned}$$

Luego, $\dot{g}(\boldsymbol{\theta}) = \kappa_2 \begin{pmatrix} -c_3/\mu'_3 \\ c_2/\mu'_2 \end{pmatrix}$ y la forma cuadrática resulta

$$\begin{aligned} \tau^2 &= (\dot{g}(\boldsymbol{\theta}))^t \begin{pmatrix} \mu'_6 - \mu'^2_3 & \mu'_5 - \mu'_3\mu'_2 \\ \mu'_5 - \mu'_3\mu'_2 & \mu'_4 - \mu'^2_2 \end{pmatrix} \dot{g}(\boldsymbol{\theta}) \\ &= \kappa_2^2 \left[\frac{c_3^2}{\mu'^2_3} (\mu'_6 - \mu'^2_3) + \frac{c_2^2}{\mu'^2_2} (\mu'_4 - \mu'^2_2) - 2 \frac{c_2 c_3}{\mu'_2 \mu'_3} (\mu'_5 - \mu'_2 \mu'_3) \right] \\ &= \kappa_2^2 \left[c_3^2 \frac{\mu'_6}{\mu'^2_3} + c_2^2 \frac{\mu'_4}{\mu'^2_2} - 2c_2 c_3 \frac{\mu'_5}{\mu'_2 \mu'_3} - (c_2 - c_3)^2 \right]. \end{aligned}$$

El teorema queda demostrado. □

NOTA 2.7. Del teorema 2.14 y las expresiones (1.4.7)-(1.4.13), resultan los primeros 7 momentos centrales para una distribución $LGG(0, 1, \kappa)$:

$$\begin{aligned} E(T - \mu) &= 0 \\ E(T - \mu)^2 &= (\sqrt{\kappa})^2 \psi'(\kappa) \\ E(T - \mu)^3 &= (\sqrt{\kappa})^3 \psi''(\kappa) \\ E(T - \mu)^4 &= (\sqrt{\kappa})^4 [\psi'''(\kappa) + 3(\psi'(\kappa))^2] \\ E(T - \mu)^5 &= (\sqrt{\kappa})^5 [\psi^{(4)}(\kappa) + 10\psi'(\kappa)\psi''(\kappa)] \\ E(T - \mu)^6 &= (\sqrt{\kappa})^6 [\psi^{(5)}(\kappa) + 15\psi'(\kappa)\psi'''(\kappa) + 10(\psi''(\kappa))^2 + 15(\psi'(\kappa))^3] \\ E(T - \mu)^7 &= (\sqrt{\kappa})^7 [\psi^{(6)}(\kappa) + 21\psi'(\kappa)\psi^{(4)}(\kappa) + 35\psi''(\kappa)\psi'''(\kappa) + 105\psi''(\kappa)(\psi'(\kappa))^2]. \end{aligned}$$

Del teorema 2.15 se deduce el siguiente corolario.

Corolario 2.16. *Bajo las condiciones del teorema 2.15 se concluye:*

a) $\hat{\kappa}_2$ es un estimador asintóticamente insesgado de κ .

b) $\hat{\kappa}_2$ es consistente en ECM para κ .

Con la finalidad de simplificar los cálculos para la obtención de los momentos centrales de la distribución log-gamma generalizada estándar, a partir de la demostración del teorema anterior y cálculos de la derivada de la función gamma, ver (1.4.7)-(1.4.13), se puede establecer una regla para determinar los términos que aparecerán en los momentos centrales de la distribución log-gamma generalizada en su forma estándar. Para esto se establece la siguiente notación: (i, j, k) representará al término $\psi^{(i)}(\kappa)\psi^{(j)}(\kappa)\psi^{(k)}(\kappa)$, para $i, j, k \in \mathbb{N}$, por ejemplo $(1, 3, 1) \rightarrow (\psi'(\kappa))^2\psi'''(\kappa)$. Luego, la regla para determinar el k -ésimo momento central es la siguiente:

- 1).- Todos los términos tienen al factor común $(\sqrt{\kappa})^k$.
- 2).- Tienen al término $\psi^{(k-1)}(\kappa)$.
- 3).- Los términos $(k-2) \mapsto 2$ se forman con $(1, k-3), (2, k-4), \dots, (k-3, 1)$.
- 4).- Los términos $(k-3) \mapsto 3$ se forman con $(1, 1, k-5), (1, 2, k-6), \dots$
- 5).- Así, sucesivamente hasta que $(k-s) < s$.

En este caso la notación $(3) \mapsto 2$ significa que el término con derivadas de orden 3 se forma con las combinaciones de 2 factores ψ cuyos órdenes de derivadas sumen 3 y que sean mayores a cero. En este caso solo existe una combinación de dos factores $(1, 2)$. Similarmente $(5) \mapsto 3$, se obtiene con combinaciones de 3 factores cuyas derivadas de orden sumen 5, $(1, 1, 3)$ y $(1, 2, 2)$.

EJEMPLO 2.2

Se revisarán los términos del quinto al noveno momento.

1).- $E(T - \mu)^5$, en este caso $k = 5$, luego tiene los términos:

- $\psi^{(4)}(\kappa)$,
- $(3) \mapsto 2$ se forman con $(1, 2)$, luego tendrá el término $\psi'(\kappa)\psi''(\kappa)$.

2).- $E(T - \mu)^6$, en este caso $k = 6$, luego tiene los términos:

- $\psi^{(5)}(\kappa)$,
- $(4) \mapsto 2$ se forman con $(1, 3)$ y $(2, 2)$, luego tendrá los términos $\psi'(\kappa)\psi'''(\kappa)$ y $(\psi''(\kappa))^2$
- $(3) \mapsto 3$ se forman con $(1, 1, 1)$, luego tendrá el término $(\psi'(\kappa))^3$.

3).- $E(T - \mu)^7$, en este caso $k = 7$, luego tiene los términos:

- $\psi^{(6)}(\kappa)$,
 - (5) \mapsto 2 se forman con (1, 4) y (2, 3), luego tendrá los términos $\psi'(\kappa)\psi^{(4)}(\kappa)$ y $\psi''(\kappa)\psi'''(\kappa)$
 - (4) \mapsto 3 se forman con (1, 1, 2), luego tendrá el término $(\psi'(\kappa))^2\psi''(\kappa)$.
- 4).- $E(T - \mu)^8$, en este caso $k = 8$, luego tiene los términos:
- $\psi^{(7)}(\kappa)$,
 - (6) \mapsto 2 se forman con (1, 5), (2, 4) y (3, 3), luego tendrá los términos $\psi'(\kappa)\psi^{(5)}(\kappa)$, $\psi''(\kappa)\psi^{(4)}(\kappa)$ y $(\psi'''(\kappa))^2$
 - (5) \mapsto 3 se forman con (1, 1, 3) y (1, 2, 2), luego tendrá los términos $(\psi'(\kappa))^2\psi'''(\kappa)$ y $\psi'(\kappa)(\psi''(\kappa))^2$.
 - (4) \mapsto 4 se forman con (1, 1, 1, 1), luego tendrá el término $(\psi'(\kappa))^4$.

Realizando los cálculos para los coeficientes se encontró el octavo momento central de la distribución log-gamma generalizada en su forma estándar.

$$E(T - \mu)^8 = (\sqrt{\kappa})^8 \left[\psi^{(7)}(\kappa) + 28\psi'(\kappa)\psi^{(5)}(\kappa) + 56\psi''(\kappa)\psi^{(4)}(\kappa) + 35(\psi'''(\kappa))^2 + 210(\psi'(\kappa))^2\psi'''(\kappa) + 280\psi'(\kappa)(\psi''(\kappa))^2 + 105(\psi'(\kappa))^4 \right]. \quad (2.4.5)$$

- 5).- $E(T - \mu)^9$, en este caso $k = 9$, luego tiene los términos:
- $\psi^{(8)}(\kappa)$,
 - (7) \mapsto 2 se forman con (1, 6), (2, 5) y (3, 4), luego tendrá los términos $\psi'(\kappa)\psi^{(6)}(\kappa)$, $\psi''(\kappa)\psi^{(5)}(\kappa)$ y $\psi^{(3)}(\kappa)\psi^{(4)}(\kappa)$
 - (6) \mapsto 3 se forman con (1, 1, 4), (1, 2, 3) y (2, 2, 2), luego tendrá los términos $(\psi'(\kappa))^2\psi^{(4)}(\kappa)$, $\psi'(\kappa)\psi''(\kappa)\psi'''(\kappa)$ y $(\psi''(\kappa))^3$.
 - (5) \mapsto 4 se forman con (1, 1, 1, 2), luego tendrá el término $(\psi'(\kappa))^3\psi''(\kappa)$.

En las gráficas (2.6) se muestran algunas aproximaciones para la distribución asintótica del estimador de momentos para el parámetro de forma de la distribución $LGG(0, 1, \kappa)$. En general, las curvas representan lo siguiente:

- La curva en negro representa la distribución de los estimadores de κ , obtenida con 10,000 realizaciones y un tamaño de muestra $n = 10,000$,
- la curva en azul ($\hat{\kappa}_{2m}$) corresponde a la distribución normal con los valores de las realizaciones de los estimadores y
- la gráfica en rojo ($\hat{\kappa}_{2a}$) corresponde a la aproximación de la distribución normal, obtenida con los valores del teorema 2.15.

Para el caso de interés, dado en (2.4.1), se tiene $c_1 = 1.0517$, $c_2 = 3.0495$ y $c_3 = 2.033$.

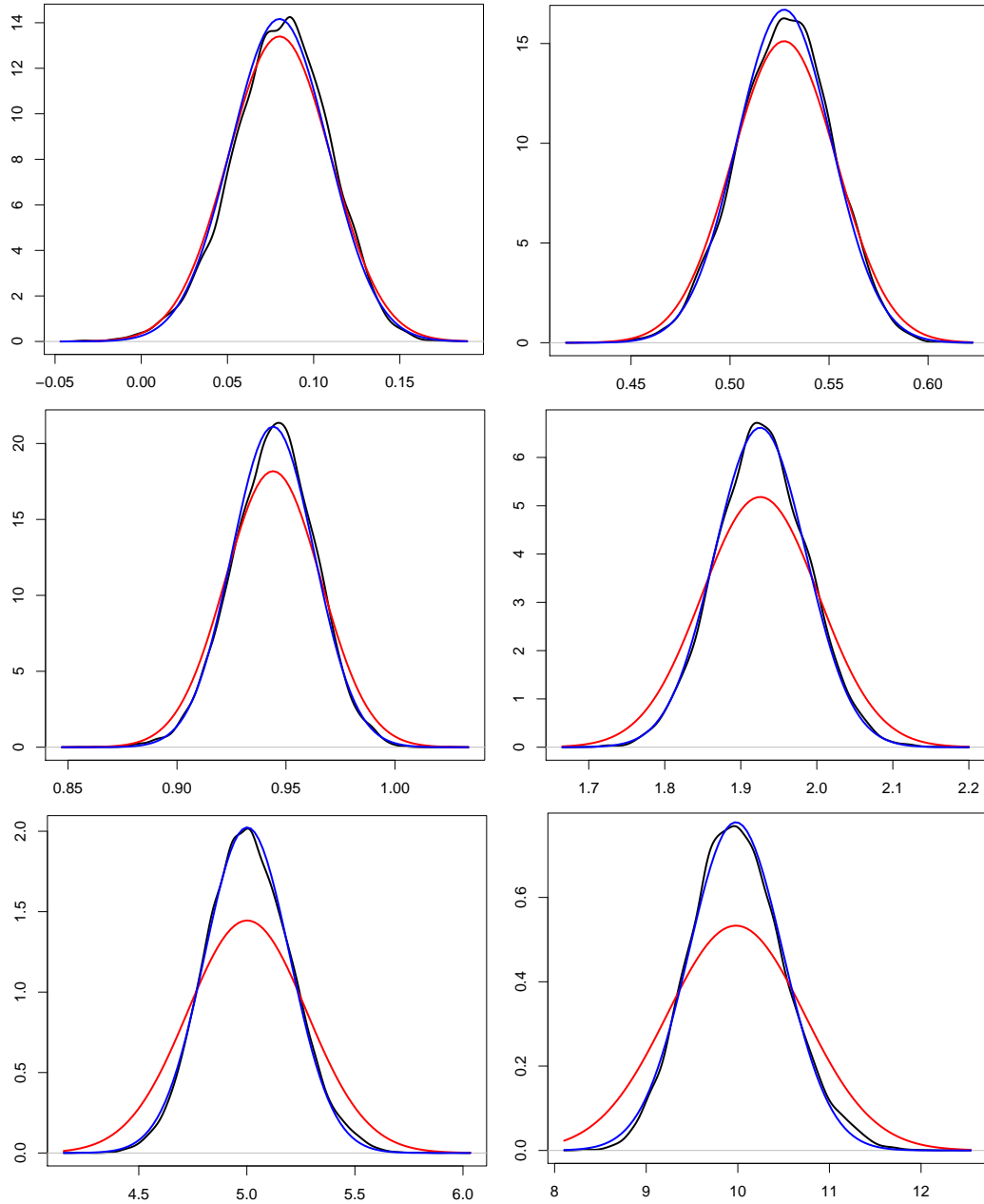


Figura 2.6: Muestran las distribuciones aproximadas del parámetro κ con valores 0.1, 0.5, 1, 2, 5 y 10, obteniendo: $\hat{\kappa}_{2a} \sim N(0.0802, 0.000886)$ y $\hat{\kappa}_{2m} \sim N(0.0810, 0.000793)$; $\hat{\kappa}_{2a} \sim N(0.5273, 0.000696)$ y $\hat{\kappa}_{2m} \sim N(0.5277, 0.000571)$; $\hat{\kappa}_{2a} \sim N(0.9441, 0.000482)$ y $\hat{\kappa}_{2m} \sim N(0.9446, 0.000407)$; $\hat{\kappa}_{2a} \sim N(1.9256, 0.005927)$ y $\hat{\kappa}_{2m} \sim N(1.9275, 0.003634)$; $\hat{\kappa}_{2a} \sim N(5.0016, 0.076203)$ y $\hat{\kappa}_{2m} \sim N(5.0095, 0.038852)$; $\hat{\kappa}_{2a} \sim N(9.9790, 0.559955)$ y $\hat{\kappa}_{2m} \sim N(10.0027, 0.262725)$, respectivamente.

Las gráficas muestran las distribuciones del estimador de momentos, nótese que en general la distribución de la estadística ajustada tiene colas más pesadas que la distribución del estimador de momentos.

2.4.1. ¿Cómo determinar la distribución asintótica de $\hat{\kappa}_2$?

La distribución asintótica de $\hat{\kappa}_2$ está dada en el teorema 2.15

$$\hat{\kappa}_2 \xrightarrow{n \rightarrow +\infty} N\left(\kappa_2, \frac{1}{n}\tau^2\right), \quad (2.4.6)$$

en donde el valor de τ está dado en el mismo teorema 2.15. Para determinar la distribución asintótica de $\hat{\kappa}_2$ se puede establecer la siguiente metodología.

- 1).- Calcular el estimador de momentos del parámetro κ o del coeficiente de asimetría γ , sea éste $\hat{\kappa}_0$.
- 2).- Con el valor calculado y la función seccionada que mejor ajusta la inversa de G , dada en la nota 2.6, determinar los valores de las constantes c_1 , c_2 , c_3 .
- 3).- Con las constantes determinadas y las expresiones del teorema 2.15 calcular la media y varianza de $\hat{\kappa}_2$.

Con esta metodología se trazaron las gráficas (2.6).

Por ejemplo, para $n = 1000$ resultó $\hat{\kappa}_0 = 0.4$, entonces la distribución del estimador $\hat{\kappa}_2 \sim N(0.40223, 0.05935)$. Para los cálculos se utilizó la función en el proyecto R dada en el anexo correspondiente al capítulo 2.

Estudio de los estimadores de κ por Simulación

3.1. Introducción

En el estudio de las distribuciones muestrales es muy importante el conocimiento sobre la distribución de la población. Como se sabe esto se puede lograr por medio de una prueba de hipótesis, para el caso de estudio se llevará a cabo en el siguiente capítulo.

En el presente capítulo se continuará revisando la parte que está atrás de una prueba de bondad de ajuste, puesto que como se sabe ésta inicia desde la búsqueda de los estimadores que se utilizarán en la prueba de ajuste. En el capítulo previo se propuso tres estimadores, sobre los que se llevó a cabo un desarrollo analítico para demostrar su existencia y unicidad, también se estudió el comportamiento asintótico del estimador de momentos y el de máxima verosimilitud obteniendo su distribución y demostrando que son consistentes en ECM y asintóticamente insesgados. Ahora con los estimadores propuestos la etapa que sigue consiste en realizar un estudio estadístico sobre su comportamiento, porque se debe buscar al estimador con base en las propiedades deseables de un estimador puntual, como son: suficiencia, insesgamiento, consistencia en probabilidad y consistencia en error cuadrado medio.

El problema en el estudio de las propiedades de un estimador, reside en que existe una gran cantidad de distribuciones de las que no se tiene un conocimiento profundo sobre su comportamiento o en ocasiones éstas son bastante complicadas. Por tales razones se crean desarrollos por métodos de simulación para dichas distribuciones, tal es el caso de la distribución que se está estudiando, la log-gamma generalizada. En el capítulo anterior se inició con el estudio sobre su distribución asintótica, ahora se analizarán sus propiedades de estimación, insesgamiento y consistencia en error cuadrado medio, pero debido a que no se conoce la distribución de cada uno de los estimadores esto se realizará por medio de la simulación.

Así, en esta parte del trabajo se lleva a cabo un estudio detallado de simulación sobre la distribución log-gamma generalizada en su forma estándar, para esto se hace referencia y en ocasiones (cuando sea un poco más complicado) se detallan los métodos numéricos requeridos

para realizar toda la simulación. El análisis inicia desde la generación de números aleatorios de la distribución en estudio, continuando con la búsqueda de cuantiles y con estas técnicas estudiadas se pasa a la parte estadística, en donde se corroboran los resultados asintóticos obtenidos en el capítulo 2. Para esto último se calcula por simulación el error cuadrado medio y varianza, para determinar si son o no insesgados los estimadores. Lo anterior se lleva a cabo tanto para muestras pequeñas como muestras grandes.

3.2. La generación de números aleatorios

La generación de números aleatorios aunque es una tarea bastante conocida, sigue teniendo dificultades en su aplicación para las distribuciones poco conocidas o de nueva creación, según el tipo de fenómeno en estudio. Es conocido que los paquetes estadísticos tienen varias funciones para generar números aleatorios de las distribuciones más populares. Un paquete bastante completo en este sentido es el Proyecto R, puesto que tiene funciones para un gran número de distribuciones.

La cantidad de distribuciones es tan amplia como los mismos fenómenos que ocurren en nuestro alrededor, por tales motivos es importante tener en mente algunos métodos o técnicas para generar la distribución que se desea estudiar. Por ejemplo, algunos de los métodos más utilizados son:

- 1).- Método de Aceptación-Rechazo.
- 2).- Método de Transformación inversa.
- 3).- Método de Composición.
- 4).- Método de transformaciones.
- 5).- Método de congruencias multiplicativo.
- 6).- Método de Fibonacci.
- 7).- Método de Generación de vectores aleatorios.
- 8).- Método Polar para la generación de variables aleatorias normales.

Así, cuando se tiene una situación en la que no se conoce cómo generar la distribución aleatoria, es posible intentar la generación de la distribución a estudiar con alguno de estos métodos o combinación de ellos. Una consulta de estos métodos se puede llevar a cabo en el libro de Sheldon [35].

3.2.1. Método de aceptación-rechazo, generación de la distribución $LGG(0, 1, \kappa)$

Para la generación de números aleatorios de la distribución log-gamma generalizada estándar, después de estudiar su forma y propiedades en el capítulo 1, ver figuras 1.2, 1.3 y 1.4 y el método de aceptación-rechazo junto con el soporte de esta distribución, se puede pensar en la distribución normal como función envolvente.

Al estudiar la familia log-gamma generalizada se puede observar que su cola izquierda es sumamente pesada, por tal razón la función envolvente en la que se pensó al inicio (normal estándar) puede traer dificultades como función envolvente. Luego, del conocimiento de la media y varianza de la log-gamma generalizada se puede pensar, como función envolvente, en la distribución normal con la media y varianza de la log-gamma generalizada.

Así, para la generación de números aleatorios de la distribución log-gamma generalizada

se utilizará la distribución normal, con media $\mu = \sqrt{\kappa}(\psi(\kappa) - \log(\kappa))$ y varianza $\sigma^2 = \kappa\psi'(\kappa)$, como función envolvente. Por lo tanto, para poder aplicar el método se tiene que maximizar

$$\max_x \frac{f_X(x; 0, 1, \kappa)}{\phi(x; \mu, \sigma^2)} = \max_x \frac{\frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \exp\left(x\sqrt{\kappa} - \kappa \exp\left(\frac{x}{\sqrt{\kappa}}\right)\right)}{\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}. \quad (3.2.1)$$

Reduciendo la expresión (3.2.1), resulta que se tiene que maximizar con respecto a x la función

$$\max_x \left\{ \exp\left(x\sqrt{\kappa} + \frac{(x-\mu)^2}{2\sigma^2} - \kappa \exp\left(\frac{x}{\sqrt{\kappa}}\right)\right) \right\}. \quad (3.2.2)$$

Al tratar de obtener el máximo, se observa que la función depende del valor κ , luego es propicio analizar cómo influye este parámetro en la maximización.

Se observó que las colas de la distribución log-gamma generalizada eran muy pesadas y la influencia puede ser grande, esto se aprecia al analizar el exponente de la función exponencial.

Cuando $x < 0$ el crecimiento lo dicta $(x - \mu)^2$, es decir, la función será creciente tipo exponencial x^2 , en el caso de $x > 0$ el crecimiento lo dicta $e^{x/\sqrt{\kappa}}$, pero tiene su coeficiente negativo y está como exponente de la función exponencial. Por consiguiente, la función en el sub-eje positivo se aproxima a cero rápidamente.

Gráficamente la función (3.2.2) se tiene una idea más precisa de cómo influye el valor de κ en la búsqueda un valor que acote el cociente entre la función de densidad de la log-gamma generalizada y su envolvente, en donde se puede apreciar que no es posible encontrar la constante del método de aceptación y rechazo. En la figura 3.1 se puede apreciar el análisis planteado, por lo tanto, no es posible aplicar el método de aceptación-rechazo para esta familia de distribuciones.

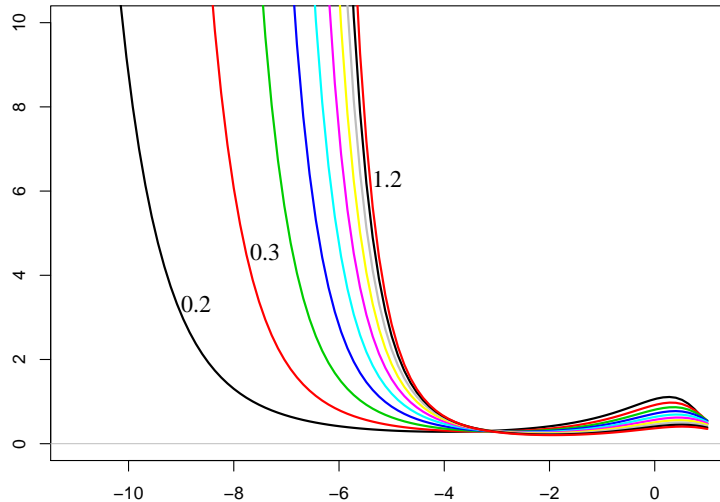


Figura 3.1: Gráficas de la función que se debe maximizar para que la función envolvente sea una $N(\mu, \sigma^2)$, para valores de $\kappa = 0.2, 0.3, \dots, 1.2$, no tienen un máximo absoluto para generar números aleatorios de la distribución $LGG(0, 1, \kappa)$.

3.2.2. Método de transformaciones, generación de la distribución $LGG(0, 1, \kappa)$

Las dificultades que se tienen en la aplicación del método de aceptación y rechazo, consisten en que la cola izquierda es muy pesada, sobre todo para valores pequeños del parámetro de forma. Luego, se recurre a un método alternativo por medio transformaciones.

En el desarrollo del capítulo 1, se revisó la procedencia por medio de transformaciones de la distribución $\Gamma(\kappa, 1)$ a la distribución $LGG(0, 1, \kappa)$. La idea ahora consiste en aplicar dichas transformaciones a la inversa, realizando el seguimiento de la sección 1.3 se tendrá lo siguiente.

Sea Z una variable aleatoria de la distribución $LGG(0, 1, \kappa)$, realizando la transformación $Y = \kappa \exp(Z/\sqrt{\kappa})$ se obtiene que $Y \in \Gamma(\kappa, 1)$. Inversamente, se puede regresar a la distribución $LGG(0, 1, \kappa)$, partiendo de Y , por medio de $Z = \sqrt{\kappa} \log(Y/\kappa)$. Es decir, para generar números aleatorios con distribución $LGG(0, 1, \kappa)$, se pueden generar números aleatorios con distribución $\Gamma(\kappa, 1)$, posteriormente con $Z = \sqrt{\kappa} \log(Y/\kappa)$ se obtendrán los números aleatorios deseados.

Una forma de comparar gráficamente si los números generados provienen de una distribución dada, se traza la gráfica de la distribución generada y compara con la función de densidad correspondiente. En el caso de estudio se tiene la figura 3.2.

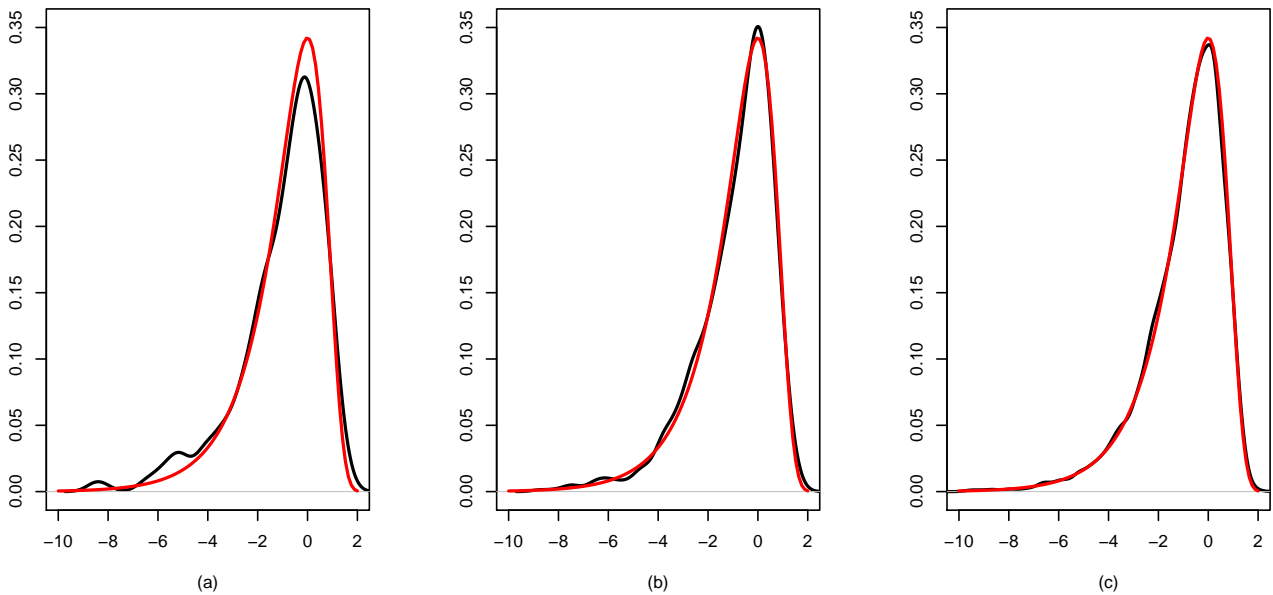


Figura 3.2: Generación de números aleatorios $LGG(0, 1, \kappa)$ por el método de la transformación, $\kappa = 0.5$: (a) $n = 500$, (b) $n = 1000$ y (c) $n = 10,000$ repeticiones. La línea negra es el resultado de la generación de números aleatorios, la otra línea (roja) la función de densidad $LGG(0, 1, 0.5)$.

Interpretación de la figura 3.2, cuando se generan 500 números aleatorios de la distribución $LGG(0, 1, 0.5)$, se puede apreciar que se asemejan un poco a la función de distribución (ver figura 3.2a), con 1000 números aleatorios las distribuciones casi coinciden (ver figura 3.2b). Se sabe de la teoría de simulación que una cantidad apropiada para la generación de una distribución aleatorio debe ser mayor a 5,000 repeticiones, en este caso se realizaron 10,000 repeticiones y las distribuciones casi coinciden, se observa en la figura (3.2c) que ambas funciones casi coinciden una está arriba de la otra.

3.2.3. Comentarios sobre la generación de muestras aleatorias de $LGG(0, 1, \kappa)$

La estimación del parámetro κ en general resulta un problema con grandes dificultades numéricas, puesto que una buena generación depende no sólo del valor de κ , sino del tamaño de la muestra n . Por ejemplo en el cuadro 2.1 se mostraron los porcentajes teóricos de valores de z que deben ser negativos en la muestra, con esto se puede concluir que en general el promedio de estos valores debe ser negativo, además de que son sesgados a la izquierda. Pero al generar muestras de tamaños pequeños, puede ocurrir que dichos valores no se distribuyan conforme al porcentaje de la tabla 2.1, llegando incluso a ser contrario, lo que ocasionaría sesgos derechos y en consecuencia el valor de $\kappa < 0$, lo cual no puede ocurrir. Este problema disminuye conforme se tomen muestra grandes.

En la tabla 3.1 se muestran los porcentajes en 100,000 realizaciones que resultaron con promedios positivos, luego darían valores estimados de κ negativos, que indicarían resultados erróneos.

Valores de κ															
n	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	2	3	4	5	10
10	0.006	0.111	0.509	1.138	1.847	2.851	3.660	4.762	5.651	6.585	14.13	18.9	22.2	24.6	31.5
20	0	0.001	0.010	0.055	0.126	0.307	0.558	0.770	1.130	1.432	6.13	10.3	13.7	16.3	24.4
30	0	0	0	0	0.014	0.024	0.077	0.159	0.222	0.395	2.77	5.9	8.7	11.3	19.5
40	0	0	0	0	0.002	0.007	0.008	0.034	0.060	0.090	1.35	3.5	5.9	8.1	16.1
50	0	0	0	0	0	0.001	0.004	0.007	0.022	0.027	0.67	2.1	3.8	5.9	13.3
60	0	0	0	0	0	0	0.001	0.002	0.003	0.005	0.31	1.3	2.8	4.2	11.0
70	0	0	0	0	0	0	0	0	0.001	0.004	0.17	0.8	1.9	3.2	9.4
80	0	0	0	0	0	0	0	0	0	0.002	0.09	0.5	1.4	2.3	7.9
90	0	0	0	0	0	0	0	0	0	0.001	0.04	0.3	0.9	1.7	6.6
100	0	0	0	0	0	0	0	0	0	0	0.02	0.2	0.6	1.3	5.7

Cuadro 3.1: Porcentaje en 100,000 corridas con promedio de valores positivos

3.3. Métodos para calcular valores de una función inversa

El problema del cálculo de valores para una función inversa es el mismo que la solución de una ecuación $f(x) = c$, c constante. Es decir, cuando la ecuación no se puede resolver en forma analítica o es demasiado compleja, se recurre a los métodos numéricos de aproximación, como son: Bisecciones, regla falsa, punto fijo, Newton-Raphson, secantes, Steffensen, etcétera.

El método más empleado en estos casos es el de Newton-Raphson, por su sencillez y poderío, está restringido a las siguientes condiciones para que sea convergente en (a, b) :

- 1).- La función $f(x) - c$ sea continua en (a, b) .
- 2).- La derivada también sea continua y además diferente de cero en (a, b) .
- 3).- La segunda derivada sea continua y no cambie de signo en (a, b) .
- 4).- El signo de la segunda derivada, debe coincidir con el de la función evaluada en la semilla.

El método es tan poderoso que generalmente se usa sin comprobar las condiciones anteriores, en caso de no converger rápidamente es que alguna de las restricciones no se cumple.

La fórmula recursiva del método en el caso de una variable es la siguiente

$$x_{n+1} = x_n - \frac{f(x_n) - c}{f'(x_n)}. \tag{3.3.1}$$

Aunque el método de Newton-Raphson es tan poderoso, en ocasiones es necesario ayudar al método en la búsqueda de las raíces, puesto que en situaciones en donde intervienen funciones exponenciales los valores de la función suelen ser muy extremos, o muy grandes o muy pequeños, dependiendo del signo del exponente. En la distribución que se está revisando en el trabajo se tienen ambas situaciones, puesto que contienen al exponente $x/\sqrt{\kappa}$, en donde x puede tomar valores en todo \mathbb{R} , luego cuando κ es pequeño los valores del cociente $x/\sqrt{\kappa}$ pueden cambiar abruptamente de positivos a negativos, en el primer caso la exponencial crecería mucho muy rápido y en el otro se aproxima también mucho muy rápido a cero, situaciones que dificultan cálculos con la distribución $L(0, 1, \kappa)$.

En la búsqueda del valor de la función inversa, el método de Newton-Raphson en ocasiones resulta difícil de implementar porque las colas de las distribuciones son pesadas y en tal caso se debe aproximar en valores del argumento grandes, lo que haría que el valor de la función de densidad sea muy pequeño, de tal forma que numéricamente se sustituye por cero. En esta situación se recomienda utilizar el método más sencillo para la solución de ecuaciones, el método de bisecciones.

El método de bisecciones la única condición que pide es la continuidad de la función, el problema que representa es que la rapidez de convergencia es muy lenta.

En ocasiones otro problema que se tiene en la búsqueda de las raíces de una ecuación es que los valores pueden ser muy pequeños, en estos casos lo que se recomienda en los cálculos es utilizar transformaciones homotéticas, que amplíen el segmento de la búsqueda y utilizar el hecho de que si $f(\alpha x) = 0$ tiene como raíz a x_0 , entonces $f(x) = 0$ tiene como raíz αx_0 .

EJEMPLO 3.1

Encontrar las raíces de la ecuación $f(x) = 2e^{-80000x} - e^{-300000x} - 1 = 0$.

Solución. Cuando se resuelve la ecuación por algún método numérico o paquete matemático la raíz por default es $x = 0$, pero tiene otra raíz muy próxima a cero que nunca la encontrarán los paquetes, a no ser que se límite la búsqueda. Una forma sencilla para determinar la raíz, consiste en realizar una homotecia y analizar, por ejemplo la ecuación $2e^{-0.08y} - e^{-0.3y} - 1 = 0$ que resulta de sustituir en la ecuación original x por $y \times 10^{-6}$ cuya raíz resulta en forma aproximada $y_0 = 7.361173037\dots$, entonces la raíz de $f(x) = 2e^{-80000x} - e^{-300000x} - 1 = 0$ es $x_0 = 7.361173037 \times 10^{-6} \dots$

En la figura 3.3 se muestran 3 funciones homotéticas, $2e^{-0.08y} - e^{-0.3y} - 1$, $2e^{-0.8y} - e^{-3y} - 1$ y $2e^{-8y} - e^{-30y} - 1$, se observa que las tres tienen una raíz en 0 y la otra es $7.361173037\dots$, $0.7361173037\dots$ y $0.07361173037\dots$, respectivamente.

3.3.1. Método para calcular cuantiles de una distribución

La búsqueda de los cuantiles en el estudio de las distribuciones es un tema muy importante, pero de apariencia sencillo, puesto que el problema en sí es simple de resolver en forma simbólica, como se verá a continuación.

Sea X una variable aleatoria continua y $F(x)$ su función de distribución, supóngase que se conoce el valor de la probabilidad acumulada que es denotada por α , y se desea conocer el valor de la variable X tal que $F(x) = \alpha$. La solución es sencilla debido a que la función

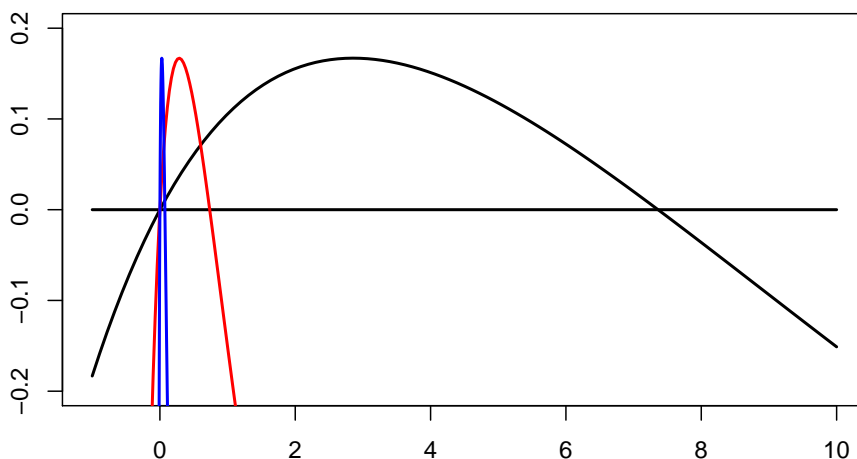


Figura 3.3: Gráficas de las funciones homóteticas para resolver la ecuación $2e^{-80000x} - e^{-300000x} - 1 = 0$.

de distribución es continua en todo \mathbb{R} , monótona no decreciente y diferenciable en todos los puntos, excepto en los puntos de discontinuidad de la función de densidad de X , luego con estas propiedades se concluye que $F(x)$ tiene inversa excepto en los segmentos en donde es constante. En el caso particular en que $F(x)$ sea monótona creciente existe su inversa y el problema de encontrar el cuantil α queda expresado de la siguiente forma

$$F(x) = \alpha \Rightarrow x = F^{-1}(\alpha). \quad (3.3.2)$$

Entonces el problema de encontrar un cuantil determinado está resuelto, pero en realidad las dificultades que se tienen en este caso se refieren a encontrar la función inversa de la función de distribución acumulada. Por ejemplo en el caso de la distribución normal estándar, es conocido que existe la función inversa de la distribución acumulada, pero ésta es la inversa de la integral $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz$, que como se sabe dicha integral no se puede resolver analíticamente.

La pregunta obvia es qué hacer en estos casos, la respuesta resulta también obvia, emplear los métodos numéricos para resolver la ecuación $F(x) = \alpha$. Entonces la solución de la ecuación anterior cuando no se tiene la expresión de la función inversa para F , se puede resolver por métodos numéricos.

En el caso particular cuando $F(x)$ es diferenciable en todo \mathbb{R} la respuesta se da en la siguiente proposición.

Proposición 3.1. *Sea X una variable aleatoria continua con función de densidad $f(x)$ y distribución $F(x)$ diferenciable en todo \mathbb{R} , entonces la ecuación $F(x) = \alpha$ se puede resolver por el método iterativo de Newton-Raphson dado en 3.3.1*

$$x_{n+1} = x_n - \frac{F(x_n) - \alpha}{f(x_n)}, \text{ con semilla } x_0 \text{ si } f(x_0) \neq 0. \quad (3.3.3)$$

Demostración. La comprobación resulta inmediata del método de Newton-Raphson y la propiedad de que $F(x)$ es diferenciable en todo \mathbb{R} . \square

Note que al ser x_n un valor específico y en caso de no conocer o poder resolver analíticamente la inversa de la función de distribución acumulada se puede calcular la integral representada por $F(x_n)$ con algún método numérico, por ejemplo el de simpson 1/3.

En la expresión **3.3.3** se divide entre la función de densidad y como se mencionó al final de la sección 3.3, en ocasiones esto puede ocasionar problemas porque las colas de la distribución pueden ser pesadas, de tal manera que para obtener el cuantil requerido es necesario considerar valores muy negativos en caso de cuantiles pequeños y distribuciones con soportes en \mathbb{R} o de valores muy grandes en caso de cuantiles próximos a 1, lo que ocasiona que la función de densidad tenga que ser evaluada en valores que resultan casi cero, por cuestiones de cómputo se sustituyen los valores por cero y los programas mandarían un error de cálculo.

En esta situación se recomienda utilizar el método de bisecciones, puesto que la función es continua y ésta es la única restricción del método (con exclusión del cambio de signo en un intervalo cerrado).

3.3.2. Método para calcular cuantiles de la distribución $LGG(0, 1, \kappa)$

Al generar la distribución aleatoria por medio de una transformación se abre otro camino para calcular los cuantiles de la distribución $LGG(0, 1, \kappa)$.

Forma 1.

En la subsección 3.2.2 al generar números aleatorios, se realizó una transformación a la distribución gamma con parámetros κ y 1. Ahora se puede utilizarla en orden inverso para regresar a los cuantiles de la distribución $LGG(0, 1, \kappa)$.

Sea $P(Z \leq z) = F_Z(z) = \alpha$ con $Z \in LGG(0, 1, \kappa)$, y se desea conocer el valor z , esto se puede lograr, calculando $y = F_Y^{-1}(\alpha)$ en donde $Y \in \Gamma(\kappa, 1)$, posteriormente se realiza el cambio de variable $z = \sqrt{\kappa} \log(y/\kappa)$.

Forma 2.

En ocasiones los valores de los parámetros son demasiado pequeños o grandes, de tal forma que pueden ocasionar problemas de desbordamientos en el proyecto R, en estos casos se recomienda realizar los cálculos con la fórmula recursiva **(3.3.3)**, sustituyendo la integración por algún método numérico, como puede ser simpson 1/3.

$$x_{n+1} = x_n - \frac{F(x_n) - \alpha}{f(x_n)} = x_n - \frac{\int_{-\infty}^{x_n} \frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \exp\left(z\sqrt{\kappa} - \kappa e^{z/\sqrt{\kappa}}\right) dz - \alpha}{\frac{\kappa^{\kappa-1/2}}{\Gamma(\kappa)} \exp\left(x_n\sqrt{\kappa} - \kappa e^{x_n/\sqrt{\kappa}}\right)}, \quad x_0. \quad (3.3.4)$$

La elección de la semilla x_0 es sencilla cuando el valor de $\kappa > 1$, se puede utilizar la semilla conforme a los valores de sus puntos de inflexión calculados en **(1.3.6)**. Es decir, para cuantiles pequeños se utiliza el punto de inflexión izquierdo $z_1 = \sqrt{\kappa} \log\left(1 + \frac{1}{2\kappa} - \frac{1}{2\kappa} \sqrt{4\kappa + 1}\right)$ y para cuantiles grandes el punto de inflexión derecho $z_2 = \sqrt{\kappa} \log\left(1 + \frac{1}{2\kappa} + \frac{1}{2\kappa} \sqrt{4\kappa + 1}\right)$. El uso de los puntos de inflexión es para que se cumplan las propiedades 3 y 4 del método de Newton-Raphson. En caso de valores pequeños de κ resulta siempre un poco complicado, porque la distribución es casi horizontal, situación que dificulta encontrar la semilla, puesto

que deberá ser un valor negativo grande en valor absoluto y como se ha comentado en las secciones previas, la evaluación de la función de densidad en estos valores es sustituida por cero mandando el programa errores numéricos por la indefinición de las operaciones.

Se ha mencionado que el problema anterior se puede resolver utilizando el método de bisecciones para resolver la ecuación $F(x) - c = 0$, en este caso no hay problemas porque la función es continua.

3.4. Estimadores para la simulación

En el capítulo previo se propusieron tres estimadores para el parámetro de forma de la distribución $LGG(0, 1, \kappa)$.

3.4.1. Estimador $\hat{\kappa}_1$ de máxima verosimilitud

En el capítulo previo se planteó el problema de los estimadores de máxima verosimilitud para los parámetros de la distribución $LGG(\xi, \sigma, \kappa)$ y $LGG(0, 1, \kappa)$ mostrando que dichos estimadores existen, en muestras de tamaños considerables en donde se conservan los porcentajes de valores negativos y positivos dados en la tabla 2.1, pero no se pueden obtener en forma analítica. Por otro lado, en la tabla 3.1 se dieron los porcentajes de promedios positivos para tamaños de muestra de 10 hasta 100 con saltos de 10, por ejemplo en el caso de un tamaño de muestra 10 y $\kappa = 10$, los cálculos se pueden interrumpir casi una tercera parte del tiempo, porque aproximará a valores del estimador negativos, pero éste no puede ser negativo, además que para valores pequeños de las muestras en cada corrida puede dar valores muy alejados del valor del parámetro. En estos casos se recomienda realizar una cantidad considerable de corridas y como estimador tomar el promedio de éstos.

Para determinar el estimador de máxima verosimilitud en el caso estándar se puede utilizar la función `optimize` del proyecto R o como se estableció en el estimador 1 en la ecuación resultante para la búsqueda del máximo relativo.

En el caso de dos o más parámetros la búsqueda se complica considerablemente porque tienen una infinidad de dirección para realizar la aproximación de las estimaciones para los parámetros. En estos casos se recomienda utilizar la función `optim` de R, de la siguiente forma.

Paso 1. Buscar una estimación de los parámetros, por medio de un recorrido por ciclos en cada variable, y acotar los valores de los parámetros que dan un valor mayor en la función de máxima verosimilitud.

Paso 2. Después de acotar los valores de los estimadores en un intervalo, cuya longitud depende del comportamiento de la función de verosimilitud. Se generan 5,000 valores uniformes, dentro de los segmentos determinados en el paso 1, para cada variable. Se evalúa la función de verosimilitud en combinaciones de estos valores y se elige la que da un valor máximo de verosimilitud, que representará una estimación para los EMV de los parámetros.

Note que este método es de bombardeos aleatorios en la región de acotación de los valores de los parámetros en que se encuentra el óptimo, de tal forma que mientras mayor sean los recorridos y menores los intervalos de búsqueda del óptimo es muy probable que los valores obtenidos sean los estimadores de máxima verosimilitud para los parámetros.

3.4.2. Estimador $\hat{\kappa}_2$ de momentos

El segundo estimador propuesto en el capítulo previo, se refiere al estimador de momentos que se calcula con el coeficiente de asimetría para la muestra. La forma de calcular a este estimador hace que tenga una importancia en el estudio de la simulación, porque prácticamente los dos primeros estimadores inician con éste.

Lo anterior se debe a que el estimador

$$\hat{\kappa}_2 = G^{-1} \left(\frac{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^3}{\left(\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 \right)^{3/2}} \right),$$

se basa únicamente en los valores muestrales.

Por lo anterior el estudio del estimador $\hat{\kappa}_2$ es de suma importancia para la obtención de los resultados en el siguiente capítulo, en donde se calcularán el tamaño y la potencia de la prueba.

En el cálculo de este estimador se pueden tener problemas cuando el coeficiente de asimetría sea negativo, porque en dichas condiciones el parámetro estimado resulta negativo, lo que no puede ocurrir. En estas condiciones hay que hacer un filtro inicial para realizar simulaciones bootstrap.

Otro problema en su cálculo consiste en decidir que fórmula utilizar para encontrarlo, puesto que se pueden emplear las técnicas revisadas en la sección 3.3. Pero también es posible utilizar la fórmula (2.4.3) encontrada al realizar el ajuste de la función inversa para determinar el estimador de momentos.

3.4.3. Estimador $\hat{\kappa}_3$ de momentos

El tercer estimador propuesto en el capítulo 2 se refiere al estimador de momentos obtenido mediante una transformación, mismo que resulta de resolver la ecuación

$$\frac{1}{n} \sum_{i=1}^n \exp\left(\frac{z_i}{\sqrt{\hat{\kappa}}}\right) = 1. \quad (3.4.1)$$

El cálculo de este estimador es el que presenta mayores dificultades en muestras pequeñas, debido a que los porcentajes de valores negativos y positivos de las variables generadas con frecuencia no se conservan, en consecuencia la suma de los términos en (3.4.1) nunca cambiará de signo, es decir no tendrá solución la ecuación.

Después de revisar las tres subsecciones sobre los estimadores del parámetro κ se puede notar que en caso del análisis asintótico que se revise, será suficiente estudiar una sola muestra de tamaño grande y ésta dará resultados apropiados de los estimadores. Pero no será lo mismo con muestras pequeñas, en donde se deberá considerar varias realizaciones para obtener buenos estimadores puntuales.

3.5. Propiedades asintóticas de los estimadores

En el caso de los 3 estimadores propuestos serán revisadas sus propiedades asintóticas deseables de un estimador, como son: el insesgamiento, la consistencia en probabilidad y en error cuadrado medio. Esto se comprobará por simulación, y se realizará en etapas la primera consistirá en una comparación de la estimación de cada uno de los tres parámetros, posteriormente se comparan sus ECM y se prueba si hay o no insesgamiento asintótico. Así, en esta sección se comprobarán las propiedades asintóticas que se obtuvieron en el capítulo 2.

3.5.1. Comparación de los estimadores

En la siguiente tabla se muestran algunos valores de κ y la estimación que realiza cada uno de los estimadores.

κ	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$		κ	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
0.1	0.10132	0.09209	0.09662		0.1	0.10069	0.09789	0.09998
0.2	0.18257	0.19438	0.20084		0.2	0.20215	0.20447	0.20081
0.4	0.41180	0.44334	0.40798		0.4	0.39943	0.39794	0.40662
0.6	0.60048	0.59856	0.62950		0.6	0.59615	0.60456	0.59225
0.8	0.80394	0.74452	0.81224		0.8	0.79487	0.79845	0.79870
1	1.01399	1.01314	0.94867		1	0.99970	0.98069	1.01304
2	2.06890	2.06656	1.91683		2	2.01405	2.05815	1.97441
3	3.03609	2.79252	3.11640		3	3.00412	2.99602	2.98931
4	3.97202	4.19916	3.64523		4	4.02002	4.14934	4.05329
5	5.18357	4.22332	5.57464		5	5.01373	5.10084	4.92933
10	9.62879	10.08474	9.98425		10	10.04422	10.25044	9.96938
15	15.01258	14.32791	14.01153		15	14.86974	15.70511	15.04783
20	19.71464	20.28032	19.43753		20	19.78095	20.10249	20.09782

Cuadro 3.2: Valores de los estimadores con tamaño de muestra 10,000 tabla izquierda (sin repeticiones) y en la tabla derecha se hacen 10 repeticiones.

En la tabla 3.2 se puede apreciar que en general con los 3 estimadores se obtienen buenas estimaciones, aunque resultan ligeramente en orden ascendente de error $\hat{\kappa}_1$, $\hat{\kappa}_3$ y $\hat{\kappa}_2$.

Note que se tomó una sola muestra de tamaño 10,000 en la tabla izquierda y en la tabla derecha las muestra también fue de tamaño 10,000, pero con 10 repeticiones, lo que mejora la estimación anterior.

3.5.2. Comparación de los ECM

Para determinar los errores cuadrados medios de cada uno de los estimadores se utiliza el programa en el proyecto R, que se puede consultar en el anexo, en la parte correspondiente al capítulo 3. Para los cálculos se tomaron valores más representativos del comportamiento del parámetro κ .

En la tabla 3.3 se pueden apreciar los ECM de los 3 estimadores en donde los valores de ECM en forma ascendente para los estimadores son $\hat{\kappa}_1$, $\hat{\kappa}_3$ y $\hat{\kappa}_2$ (ver figura 3.4).

En la figura 3.4 se puede apreciar que los ECM son directamente proporcionales con el valor de κ , para cada uno de los estimadores. Esto se aprecia con las rectas ascendentes, estando por debajo la recta correspondiente al ECM para el estimador $\hat{\kappa}_1$ y aproximadamente con dos veces su valor para el estimador $\hat{\kappa}_3$ y finalmente el estimador que proporciona un ECM

κ	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$		κ	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
0.1	0.000102	0.000473	0.000224		2	0.00195	0.00977	0.004485
0.2	0.000199	0.000960	0.000451		3	0.00295	0.01454	0.006934
0.3	0.000293	0.001418	0.000661		4	0.00406	0.01940	0.009002
0.4	0.000404	0.001901	0.000855		5	0.00496	0.02383	0.011126
0.5	0.000484	0.002382	0.001110		6	0.00583	0.02859	0.013554
0.6	0.000604	0.002877	0.001366		7	0.00701	0.03340	0.015700
0.7	0.000719	0.003324	0.001602		8	0.00786	0.03811	0.018078
0.8	0.000826	0.003827	0.001816		9	0.00871	0.04264	0.020163
0.9	0.000898	0.004349	0.002024		10	0.00958	0.04747	0.022659
1	0.001034	0.004670	0.002229		20	0.01981	0.09502	0.044932

Cuadro 3.3: Valores de los ECM para cada estimador con un tamaño de muestra $n = 10,000$ y $m = 5,000$ repeticiones; error máximo en las aproximaciones de 0.00001 (ver gráfica 3.4).

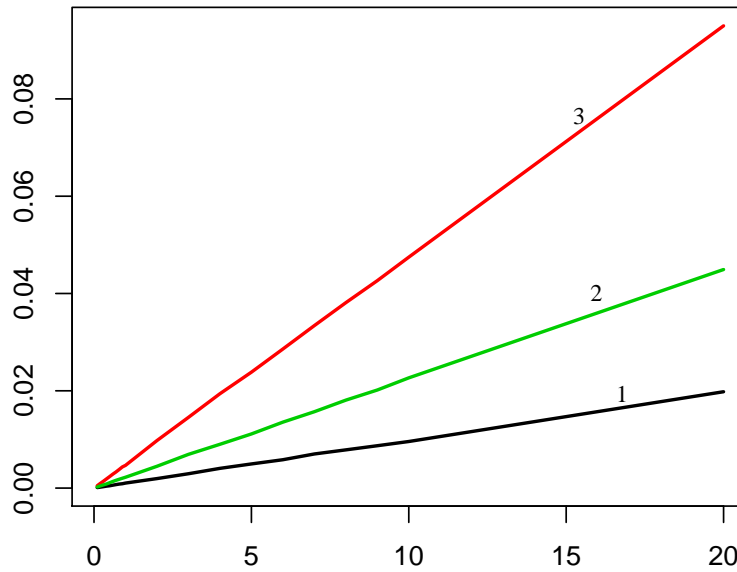


Figura 3.4: Valores de los ECM para cada estimador con un tamaño de muestra $n = 10,000$; $m = 5,000$ repeticiones; error máximo de 0.00001

más grande de los tres es $\hat{\kappa}_2$, aunque las diferencias en valor relativo entre cada uno de los estimadores son pequeñas.

Para este análisis (tabla 3.3 y figura 3.4) se realizaron 5,000 repeticiones con un tamaño muestra $n = 10,000$.

3.5.3. ECM y varianza de los estimadores

En esta subsección se analizan los ECM para cada estimador y se comparan con su varianza, obteniendo los resultados de las tablas 3.4, 3.5 y 3.6, en estas tablas se puede apreciar que la varianza coincide con ECM, luego los estimadores serán asintóticamente insesgados.

Para la ejecución en cada estimador se realizaron 5,000 repeticiones con tamaños de muestra que van de 1,000 a 15,000, variando de 1000 en 1000, para los valores de κ mostrados arriba.

VALORES de $\hat{\kappa}_1$

n	0.1		0.5		1		5		10	
	VAR	ECM	VAR	ECM	VAR	ECM	VAR	ECM	VAR	ECM
1	0.00105	0.00106	0.00509	0.00512	0.01102	0.01110	0.05322	0.05374	0.10200	0.10270
2	0.00050	0.00051	0.00246	0.00246	0.00508	0.00510	0.02497	0.02510	0.05148	0.05166
3	0.00033	0.00033	0.00169	0.00169	0.00339	0.00340	0.01656	0.01657	0.03213	0.03224
4	0.00025	0.00025	0.00127	0.00127	0.00251	0.00252	0.01207	0.01209	0.02567	0.02575
5	0.00020	0.00020	0.00103	0.00103	0.00191	0.00192	0.00984	0.00984	0.01993	0.01995
6	0.00017	0.00017	0.00085	0.00085	0.00159	0.00159	0.00835	0.00838	0.01651	0.01652
7	0.00014	0.00014	0.00072	0.00072	0.00145	0.00146	0.00691	0.00692	0.01424	0.01425
8	0.00013	0.00013	0.00063	0.00063	0.00120	0.00121	0.00617	0.00617	0.01283	0.01285
9	0.00011	0.00011	0.00055	0.00055	0.00109	0.00109	0.00549	0.00548	0.01093	0.01093
10	0.00010	0.00010	0.00051	0.00051	0.00098	0.00098	0.00499	0.00499	0.01015	0.01015
11	0.00009	0.00009	0.00046	0.00046	0.00089	0.00089	0.00443	0.00443	0.00922	0.00922
12	0.00009	0.00009	0.00040	0.00040	0.00080	0.00080	0.00409	0.00409	0.00838	0.00839
13	0.00008	0.00008	0.00038	0.00038	0.00077	0.00077	0.00370	0.00370	0.00779	0.00779
14	0.00007	0.00007	0.00036	0.00036	0.00072	0.00072	0.00352	0.00352	0.00715	0.00715
15	0.00007	0.00007	0.00032	0.00032	0.00067	0.00067	0.00336	0.00336	0.00671	0.00670

Cuadro 3.4: Valores de la varianza y ECM para el estimador $\hat{\kappa}_1$ con un tamaño de muestra n en miles; $m = 5,000$ repeticiones; error máximo de 0.00001 (ver gráfica 3.7).

VALORES de $\hat{\kappa}_2$

n	0.1		0.5		1		5		10	
	VAR	ECM	VAR	ECM	VAR	ECM	VAR	ECM	VAR	ECM
1	0.00489	0.00551	0.02312	0.02595	0.04885	0.05475	0.23348	0.26303	0.46555	0.51961
2	0.00240	0.00256	0.01150	0.01228	0.02270	0.02412	0.12125	0.12966	0.24037	0.25701
3	0.00151	0.00158	0.00796	0.00831	0.01556	0.01637	0.07669	0.08040	0.15200	0.15980
4	0.00118	0.00121	0.00576	0.00593	0.01175	0.01219	0.05868	0.06087	0.11747	0.12101
5	0.00097	0.00100	0.00461	0.00474	0.00961	0.00988	0.04817	0.04969	0.09379	0.09646
6	0.00075	0.00077	0.00400	0.00411	0.00786	0.00805	0.03783	0.03896	0.08427	0.08614
7	0.00069	0.00071	0.00333	0.00345	0.00658	0.00670	0.03373	0.03435	0.06798	0.06986
8	0.00058	0.00060	0.00295	0.00301	0.00584	0.00592	0.03019	0.03073	0.05991	0.06137
9	0.00052	0.00053	0.00271	0.00276	0.00544	0.00554	0.02574	0.02622	0.05205	0.05311
10	0.00045	0.00046	0.00229	0.00232	0.00483	0.00489	0.02256	0.02301	0.04791	0.04850
11	0.00044	0.00044	0.00216	0.00218	0.00424	0.00430	0.02193	0.02223	0.04158	0.04213
12	0.00040	0.00040	0.00200	0.00203	0.00386	0.00392	0.01983	0.02017	0.03896	0.03939
13	0.00036	0.00037	0.00178	0.00181	0.00355	0.00360	0.01838	0.01857	0.03835	0.03869
14	0.00034	0.00034	0.00171	0.00174	0.00336	0.00340	0.01701	0.01725	0.03406	0.03436
15	0.00031	0.00032	0.00157	0.00158	0.00323	0.00327	0.01584	0.01598	0.03193	0.03233

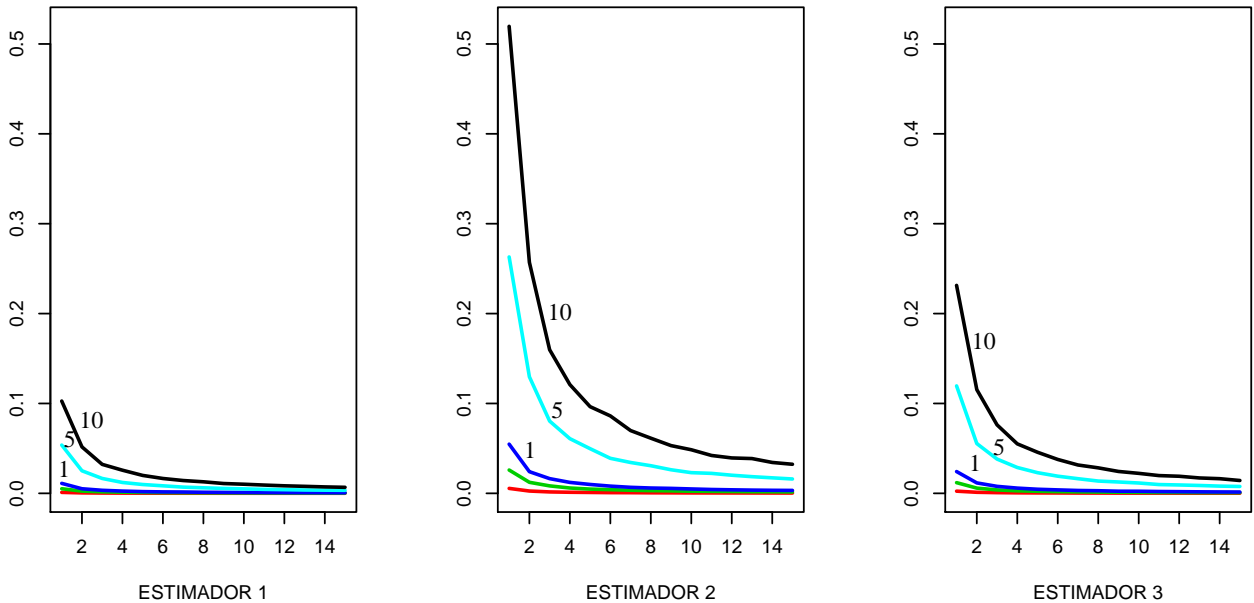
Cuadro 3.5: Valores de la varianza y ECM para el estimador κ_2 con un tamaño de muestra n en miles; $m = 5,000$ repeticiones; error máximo de 0.00001 (ver gráfica 3.7).

En la figura 3.7 se muestran los comportamientos asintóticos de cada uno de los ECM para cada estimador con los tamaños de muestra $n = 1000, 2000, \dots, 15000$. En las figuras de los ECM para cada uno de los tres estimadores, se puede notar cierta dependencia del ECM, no sólo con el tamaño de muestra, sino también con el valor del parámetro.

VALORES de $\hat{\kappa}_3$

n	0.1		0.5		1		5		10	
	VAR	ECM	VAR	ECM	VAR	ECM	VAR	ECM	VAR	ECM
1	0.00240	0.00242	0.01182	0.01190	0.02414	0.02426	0.11856	0.11954	0.23000	0.23138
2	0.00118	0.00118	0.00580	0.00582	0.01159	0.01161	0.05528	0.05537	0.11505	0.11540
3	0.00078	0.00078	0.00381	0.00382	0.00773	0.00776	0.03789	0.03814	0.07614	0.07620
4	0.00055	0.00055	0.00280	0.00280	0.00581	0.00583	0.02871	0.02875	0.05501	0.05507
5	0.00045	0.00045	0.00225	0.00226	0.00457	0.00458	0.02298	0.02300	0.04570	0.04572
6	0.00039	0.00039	0.00189	0.00189	0.00385	0.00385	0.01908	0.01910	0.03759	0.03771
7	0.00031	0.00031	0.00162	0.00162	0.00321	0.00321	0.01628	0.01629	0.03155	0.03155
8	0.00029	0.00029	0.00142	0.00142	0.00290	0.00290	0.01367	0.01367	0.02834	0.02836
9	0.00024	0.00025	0.00123	0.00123	0.00237	0.00237	0.01263	0.01263	0.02432	0.02435
10	0.00022	0.00022	0.00109	0.00109	0.00233	0.00233	0.01149	0.01150	0.02224	0.02225
11	0.00021	0.00021	0.00099	0.00099	0.00203	0.00203	0.00977	0.00977	0.01976	0.01976
12	0.00019	0.00019	0.00097	0.00097	0.00186	0.00186	0.00931	0.00932	0.01901	0.01901
13	0.00017	0.00017	0.00086	0.00086	0.00167	0.00167	0.00877	0.00877	0.01709	0.01711
14	0.00016	0.00016	0.00080	0.00080	0.00156	0.00156	0.00805	0.00805	0.01630	0.01631
15	0.00015	0.00015	0.00077	0.00077	0.00148	0.00148	0.00766	0.00767	0.01431	0.01430

Cuadro 3.6: Valores de las varianzas y ECM para los estimadores κ_3 con un tamaño de muestra n en miles; $m = 5,000$ repeticiones; error máximo de 0.00001 (ver gráfica 3.7).



Cuadro 3.7: Valores de los ECM para el estimador $\hat{\kappa}_1$, $\hat{\kappa}_2$ y $\hat{\kappa}_3$ en los 5 valores de $\kappa = 0.1, 0.5, 1, 5, 10$, se tomaron con un tamaño de muestra n dada en miles y $m = 5,000$ repeticiones

Para comprobar lo anterior sobre la proporcionalidad, se busca en cada tamaño de muestra el coeficiente de proporcionalidad entre el tamaño del ECM y n, κ , encontrando los valores que se muestran en la tabla 3.8 y como se puede observar son muy similares. Es decir, se comprobó de esta forma que los ECM si dependen en forma proporcional, tanto del tamaño de la muestra n como del parámetro κ .

$n \backslash \kappa$	Estimador $\hat{\kappa}_1$					Estimador $\hat{\kappa}_2$					Estimador $\hat{\kappa}_3$				
	0.1	0.5	1	5	10	0.1	0.5	1	5	10	0.1	0.5	1	5	10
1	10.6	10.2	11.1	10.7	10.3	55.4	54.8	53.6	51.9	54.0	24.2	23.8	24.3	23.9	23.1
2	10.1	9.8	10.2	10.0	10.3	51.5	51.0	50.0	48.9	50.6	23.6	23.3	23.2	22.1	23.1
3	10.0	10.2	10.2	9.9	9.7	49.6	48.5	48.4	48.8	48.8	23.4	22.9	23.3	22.9	22.9
4	10.0	10.1	10.1	9.7	10.3	48.2	47.9	48.3	48.3	49.8	21.9	22.4	23.3	23.0	22.0
5	10.2	10.3	9.6	9.8	10.0	47.5	47.8	48.2	47.1	50.0	22.3	22.6	22.9	23.0	22.9
6	10.0	10.2	9.5	10.1	9.9	47.6	46.7	46.7	47.7	46.0	23.4	22.7	23.1	22.9	22.6
7	9.7	10.1	10.2	9.7	10.0	48.3	49.3	49.2	47.5	47.5	21.5	22.7	22.5	22.8	22.1
8	10.1	10.2	9.6	9.9	10.3	47.1	47.2	48.0	47.0	48.7	23.6	22.8	23.2	21.9	22.7
9	10.1	9.9	9.8	9.9	9.8	47.2	47.0	46.8	47.9	49.0	22.1	22.2	21.4	22.7	21.9
10	9.7	10.2	9.8	10.0	10.2	46.7	47.8	46.9	47.9	47.9	22.5	21.9	23.3	23.0	22.2
11	10.0	10.2	9.8	9.7	10.1	48.6	48.5	49.3	48.0	49.3	22.7	21.7	22.4	21.5	21.7
12	10.2	9.7	9.6	9.8	10.1	49.0	47.7	47.3	47.2	47.0	22.4	23.3	22.4	22.4	22.8
13	10.1	9.8	9.9	9.6	10.1	47.0	48.4	47.4	47.5	49.5	22.1	22.2	21.7	22.8	22.2
14	9.9	10.0	10.0	9.9	10.0	47.3	49.8	48.9	47.5	45.8	22.0	22.4	21.8	22.6	22.8
15	9.9	9.7	10.1	10.1	10.1	48.7	47.2	46.8	48.2	50.7	22.9	23.2	22.2	23.0	21.5

Cuadro 3.8: Valores de proporcionalidad entre ECM, tamaño de la muestra n en miles y el valor del parámetro κ .

Después de calcular el promedio de estos valores se obtienen las proporciones asintóticas de decrecimiento del ECM para cada uno de los tres estimadores:

- $ECM(\hat{\kappa}_1) \asymp \frac{10\kappa}{n}$.
- $ECM(\hat{\kappa}_2) \asymp \frac{49\kappa}{n}$.
- $ECM(\hat{\kappa}_3) \asymp \frac{23\kappa}{n}$.

Luego, se concluye que asintóticamente es preferible en ECM el estimador $\hat{\kappa}_1$ sobre los otros dos, esto se corrobora de diferentes formas, con la tabla 3.8, las figura 3.4 y 3.7.

3.6. Propiedades de los estimadores en tamaños de muestra pequeños

En la sección anterior se revisaron los comportamientos asintóticos de los tres estimadores, ahora se revisarán los comportamientos en muestras pequeñas. En los casos de muestras pequeñas los cálculos se dificultan considerablemente, puesto que el problema de la conservación del porcentaje de números negativos para la distribución $LGG(0, 1, \kappa)$ mostrado en la tabla 2.1 (que es importante para la existencia de los EMV), constantemente puede no cumplirse. Para esto se trabaja en los cálculos con más restricciones que en el caso asintótico. Además las estimaciones requieren de un número determinado de corridas a diferencia del caso asintótico en donde únicamente se realizó una corrida con tamaño de muestra grande (dados en miles), aquí se realizaron 10,000 corridas para cada tamaño de muestra.

3.6.1. Comparación de los estimadores

En la siguiente tabla se muestran algunos valores de κ y la estimación que realiza para cada uno de los estimadores.

$\kappa \backslash n$	30	40	50	75	100	200	30	40	50	75	100	200
0.1	0.132	0.123	0.119	0.112	0.109	0.104	0.117	0.114	0.113	0.109	0.108	0.105
0.2	0.266	0.246	0.240	0.223	0.217	0.208	0.234	0.225	0.229	0.223	0.216	0.212
0.4	0.526	0.494	0.475	0.445	0.433	0.418	0.470	0.446	0.455	0.439	0.432	0.422
0.6	0.804	0.740	0.716	0.669	0.648	0.625	0.716	0.682	0.688	0.663	0.654	0.637
0.8	1.067	0.981	0.949	0.887	0.865	0.837	0.955	0.900	0.912	0.875	0.871	0.851
1	1.331	1.229	1.185	1.110	1.088	1.042	1.179	1.125	1.142	1.098	1.085	1.052
2	2.665	2.473	2.377	2.228	2.166	2.093	2.397	2.223	2.262	2.188	2.155	2.107
3	3.984	3.682	3.560	3.336	3.248	3.129	3.516	3.434	3.343	3.287	3.264	3.189
4	5.303	4.933	4.725	4.430	4.337	4.189	4.710	4.474	4.567	4.367	4.304	4.229
5	6.637	6.153	5.938	5.543	5.394	5.227	5.913	5.695	5.691	5.485	5.447	5.272
10	13.308	12.356	11.785	11.192	10.824	10.453	11.750	11.366	11.176	10.965	10.829	10.564
15	19.956	18.540	17.849	16.705	16.123	15.592	17.863	16.864	17.029	16.279	16.345	15.783

Cuadro 3.9: Valores de estimación para cada valor κ y tamaño de muestra n proporcionado para los estimadores $\hat{\kappa}_1$ (tabla izquierda) y $\hat{\kappa}_3$ (tabla derecha). En cada valor se realizaron 10,000 repeticiones y se tomó el promedio para la estimación.

En las tablas 3.9 y 3.10 puede notarse que se siguen cumpliendo los tamaños de los errores obtenidos en el caso asintótico. Como era de esperarse las estimaciones dependen, tanto del tamaño de la muestra, como del valor κ y el error en el estimador $\hat{\kappa}_2$ sigue siendo grande comparado con los otros dos estimadores. Por ejemplo los errores del estimador $\hat{\kappa}_2$ son del mismo orden que en los otros dos estimadores hasta que el tamaño de muestra es 5 veces más grande, observe que para $n = 200$ y $\kappa = 0.1$ se tienen las estimaciones $\hat{\kappa}_1 = 0.104$ y $\hat{\kappa}_3 = 0.105$, mientras que la estimación correspondiente para $\hat{\kappa}_2$ la muestra debe ser de tamaño 5 veces mayor ($n = 1000$) $\hat{\kappa}_2 = 0.108$.

$\kappa \backslash n$	75	100	150	200	250	300	500	1000
0.1	3.79	0.22	0.159	0.140	0.132	0.125	0.115	0.108
0.2	0.91	0.56	0.318	0.282	0.263	0.248	0.231	0.215
0.4	1.96	0.85	0.634	0.563	0.523	0.503	0.463	0.430
0.6	14.1	1.3	0.960	0.838	0.797	0.756	0.696	0.646
0.8	12.7	1.7	1.261	1.117	1.047	1.006	0.928	0.863
1	3.70	2.27	1.798	1.396	1.320	1.261	1.150	1.079
2	14.34	7.01	3.184	2.796	2.648	2.544	2.299	2.151
3	71.72	6.36	4.670	4.201	3.979	3.750	3.451	3.234
4	37.0	63.1	6.407	5.605	5.278	5.028	4.626	4.301
5	20.6	14.2	10.19	6.97	6.59	6.33	5.78	5.37
10	135.9	123.6	15.64	14.02	13.15	12.58	11.49	10.78
15	219.1	40.0	56.9	21.0	19.6	18.9	17.3	16.16

Cuadro 3.10: Valores de estimación para cada valor κ y tamaño de muestra n proporcionado para el estimador $\hat{\kappa}_2$. En cada valor se realizaron 10,000 repeticiones y se tomó el promedio para la estimación.

3.6.2. ECM y varianza de los estimadores

En el caso asintótico se encontró que los tres estimadores propuestos eran insesgados y consistentes en error cuadrado medio. Ahora se estudiará el comportamiento sobre la varianza y el error cuadrado medio cuando el tamaño de muestra es pequeño.

En el caso de muestras pequeñas el estudio sobre los estimadores resulta bastante complicado, sin importar que se tienen los programas, la parte numérica es complicada por las condiciones de la distribución. Cuando las muestras son pequeñas ocurre con frecuencia que

los porcentajes de valores positivos y negativos, dados en la tabla 2.1, pueden no cumplirse, llegando a ser en ocasiones o todos los valores muestrales negativos o todos positivos. En estas situaciones de entrada el sesgo de la distribución se altera influyendo grandemente en el valor encontrado para estimar al parámetro.

En caso de haber más valores negativos que las cantidades teóricas esto no repercute considerablemente en los estimadores $\hat{\kappa}_1$ y $\hat{\kappa}_2$. Cuando los valores positivos sobrepasan la cantidad teórica esperada, tampoco influyen considerablemente en el estimador $\hat{\kappa}_1$, pero no así en el estimador $\hat{\kappa}_2$, en cuyo caso la presencia de mayor cantidad de valores positivos es crucial para realizar una buena estimación, más aún los cálculos pueden interrumpirse debido a que en esta situación el coeficiente de asimetría podría resultar positivo, lo que ocasiona que el estimador de κ sea negativo, lo que no puede ser.

En el caso del estimador $\hat{\kappa}_3$ se tienen más dificultades, debido a que en ambos casos analizados sobre las cantidades de positivos y negativos debe ser muy balanceada, en caso contrario no existe el estimador.

Teorema 3.2. *Los teoremas 2.2 y 2.3 en muestras pequeñas se cumplen si y sólo si son conservados los porcentajes teóricos de las cantidades positivas y negativas de la tabla 2.1, junto con una distribución de ellos de acuerdo a la LGG.*

Demostración. En caso de cumplirse los porcentajes las demostraciones de los teoremas 2.2 y 2.3 se repiten.

Cuando no se cumplen dichos porcentajes puede ocurrir que la media de los valores sea positiva, lo que lleva a un coeficiente de asimetría positivo y esto sólo ocurre cuando el valor de $\kappa < 0$ lo que no puede ser, luego no existiría el estimador $\hat{\kappa}_2$.

Para el tercer estimador, la ecuación que se tiene que resolver

$$\frac{1}{n} \sum_{i=1}^n \exp\left(\frac{z_i}{\sqrt{\hat{\kappa}}}\right) = 1.$$

La expresión de la izquierda siempre es positiva, luego si existen únicamente valores de z_i positivos, la suma siempre será mayor a 1 y no tendrá raíz. En caso de ser todos los z_i negativos la suma siempre será menor a 1 y no tendrá raíz la ecuación. Luego, en estos casos no existe un estimador $\hat{\kappa}_3$. \square

En las tablas 3.11, 3.12 y 3.13 se muestra un estudio detallado de los comportamientos de los ECM y varianzas para cada uno de los tres estimadores propuestos cuando las muestras son de tamaño pequeño ($n = 30, 40, 50, 75, 100, 150, 200$) en el caso de los estimadores $\hat{\kappa}_2$ se eliminan los valores 30, 40 y 50 porque sus ECM's son muy grandes, sustituyendo por tamaños 300, 400 y 500. En cada caso se muestran tres tablas que están dadas en una misma y que están separadas por los valores, cada una de ellas se identifica porque responde a los valores de $\kappa = 0.1, 0.2, \dots, 1, 2, 3, 4, 5, 10, 15, 20$ y se identifican de la siguiente forma:

- La tabla más grande se encuentra a la izquierda, corresponde a los valores del ECM y

varianza para valores de $\kappa = 0.1, 0.2, \dots, 1, 2, 3, 4, 5, 10, 15, 20$ en ella se puede observar que los estimadores no son insesgados, pero su ECM aumenta en forma proporcional al valor correspondiente a $\kappa = 0.1$ para cualquier tamaño de muestra.

- La tabla que se encuentra arriba a la derecha, corresponde a los resultados de $ECM \times n/\kappa$ y ésta depende del estimador.
 - 1).- Estimador $\hat{\kappa}_1$. Nótese que dado un valor de n los valores $ECM \times n/\kappa$ son constante para cualquier valor de κ , de tal forma que se tiene la relación para muestras menores a 200: $ECM_1 \doteq \frac{154.7\kappa}{n^{1.13}}$.
 - 2).- Estimador $\hat{\kappa}_2$. Nótese que dado un valor de n los valores $ECM \times n/\kappa$ son constante para cualquier valor de κ , de tal forma que se tiene la relación para muestras entre 200 y 500: $ECM_2 \doteq \frac{83250\kappa}{n^{2.43}}$.
 - 3).- Estimador $\hat{\kappa}_3$. Nótese que dado un valor de n los valores $ECM \times n/\kappa$ son constante para cualquier valor de κ , pero no presentan un comportamiento monótono.

En cualquiera de los casos se observa el resultado esperado, los ECM disminuyen conforme aumenta el tamaño de la muestra.

- En la tabla de abajo a la derecha se muestran los cuadrados de los sesgos en ella se puede observar que éstos aumentan en forma proporcional al valor correspondiente a $\kappa = 0.1$ para cualquier tamaño de muestra.

NOTA 3.1. De los resultados obtenidos en el presente capítulo, resulta que el estimador con mejores propiedades es $\hat{\kappa}_1$, pero sin embargo el estimador $\hat{\kappa}_2$ representa la bondad de depender únicamente de los momentos muestrales y sus estimaciones son muy buenas comparadas con los otros dos estimadores.

El estimador $\hat{\kappa}_3$ en definitiva no es apropiado para el estudio que se realizará en el capítulo siguiente, por los problemas que presenta en sus cálculos, sin embargo hay que notar que presentó buenas propiedades asintóticas casi del orden que el estimador de máxima verosimilitud.

Por otro lado, los cálculos para las simulaciones que se realizarán con los dos primeros estimadores en el capítulo siguiente, para el tamaño de prueba y potencia de la prueba por el método de Bootstrap, se basan en su primer paso en el estimador de momentos para el coeficiente de asimetría. Por tales razones es de suma importancia llevar a cabo un estudio en la inferencia de la distribución log-gamma generalizada con el segundo estimador.

κ	TAMAÑO DE LA MUESTRA n						TAMAÑO DE LA MUESTRA n							
	30	40	50	75	100	200	κ	30	40	50	75	100	200	
0.1	Var	0.00094	0.00012	0.00009	0.00005	0.00003	0.00001	0.1	3.58	2.21	1.81	1.00	0.74	0.53
	ECM	0.01194	0.00551	0.00362	0.00133	0.00074	0.00020	0.2	3.48	2.17	1.81	1.02	0.73	0.53
0.2	Var	0.00099	0.00025	0.00019	0.00010	0.00006	0.00002	0.3	3.43	2.19	1.79	1.02	0.73	0.54
	ECM	0.02320	0.01083	0.00725	0.00273	0.00146	0.00040	0.4	3.45	2.19	1.81	1.01	0.73	0.54
0.3	Var	0.00144	0.00038	0.00028	0.00015	0.00010	0.00004	0.5	3.54	2.19	1.80	0.99	0.73	0.53
	ECM	0.03433	0.01645	0.01075	0.00408	0.00219	0.00060	0.6	3.43	2.20	1.82	1.01	0.74	0.53
0.4	Var	0.00201	0.00052	0.00039	0.00020	0.00013	0.00005	0.7	3.46	2.21	1.80	1.01	0.74	0.54
	ECM	0.04594	0.02192	0.01446	0.00539	0.00291	0.00080	0.8	3.46	2.20	1.78	1.02	0.73	0.53
0.5	Var	0.00502	0.00062	0.00047	0.00024	0.00016	0.00006	0.9	3.47	2.23	1.81	1.02	0.73	0.53
	ECM	0.05900	0.02739	0.01805	0.00661	0.00363	0.00103	1	3.57	2.19	1.79	1.01	0.72	0.55
0.6	Var	0.00298	0.00080	0.00059	0.00029	0.00020	0.00008	2	3.59	2.39	1.99	1.02	0.74	0.53
	ECM	0.06867	0.03304	0.02185	0.00805	0.00446	0.00122	3	3.63	2.43	1.99	1.01	0.74	0.54
0.7	Var	0.00357	0.00094	0.00067	0.00036	0.00023	0.00009	4	3.58	2.41	2.01	1.02	0.74	0.53
	ECM	0.08079	0.03867	0.02516	0.00940	0.00517	0.00142	5	3.54	2.18	1.81	1.03	0.75	0.53
0.8	Var	0.00407	0.00104	0.00080	0.00043	0.00026	0.00010	10	3.55	2.22	1.80	1.02	0.74	0.53
	ECM	0.09215	0.04400	0.02855	0.01086	0.00586	0.00165	15	3.55	2.42	1.97	1.02	0.74	0.52
0.9	Var	0.00453	0.00121	0.00091	0.00046	0.00029	0.00011	20	3.52	2.40	2.02	1.01	0.74	0.54
	ECM	0.10403	0.05026	0.03262	0.01223	0.00661	0.00181	0.1	0.0110	0.0054	0.0035	0.0026	0.0007	0.0003
1	Var	0.01005	0.00129	0.00091	0.00053	0.00032	0.00019	0.2	0.0222	0.0106	0.0071	0.0026	0.0014	0.0007
	ECM	0.11889	0.05477	0.03576	0.01350	0.00723	0.00198	0.3	0.0329	0.0161	0.0105	0.0039	0.0021	0.0010
2	Var	0.01988	0.01304	0.00969	0.00106	0.00067	0.00035	0.4	0.0439	0.0214	0.0141	0.0052	0.0028	0.0014
	ECM	0.23930	0.11952	0.07945	0.02728	0.01471	0.00402	0.5	0.0540	0.0268	0.0176	0.0064	0.0035	0.0017
3	Var	0.02943	0.01970	0.01443	0.00153	0.00098	0.00055	0.6	0.0657	0.0322	0.0213	0.0078	0.0043	0.0020
	ECM	0.36281	0.18189	0.11965	0.04060	0.02207	0.00614	0.7	0.0772	0.0377	0.0245	0.0090	0.0049	0.0024
4	Var	0.04091	0.02599	0.01956	0.00210	0.00127	0.00050	0.8	0.0881	0.0430	0.0278	0.0104	0.0056	0.0027
	ECM	0.47795	0.24079	0.16049	0.05450	0.02951	0.00821	0.9	0.0995	0.0491	0.0317	0.0118	0.0063	0.0030
5	Var	0.04873	0.00636	0.00475	0.00263	0.00149	0.00065	1	0.1088	0.0535	0.0349	0.0130	0.0069	0.0035
	ECM	0.58923	0.27261	0.18122	0.06893	0.03736	0.01001	2	0.2194	0.1065	0.0698	0.0262	0.0140	0.0067
10	Var	0.09936	0.01374	0.00935	0.00523	0.00320	0.00132	3	0.3334	0.1622	0.1052	0.0391	0.0211	0.0103
	ECM	1.18413	0.55528	0.36052	0.13545	0.07384	0.02037	4	0.4370	0.2148	0.1409	0.0524	0.0282	0.0135
15	Var	0.14730	0.09871	0.07159	0.00760	0.00497	0.00186	5	0.5405	0.2662	0.1765	0.0663	0.0359	0.0166
	ECM	1.77734	0.90718	0.59154	0.20431	0.11028	0.03078	10	1.0848	0.5415	0.3512	0.1302	0.0706	0.0333
20	Var	0.18911	0.12856	0.09980	0.01022	0.00637	0.00247	15	1.6300	0.8085	0.5199	0.1967	0.1053	0.0493
	ECM	2.34589	1.19916	0.80970	0.26910	0.14722	0.04067	20	2.1568	1.0706	0.7099	0.2589	0.1409	0.0680

Cuadro 3.11: Valores de la varianza y error cuadrado medio para cada valor κ y tamaño de muestra n proporcionados para el estimador $\hat{\kappa}_1$. En cada valor se realizaron 10,000 repeticiones y se tomó el promedio para la estimación. Las tablas de la derecha, representan: La de arriba el resultado de multiplicar el ECM por el tamaño de la muestra n y dividir entre el valor de κ , mientras que la tabla de abajo el cuadrado del sesgo.

TAMAÑO DE LA MUESTRA n

TAMAÑO DE LA MUESTRA n

κ	75	100	150	200	300	400	500	κ	75	100	150	200	300	400	500
0.1	Var	144.1	27.42	1.34850	0.00028	0.00008	0.00004	0.1	119783.9	28468.8	2092.1	34.0	20.5	14.9	12.0
	ECM	159.7	28.47	1.39473	0.01698	0.00684	0.00240	0.2	110561.0	17379.3	583.8	33.8	20.5	15.0	12.0
0.2	Var	266.9	33.38	0.69453	0.00043	0.00017	0.00008	0.3	101082.5	30948.8	2285.0	34.0	20.5	14.9	12.1
	ECM	294.8	34.76	0.77835	0.03379	0.01365	0.00480	0.4	102131.7	25017.4	74.4	34.1	20.5	15.1	12.0
0.3	Var	364.1	89.31	4.42199	0.00061	0.00026	0.00012	0.5	118138.4	27401.4	188.0	34.3	20.5	15.0	12.1
	ECM	404.3	92.85	4.56992	0.05102	0.02047	0.00723	0.6	121698.3	26515.0	2096.2	34.0	20.5	15.0	12.1
0.4	Var	492.2	96.41	0.04895	0.00136	0.00034	0.00022	0.7	99764.5	32043.8	116.4	40.1	20.3	14.9	11.9
	ECM	544.7	100.07	0.19842	0.06813	0.02729	0.00958	0.8	87519.6	37313.9	2096.3	33.9	20.4	15.1	12.0
0.5	Var	711.4	132.19	0.43538	0.00146	0.00039	0.00021	0.9	107874.1	24584.7	2067.6	34.0	20.3	15.0	12.0
	ECM	787.6	137.01	0.62667	0.08568	0.03424	0.01874	1	109594.2	32435.7	597.0	34.0	20.3	15.1	12.0
0.6	Var	880.0	153.89	8.11016	0.00131	0.00052	0.00025	2	97631.7	29576.6	294.9	34.0	20.4	15.0	11.9
	ECM	973.6	159.09	8.38470	0.10195	0.04107	0.01448	3	108688.0	28407.7	118.1	44.5	20.4	15.0	12.0
0.7	Var	841.3	216.73	0.27098	0.02094	0.00056	0.00038	4	107110.4	19724.6	6101.4	33.9	20.6	15.1	12.0
	ECM	931.1	224.31	0.54319	0.14042	0.04729	0.01668	5	125389.0	28500.5	205.1	34.1	20.5	15.0	11.9
0.8	Var	847.8	288.28	10.81226	0.00176	0.00067	0.00034	10	90381.7	24827.2	898.2	33.6	20.4	15.1	12.1
	ECM	933.5	298.51	11.18008	0.13561	0.05448	0.03018	15	106349.4	23666.0	6198.6	33.9	20.4	15.0	12.1
0.9	Var	1167.4	213.75	12.00732	0.00181	0.00075	0.00035	20	104223.5	32160.7	1393.8	202.1	20.4	15.1	12.0
	ECM	1294.5	221.26	12.40530	0.15280	0.06100	0.03370	0.2	15.6	1.0488	0.0462	0.0167	0.0068	0.0037	0.0024
1	Var	1329.0	312.88	3.55616	0.00255	0.00082	0.00042	0.2	28.0	1.3820	0.0838	0.0334	0.0135	0.0074	0.0047
	ECM	1461.3	324.36	3.97988	0.17013	0.06776	0.03784	0.3	40.3	3.5323	0.1479	0.0504	0.0202	0.0110	0.0071
2	Var	2359.7	570.93	3.12850	0.00540	0.00163	0.00081	0.4	52.5	3.6597	0.1495	0.0668	0.0269	0.0149	0.0094
	ECM	2603.5	591.53	3.93167	0.34043	0.13607	0.07509	0.5	76.2	4.8123	0.1913	0.0842	0.0338	0.0185	0.0118
3	Var	3922.4	823.23	1.19655	0.16013	0.00246	0.00124	0.6	93.6	5.2014	0.2745	0.1006	0.0406	0.0222	0.0142
	ECM	4347.5	852.23	2.36230	0.66809	0.20398	0.11225	0.7	89.9	7.5769	0.2722	0.1195	0.0467	0.0258	0.0164
4	Var	5154.0	753.81	160.14880	0.00972	0.00331	0.00216	0.8	85.8	10.2269	0.3678	0.1339	0.0538	0.0298	0.0188
	ECM	5712.6	788.98	162.70370	0.67762	0.27408	0.15093	0.9	127.1	7.5144	0.3980	0.1510	0.0602	0.0332	0.0213
5	Var	7526.3	1370.69	4.88395	0.01119	0.00416	0.00273	1	132.2	11.4742	0.4237	0.1676	0.0669	0.0373	0.0237
	ECM	8359.3	1425.03	6.83752	0.85363	0.34224	0.18726	2	243.8	20.6044	0.8032	0.3350	0.1344	0.0740	0.0468
10	Var	10908.3	2391.66	55.60507	0.01943	0.00840	0.00402	3	425.1	29.0048	1.1657	0.5080	0.2015	0.1106	0.0705
	ECM	12050.9	2482.72	59.88139	1.67786	0.68055	0.37807	4	558.6	35.1733	2.5549	0.6679	0.2708	0.1488	0.0947
15	Var	19187.8	3420.06	609.76500	0.03269	0.01269	0.00841	5	832.9	54.3400	1.9536	0.8424	0.3381	0.1845	0.1172
	ECM	21269.9	3549.91	619.85570	2.54395	1.01833	0.56160	10	1142.6	91.0650	4.2763	1.6584	0.6722	0.3728	0.2380
20	Var	25101.8	6207.69	176.59260	0.01660	0.01087	0.00844	15	2082.1	129.8510	10.0907	2.5113	1.0056	0.5532	0.3560
	ECM	27792.9	6432.14	185.83890	20.21140	1.35829	0.75536	20	2691.1	224.4530	9.2463	3.6176	1.3417	0.7445	0.4702

Cuadro 3.12: Valores de la varianza y error cuadrado medio para cada valor κ y tamaño de muestra n proporcionados para el estimador $\hat{\kappa}_2$. En cada valor se realizaron 10,000 repeticiones y se tomó el promedio para la estimación. Las tablas de la derecha, representan: La de arriba el resultado de multiplicar el ECM por el tamaño de la muestra n y dividir entre el valor de κ , mientras que la tabla de abajo el cuadrado del sesgo.

κ	TAMAÑO DE LA MUESTRA n										TAMAÑO DE LA MUESTRA n									
	30	40	50	75	100	150	200	κ	30	40	50	75	100	150	200					
0.1	Var	0.00043	0.00034	0.00034	0.00027	0.00022	0.00017	0.00004	0.1	1.110	0.773	1.079	0.926	0.892	0.999					
	ECM	0.00370	0.00193	0.00216	0.00124	0.00089	0.00067	0.00034	0.2	1.104	0.771	1.109	0.936	0.902	1.007					
0.2	Var	0.00085	0.00067	0.00069	0.00054	0.00047	0.00036	0.00006	0.3	1.096	0.781	1.076	0.950	0.891	1.007					
	ECM	0.00736	0.00385	0.00444	0.00250	0.00180	0.00134	0.00068	0.4	1.126	0.786	1.108	0.932	0.914	1.012					
0.3	Var	0.00123	0.00105	0.00096	0.00081	0.00070	0.00053	0.00008	0.5	1.099	0.774	1.092	0.917	0.894	1.015					
	ECM	0.01096	0.00586	0.00646	0.00380	0.00267	0.00201	0.00099	0.6	1.088	0.775	1.092	0.928	0.874	0.997					
0.4	Var	0.00173	0.00136	0.00136	0.00113	0.00092	0.00072	0.00011	0.7	1.086	0.775	1.084	0.948	0.881	1.007					
	ECM	0.01502	0.00786	0.00887	0.00497	0.00365	0.00270	0.00134	0.8	1.087	0.771	1.082	0.936	0.904	0.992					
0.5	Var	0.00211	0.00169	0.00175	0.00137	0.00115	0.00092	0.00023	0.9	1.077	0.776	1.087	0.952	0.919	1.008					
	ECM	0.01831	0.00968	0.01092	0.00611	0.00447	0.00338	0.00174	1	1.100	0.753	1.091	0.926	0.880	1.025					
0.6	Var	0.00246	0.00207	0.00200	0.00169	0.00132	0.00104	0.00017	2	1.104	0.768	1.118	0.950	0.904	1.018					
	ECM	0.02177	0.01162	0.01310	0.00742	0.00525	0.00399	0.00203	3	1.086	0.764	1.122	0.928	0.890	0.974					
0.7	Var	0.00287	0.00236	0.00234	0.00191	0.00158	0.00121	0.00019	4	1.112	0.760	1.092	0.952	0.892	1.011					
	ECM	0.02533	0.01357	0.01518	0.00885	0.00616	0.00470	0.00233	5	1.089	0.771	1.100	0.947	0.891	1.025					
0.8	Var	0.00322	0.00273	0.00264	0.00220	0.00183	0.00140	0.00021	10	1.099	0.793	1.103	0.943	0.893	1.004					
	ECM	0.02898	0.01542	0.01732	0.00998	0.00723	0.00529	0.00267	15	1.088	0.766	1.093	0.898	0.891	0.997					
0.9	Var	0.00366	0.00307	0.00292	0.00248	0.00208	0.00158	0.00024	20	1.096	0.772	1.092	0.941	0.890	1.014					
	ECM	0.03231	0.01745	0.01956	0.01142	0.00827	0.00605	0.00294	0.1	0.0327	0.00159	0.00182	0.00097	0.00067	0.00049					
1	Var	0.00402	0.00336	0.00337	0.00284	0.00224	0.00175	0.00044	0.2	0.00651	0.00318	0.00374	0.00195	0.00134	0.00099					
	ECM	0.03665	0.01884	0.02183	0.01235	0.00880	0.00683	0.00346	0.3	0.00973	0.00481	0.00550	0.00299	0.00197	0.00149					
2	Var	0.00798	0.00680	0.00693	0.00526	0.00452	0.00344	0.00052	0.4	0.01329	0.00650	0.00751	0.00384	0.00274	0.00198					
	ECM	0.07361	0.03841	0.04472	0.02533	0.01807	0.01357	0.00653	0.5	0.01620	0.00799	0.00917	0.00475	0.00332	0.00246					
3	Var	0.01216	0.01042	0.01054	0.00815	0.00680	0.00516	0.00081	0.6	0.01931	0.00955	0.01111	0.00573	0.00392	0.00295					
	ECM	0.10859	0.05731	0.06732	0.03713	0.02669	0.01949	0.00992	0.7	0.02246	0.01120	0.01284	0.00693	0.00459	0.00349					
4	Var	0.01706	0.01373	0.01379	0.01109	0.00925	0.00722	0.00110	0.8	0.02576	0.01270	0.01468	0.00778	0.00540	0.00389					
	ECM	0.14825	0.07599	0.08733	0.05077	0.03569	0.02695	0.01344	0.9	0.02865	0.01438	0.01664	0.00895	0.00619	0.00447					
5	Var	0.02029	0.01737	0.01711	0.01360	0.01128	0.00902	0.00217	1	0.03264	0.01547	0.01845	0.00972	0.00656	0.00508					
	ECM	0.18154	0.09633	0.11003	0.06311	0.04454	0.03417	0.01735	2	0.06563	0.03160	0.03779	0.02007	0.01355	0.01013					
10	Var	0.04041	0.03589	0.03313	0.02652	0.02300	0.01737	0.00449	3	0.09643	0.04689	0.05678	0.02898	0.01990	0.01433					
	ECM	0.36626	0.19829	0.22066	0.12568	0.08935	0.06695	0.03477	4	0.13119	0.06226	0.07354	0.03968	0.02645	0.01234					
15	Var	0.06057	0.05005	0.05088	0.04101	0.03326	0.02552	0.00400	5	0.16125	0.07896	0.09291	0.04951	0.03325	0.02515					
	ECM	0.54425	0.28740	0.32785	0.17969	0.13360	0.09965	0.05036	10	0.32585	0.16240	0.18754	0.09915	0.06635	0.04959					
20	Var	0.08280	0.06831	0.06882	0.05167	0.04617	0.03590	0.00544	15	0.48368	0.23734	0.27697	0.13868	0.10035	0.07413					
	ECM	0.73095	0.38602	0.43684	0.25100	0.17804	0.13520	0.06615	20	0.64815	0.31771	0.36801	0.19933	0.13187	0.09930					

Cuadro 3.13: Valores de la varianza y error cuadrado medio para cada valor κ y tamaño de muestra n proporcionados para el estimador $\hat{\kappa}_3$. En cada valor se realizaron 10,000 repeticiones y se tomó el promedio para la estimación. Las tablas de la derecha, representan: La de arriba el resultado de multiplicar el ECM por el tamaño de la muestra n y dividir entre el valor de κ , mientras que la tabla de abajo el cuadrado del sesgo.

Pruebas de bondad de ajuste

4.1. Introducción

En los capítulos previos se revisó todo el material que le antecede a una prueba de bondad de ajuste, desde las propiedades de la función gama y la distribución gama y log gamma generalizada, se propusieron tres estimadores para el parámetro de forma de la distribución log-gamma generalizada en su forma estándar y finalmente se realizó un análisis detallado de estos estimadores, decidiendo para la prueba de bondad de ajuste trabajar únicamente con dos estimadores el de máxima verosimilitud y el de momentos a través del coeficiente de asimetría. De esta forma quedó el camino despejado para realizar la prueba de bondad de ajuste sobre la distribución log-gamma generalizada.

En este capítulo se verán algunas pruebas de bondad de ajuste para la distribución log-gamma generalizada, (1.3.1). Para esto, como se mencionó, serán utilizados los dos primeros estimadores estudiados en los capítulos previos para el parámetro de forma κ .

El problema a tratar en el presente capítulo se refiere a la prueba de bondad de ajuste para una distribución log-gamma generalizada, con parámetros de localización (ξ), escala (σ) y forma (κ) desconocidos. Es decir, dada la realización de una muestra aleatoria X_1, \dots, X_n con función de densidad $f(x)$ y el contraste de hipótesis

$$\begin{aligned} H_0 : f(x) &\in \mathcal{F}_{LGG}(\xi, \sigma, \kappa) \\ H_1 : f(x) &\notin \mathcal{F}_{LGG}(\xi, \sigma, \kappa) \end{aligned} \tag{4.1.1}$$

establecer pruebas de bondad de ajuste con las que sea factible determinar si se trata de una distribución de la clase \mathcal{F}_{LGG} -densidades log-gamma generalizada con parámetros de localidad, escala y forma desconocidos.

La mecánica de las pruebas de bondad de ajuste, en sus principios generales es conocida, las dificultades que se tienen en un caso particular son varias. Una de tales dificultades ya fue resuelta en el capítulo anterior y se refiere a los estimadores de los parámetros problema que como se sabe, cuando se trata de dos o más parámetros en general es bastante complicado. Para el caso de la distribución en estudio el problema parece simplificarse, porque se trata de

una distribución que pertenece a la familia loc-escala, cuando se fija un valor para el parámetro de forma. Luego el problema de la prueba de hipótesis **4.1.1** puede plantearse como

$$\begin{aligned} H_0 : f(y) &\in \mathcal{F}_{LGG}(0, 1, \kappa) \\ H_1 : f(y) &\notin \mathcal{F}_{LGG}(0, 1, \kappa) \end{aligned} \quad (4.1.2)$$

Así, con base en la invarianza bajo la transformación de los parámetros de localización y escala, será posible establecer algunas pruebas con el coeficiente de correlación muestral, cómo una estadística de prueba. Para cada una de ellas se programó en el proyecto R una función que bajo una muestra de observaciones da respuesta sí las observaciones provienen o no de una distribución LGG para un nivel de significancia α .

En este capítulo serán propuestas las pruebas bootstrap basadas en el coeficiente de correlación como la estadística de prueba. Justamente el capítulo inicia explicando qué es una prueba para \mathcal{F}_{LGG} a través del coeficiente de correlación. Posteriormente, se explica en qué consisten las pruebas de bondad de ajuste de bootstrap y se calcula el tamaño de la prueba para cada prueba. Esta etapa termina, llevando a cabo las pruebas de hipótesis bootstrap, utilizando diferentes distribuciones en la hipótesis alternativa y calculando su potencia para cada prueba. Finalmente, se revisan algunas aplicaciones de esta familia de distribuciones, como son las cartas de control.

4.2. Una prueba para \mathcal{F}_{LGG}

Para construir una prueba para **(4.1.1)**, se observa que si H_0 se cumple, entonces la función de distribución de $f(x)$ es tal que $F(x; \xi, \sigma, \kappa) = F_*\left(\frac{x-\xi}{\sigma}; \kappa\right)$ en donde $F_*(y; \kappa) \in \mathcal{F}_{LGG}(0, 1, \kappa)$. Entonces para una variable aleatoria X con densidad $f(x)$ se tiene que para $F_*^{-1}(v; \kappa)$ la función inversa de $F_*(y; \kappa)$ en y

$$U^* = F_*^{-1}(F(X; \xi, \sigma, \kappa); \kappa) = \frac{X - \xi}{\sigma}. \quad (4.2.1)$$

Es decir, U^* es una función lineal de X .

Por otro lado, sea $F_n(x)$ la función de distribución empírica basada en una muestra aleatoria X_1, X_2, \dots, X_n de $f(x)$. Es decir,

$$F_n(x) = \begin{cases} 0, & x < x_{(1)} \\ \frac{i-0.5}{n}, & x_{(i)} \leq x < x_{(i+1)} \\ 1, & x > x_{(n)} \end{cases} \quad (4.2.2)$$

donde $x_{(1)} \leq \dots \leq x_{(n)}$ son los valores ordenados de las x'_i s, entonces si $\hat{\kappa}$ es un estimador de κ se tiene que para una realización x de X , por **(4.2.1)**

$$u := F_*^{-1}(F_n(x); \hat{\kappa}) \doteq \frac{x - \xi}{\sigma}. \quad (4.2.3)$$

ya que $F_n(x)$ es un estimador de $F(x; \xi, \sigma, \kappa)$. Entonces bajo H_0 , u es aproximadamente una función lineal de x . De donde para probar **(4.1.1)**, se propone como estadística de prueba al

coeficiente de correlación muestral en (4.2.4)

$$r_n(X, U) = \frac{\sum_{i=1}^n (X_i - \bar{X})(U_i - \bar{U})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (U_i - \bar{U})^2}}. \quad (4.2.4)$$

el cual se espera que tome valores cercanos a 1 bajo H_0 .

Es decir, la prueba rechaza H_0 si $r_n(X, U)$ toma valores alejados de 1. O sea, si $r_n(X, U) < r_\alpha$ tal que la prueba es de tamaño α , con $\alpha \in (0, 1)$ un valor conocido entonces se rechaza H_0 .

Lema 4.1. *Sea W_1, W_2, \dots, W_n una muestra aleatoria con parámetros de localidad ω y escala τ ; y V_1, V_2, \dots, V_n otra muestra aleatoria con parámetros de localidad v y escala η , entonces el coeficiente de correlación muestral es invariante bajo localidad y escala.*

Demostración. Sea el coeficiente de correlación muestral $r_n(W, V)$, dado en (4.2.4) y calcule $r_n(W^*, V^*)$ en donde $W^* = \frac{W-\omega}{\tau}$ y $V^* = \frac{V-v}{\eta}$, entonces $W_i^* - \bar{W}^* = \frac{W_i-\omega}{\tau} - \frac{\bar{W}-\omega}{\tau} = \frac{W_i-\bar{W}}{\tau}$, similarmente $V_i^* - \bar{V}^* = \frac{V_i-\bar{V}}{\eta}$. Sustituyendo las igualdades anteriores en $r_n(W^*, V^*)$

$$r_n(W^*, V^*) = \frac{\sum_{i=1}^n \left(\frac{W_i - \bar{W}}{\tau} \right) \left(\frac{V_i - \bar{V}}{\eta} \right)}{\sqrt{\sum_{i=1}^n \left(\frac{W_i - \bar{W}}{\tau} \right)^2 \sum_{i=1}^n \left(\frac{V_i - \bar{V}}{\eta} \right)^2}} = r_n(W, V). \quad \square$$

Similarmente se prueba que el coeficiente de asimetría muestral $\hat{\gamma}$ es invariante bajo localidad y escala. Entonces el estimador de κ , $\hat{\kappa}_2 = G^{-1}(\hat{\gamma})$, es invariante bajo localidad. Por lo tanto, del Lema 4.1 la distribución de $r_n(X, U)$, bajo H_0 no depende de los parámetros ξ, σ para cada $\kappa > 0$; sin embargo, dicha distribución va a depender de κ . Por lo tanto, se propone realizar una prueba de “bootstrap” paramétrica en κ basada en $r_n(X, U)$.

4.3. Pruebas de bondad de ajuste bootstrap

En el caso general cuando las variables x_i 's provienen de una distribución 1.3.1 que bajo un valor conocido del parámetro de forma pertenece a la familia de loc-escala, y el hecho de que el coeficiente de correlación muestral 4.2.4 también es invariante, conducen a simplificar el problema 4.1.1, proponiendo pruebas para el problema simplificado 4.1.2, basadas en el estimador del parámetro de forma κ .

La prueba de bootstrap con base en cualquiera de los dos estimadores, EMV (1) o de momentos (2), se realiza siguiendo los siguientes pasos.

- 1).- Con las observaciones dadas x_1, \dots, x_n calcular un valor para el estimador y considerar su valor absoluto, denótado por $\tilde{\kappa}_{j,0}$ de κ , con $j = 1, 2$.
- 2).- Con las observaciones se calcula su coeficiente de correlación muestral, r_0 .
- 3).- A partir de $\tilde{\kappa}_{j,0}$ iniciar un ciclo de bootstrap.

- a).- Generar una muestra bootstrap de tamaño n de $LGG(0, 1, \tilde{\kappa}_{j,0})$.
 - b).- Con la muestra de bootstrap del inciso (3a) calcular una estimación para el parámetro de forma y considerar su valor absoluto, denotándola por $\tilde{\kappa}_{j,1}$.
 - c).- Con $\tilde{\kappa}_{j,1}$ del inciso 3b se calcula el coeficiente de correlación muestral, \tilde{r}_1 .
- 4).- El ciclo de bootstrap del inciso 3 se repite m veces para calcular $\tilde{r}_1, \dots, \tilde{r}_m$, en donde m -cantidad de estimaciones de bootstrap para determinar el cuantil bootstrap, generalmente se recomienda $m \geq 1000$.
- 5).- Con el ciclo terminado $\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_m$, son ordenados en forma no decreciente, denotándolos con \hat{r}_i . Entonces, $\hat{r}_1 \leq \hat{r}_2 \leq \dots \leq \hat{r}_m$ y obtener el cuantil α , sea éste \hat{r}_α .
- 6).- **Regla de decisión.** Comparar \hat{r}_0 con el cuantil α del inciso 5.
- a).- Si $r_0 \leq \hat{r}_\alpha$, se rechaza H_0 al nivel de significancia α .
 - b).- Si $r_0 > \hat{r}_\alpha$, no se rechaza H_0 al nivel de significancia α .

4.3.1. Consistencia de las pruebas

Antes de iniciar a calcular el tamaño de la prueba y su potencia, falta por verificar que las pruebas utilizadas con el coeficiente de correlación son consistentes, es decir que el coeficiente de correlación muestral se aproxima al coeficiente de correlación poblacional.

Primeramente se tiene que en el capítulo 2 se probó que ambos estimadores $\hat{\kappa}_1$ y $\hat{\kappa}_2$ resultaron ser consistentes en ECM, ver Corolarios 2.6 y 2.16, respectivamente. Ahora para simplificar la comprobación que las pruebas también son consistentes se tiene que el desarrollo es el mismo para cualquiera de las dos pruebas, de tal forma que la comprobación se lleva a efecto para una sola prueba.

Sea Y una variable aleatoria con distribución F y denótese por F_n la distribución empírica correspondiente, entonces por el Teorema de Glivenko-Cantelli la distancia de Kolmogórov $\sup_y |F_n(y) - F(y)|$, ver Sen y Singer 1993 páginas 185-188, converge casi siempre a cero (a.s.). Luego, se tiene para y_1, \dots, y_n

$$F_n(y_i) \xrightarrow{P} F(y_i). \quad (4.3.1)$$

Por consiguiente, se puede formar la secuencia $\{Y_n, U_n\}$ bivariable tal que las U_i 's se obtienen de (4.7.1) al calcular la inversa

$$u_i := F^{-1}(F_n(y_i)). \quad (4.3.2)$$

Por consistencia en el estimador para la prueba, se tiene del ejemplo 3.4.7 en Sen y Singer 1993 páginas 134-136, que $r_n(Y, U)$ dado en 4.2.4 es convergente en probabilidad al parámetro del coeficiente de correlación ρ cuando $n \rightarrow \infty$.

4.4. Tamaño de la prueba

En la sección anterior se revisó el estadístico de prueba que será utilizado en las pruebas bootstrap con el coeficiente de correlación. Pero antes de iniciar con las pruebas se debe determinar si éstas tienen un tamaño de prueba adecuado.

4.4.1. Metodología para determinar el tamaño de la prueba

Como se sabe el tamaño de una prueba se relaciona con la probabilidad del error tipo I, luego se realiza la simulación para el tamaño de prueba considerando que los datos provienen de una muestra de variables aleatorias $LGG(0, 1, \kappa)$, para un valor de κ dado.

Para determinar el tamaño de la prueba con cualquiera de los 2 estimadores, $\hat{\kappa}_1$ y $\hat{\kappa}_2$, se programó una función en el proyecto R, que trabaja de la siguiente forma:

- 1).- Se genera una muestra $x_1, x_2, \dots, x_n \in LGG(0, 1, \kappa_0)$.
- 2).- Con la muestra generada en 1 calcular un valor para el estimador y considerar su valor absoluto, denótado por $\tilde{\kappa}_{j,0}$ de κ con $j = 1, 2$.
- 3).- Con la muestra de 1 y $\tilde{\kappa}_{j,0}$ se calcula el coeficiente de correlación muestral, \hat{r}_0 .
- 4).- A partir de $\tilde{\kappa}_{j,0}$ iniciar un ciclo de bootstrap.
 - a).- Generar una muestra bootstrap de tamaño n de $LGG(0, 1, \tilde{\kappa}_{j,0})$.
 - b).- Con la muestra bootstrap del inciso (4a) calcular una estimación para el parámetro de forma y considerar su valor absoluto, denotándola por $\tilde{\kappa}_{j,1}$.
 - c).- Con $\tilde{\kappa}_{j,1}$ del inciso se calcula su coeficiente de correlación muestral, \tilde{r}_1 .
- 5).- El ciclo de bootstrap del inciso 4 se repite m veces para calcular $\tilde{r}_1, \dots, \tilde{r}_m$, en donde m -cantidad de estimaciones de bootstrap para determinar el cuantil bootstrap, generalmente se recomienda $m \geq 1000$.
- 6).- Con el ciclo terminado $\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_m$ son ordenados en forma no decreciente, denotándolos con \hat{r}_i . Entonces, $\hat{r}_1 \leq \hat{r}_2 \leq \dots \leq \hat{r}_m$ y obtener los cuantiles requeridos. En este caso se tomaron los cuantiles 0.01, 0.02, 0.025, 0.03, 0.04, 0.05, 0.06, 0.07, 0.075, 0.08, 0.09 y 0.10, que se denotan en general \hat{r}_{c_i} , para $c_i = 0.01, 0.02, 0.025, \dots, 0.10$.
- 7).- Comparar \hat{r}_0 con los cuantiles del inciso 6.
 - a).- Si $\hat{r}_0 \leq \hat{r}_{c_i}$, se asigna 1 en el cuantil correspondiente.
 - b).- Si $\hat{r}_0 > \hat{r}_{c_i}$, se asigna 0 en el cuantil correspondiente.
- 8).- Repetir los incisos desde 1 hasta 7 M veces. Finalmente obtener los promedios por cuantil de estas repeticiones, dichos promedios son los tamaños de prueba buscados.

4.4.2. Tamaños de las pruebas con base en el EMV $\hat{\kappa}_1$ y de momentos $\hat{\kappa}_2$

Para los tamaños de prueba, con la metodología anterior se utilizaron $m = 1000$ y $M = 500$ repeticiones. Para la aproximación de los cuantiles se utilizó un error de 0.00001 con 1000 iteraciones y tamaños de muestra $n = 30, 40, 50, 75, 100, 150, 200$, finalmente se analizaron los valores de $\kappa = 0.1, 0.5, 1, 2, 5$. Obteniendo los resultados de los cuadros siguientes.

Tamaños de prueba para $\kappa = 0.1$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.014	0.016	0.018	0.024	0.040	0.050	0.064	0.076	0.080	0.082	0.086	0.096
40	0.010	0.015	0.020	0.020	0.025	0.025	0.060	0.070	0.070	0.075	0.105	0.115
50	0.005	0.015	0.015	0.020	0.030	0.035	0.055	0.075	0.075	0.075	0.085	0.105
75	0.000	0.000	0.005	0.005	0.030	0.030	0.040	0.055	0.065	0.065	0.095	0.100
100	0.004	0.020	0.024	0.028	0.036	0.040	0.056	0.080	0.080	0.080	0.096	0.100
150	0.010	0.020	0.024	0.024	0.038	0.044	0.052	0.056	0.058	0.064	0.072	0.082
200	0.008	0.014	0.020	0.026	0.034	0.040	0.046	0.064	0.066	0.074	0.084	0.090

Tamaños de prueba para $\kappa = 0.5$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.008	0.022	0.030	0.042	0.050	0.058	0.066	0.072	0.074	0.082	0.098	0.102
40	0.000	0.013	0.013	0.016	0.025	0.038	0.054	0.063	0.073	0.079	0.082	0.098
50	0.010	0.023	0.027	0.030	0.033	0.043	0.057	0.070	0.073	0.077	0.087	0.100
75	0.008	0.008	0.016	0.016	0.040	0.048	0.060	0.068	0.072	0.072	0.080	0.092
100	0.007	0.010	0.013	0.020	0.027	0.033	0.057	0.070	0.070	0.070	0.087	0.090
150	0.006	0.008	0.016	0.028	0.038	0.042	0.054	0.068	0.074	0.074	0.084	0.098
200	0.008	0.012	0.016	0.020	0.034	0.042	0.060	0.064	0.070	0.076	0.082	0.094

Tamaños de prueba para $\kappa = 1$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.010	0.015	0.025	0.035	0.040	0.045	0.060	0.075	0.090	0.090	0.090	0.110
40	0.010	0.020	0.030	0.035	0.040	0.040	0.055	0.070	0.080	0.090	0.100	0.110
50	0.005	0.014	0.019	0.019	0.034	0.043	0.048	0.063	0.068	0.077	0.082	0.082
75	0.010	0.010	0.024	0.029	0.034	0.048	0.063	0.072	0.077	0.077	0.077	0.101
100	0.000	0.000	0.010	0.015	0.024	0.029	0.044	0.053	0.058	0.068	0.068	0.068
150	0.005	0.020	0.030	0.030	0.030	0.035	0.045	0.055	0.060	0.080	0.085	0.090
200	0.005	0.005	0.010	0.019	0.044	0.058	0.058	0.058	0.058	0.073	0.078	0.092

Tamaños de prueba para $\kappa = 2$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.010	0.016	0.020	0.024	0.034	0.046	0.060	0.064	0.072	0.078	0.086	0.100
40	0.004	0.004	0.004	0.020	0.032	0.040	0.052	0.064	0.068	0.072	0.084	0.092
50	0.003	0.007	0.010	0.010	0.020	0.033	0.043	0.057	0.067	0.067	0.087	0.097
75	0.003	0.013	0.017	0.020	0.023	0.030	0.047	0.070	0.070	0.073	0.087	0.097
100	0.007	0.023	0.037	0.037	0.047	0.053	0.063	0.067	0.070	0.073	0.083	0.093
150	0.008	0.016	0.020	0.024	0.034	0.042	0.054	0.060	0.068	0.074	0.080	0.088
200	0.008	0.020	0.030	0.030	0.040	0.050	0.052	0.066	0.072	0.078	0.082	0.088

Tamaños de prueba para $\kappa = 5$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.010	0.026	0.026	0.036	0.050	0.052	0.068	0.076	0.082	0.084	0.092	0.102
40	0.010	0.013	0.013	0.023	0.033	0.037	0.040	0.053	0.053	0.073	0.087	0.097
50	0.010	0.013	0.013	0.017	0.023	0.037	0.040	0.040	0.047	0.053	0.057	0.063
75	0.007	0.013	0.020	0.023	0.040	0.047	0.053	0.063	0.070	0.073	0.077	0.090
100	0.007	0.020	0.030	0.030	0.040	0.047	0.057	0.067	0.073	0.080	0.090	0.100
150	0.006	0.018	0.020	0.020	0.022	0.030	0.036	0.048	0.054	0.058	0.060	0.066
200	0.006	0.014	0.022	0.026	0.038	0.040	0.046	0.052	0.056	0.074	0.084	0.092

Cuadro 4.1: Tamaños de prueba para $\kappa_1 = 0.1, 0.5, 1, 2, 5$, usando el EMV y tamaño de muestra $n = 30, 40, \dots, 200$ realizando $m = 1000$ y $M = 500$ repeticiones con aproximaciones de los cuantiles en un error menor a 0.00001 con 1000 iteraciones.

Tamaños de prueba para $\kappa = 0.1$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.020	0.057	0.069	0.083	0.111	0.154	0.166	0.189	0.206	0.217	0.240	0.251
40	0.016	0.036	0.048	0.056	0.092	0.132	0.164	0.208	0.212	0.228	0.264	0.280
50	0.010	0.037	0.063	0.087	0.107	0.133	0.150	0.187	0.213	0.227	0.233	0.257
75	0.016	0.036	0.052	0.080	0.120	0.144	0.188	0.200	0.208	0.220	0.232	0.260
100	0.013	0.033	0.043	0.047	0.087	0.123	0.150	0.173	0.177	0.180	0.190	0.217
150	0.010	0.032	0.040	0.050	0.082	0.100	0.132	0.162	0.174	0.182	0.208	0.220
200	0.002	0.010	0.022	0.032	0.062	0.086	0.122	0.144	0.158	0.172	0.196	0.212

Tamaños de prueba para $\kappa = 0.5$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.014	0.023	0.031	0.034	0.037	0.051	0.074	0.080	0.089	0.097	0.114	0.126
40	0.013	0.023	0.030	0.033	0.040	0.043	0.063	0.077	0.087	0.093	0.110	0.120
50	0.007	0.010	0.017	0.020	0.023	0.040	0.053	0.073	0.087	0.090	0.097	0.107
75	0.003	0.007	0.013	0.020	0.040	0.050	0.073	0.077	0.083	0.087	0.087	0.097
100	0.003	0.007	0.013	0.020	0.023	0.037	0.050	0.063	0.067	0.070	0.080	0.097
150	0.004	0.008	0.010	0.010	0.020	0.038	0.048	0.064	0.066	0.074	0.080	0.090
200	0.008	0.014	0.020	0.024	0.034	0.046	0.052	0.066	0.072	0.074	0.080	0.098

Tamaños de prueba para $\kappa = 1$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.002	0.006	0.010	0.016	0.034	0.042	0.048	0.056	0.068	0.070	0.082	0.092
40	0.007	0.017	0.020	0.027	0.037	0.047	0.053	0.070	0.077	0.090	0.107	0.130
50	0.002	0.008	0.008	0.012	0.012	0.016	0.032	0.046	0.052	0.056	0.072	0.092
75	0.000	0.013	0.017	0.020	0.023	0.030	0.040	0.050	0.053	0.060	0.067	0.070
100	0.010	0.010	0.013	0.013	0.020	0.027	0.033	0.040	0.050	0.050	0.057	0.070
150	0.000	0.003	0.007	0.017	0.023	0.027	0.033	0.043	0.043	0.043	0.043	0.053
200	0.007	0.010	0.013	0.017	0.017	0.023	0.043	0.043	0.043	0.050	0.060	0.070

Tamaños de prueba para $\kappa = 2$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.007	0.010	0.013	0.017	0.027	0.040	0.057	0.070	0.077	0.080	0.087	0.100
40	0.003	0.007	0.010	0.013	0.017	0.027	0.040	0.053	0.063	0.070	0.083	0.090
50	0.000	0.003	0.007	0.013	0.023	0.030	0.037	0.043	0.047	0.053	0.057	0.063
75	0.003	0.003	0.007	0.013	0.027	0.030	0.040	0.057	0.060	0.060	0.063	0.073
100	0.000	0.010	0.013	0.013	0.030	0.040	0.053	0.067	0.073	0.077	0.087	0.097
150	0.002	0.004	0.008	0.012	0.016	0.030	0.036	0.048	0.056	0.058	0.066	0.070
200	0.000	0.002	0.006	0.012	0.014	0.020	0.032	0.042	0.042	0.042	0.052	0.064

Tamaños de prueba para $\kappa = 5$ y valores de α

n	0.010	0.020	0.025	0.030	0.040	0.050	0.060	0.070	0.075	0.080	0.090	0.100
30	0.003	0.010	0.017	0.020	0.023	0.033	0.040	0.053	0.057	0.057	0.070	0.077
40	0.000	0.008	0.008	0.025	0.030	0.040	0.048	0.058	0.063	0.063	0.068	0.070
50	0.003	0.010	0.017	0.020	0.037	0.037	0.040	0.047	0.047	0.050	0.073	0.083
75	0.003	0.010	0.013	0.017	0.023	0.027	0.047	0.057	0.057	0.057	0.070	0.073
100	0.000	0.000	0.008	0.012	0.024	0.032	0.036	0.060	0.064	0.064	0.068	0.068
150	0.000	0.000	0.004	0.008	0.020	0.024	0.034	0.044	0.046	0.052	0.068	0.076
200	0.000	0.002	0.008	0.014	0.022	0.028	0.038	0.058	0.070	0.074	0.082	0.092

Cuadro 4.2: Tamaños de prueba para $\kappa_2 = 0.1, 0.5, 1, 2, 5$, usando el estimador de momentos tamaño de muestra $n = 30, 40, \dots, 200$ realizando $m = 1000$ y $M = 500$ repeticiones con aproximaciones de los cuantiles en un error menor a 0.00001 con 1000 iteraciones.

NOTA 4.1. De las tablas 4.1 y 4.2 se puede establecer:

- 1).- El estimador de máxima verosimilitud funciona bien con la prueba bootstrap para los diferentes valores de κ , aunque al aumentar su valor en muestras de tamaño 30 el tamaño de la prueba rebasa ligeramente el valor de α .
- 2).- El estimador de momentos no funciona bien para muestras pequeñas y valores de $\kappa < 0.05$, pero en valores mayores de κ sus tamaños de prueba son ligeramente mejores que los del estimador de máxima verosimilitud.
- 3).- Esto concuerda con los resultados obtenidos en el capítulo previo sobre el valor del estimador y su error cuadrado medio.

4.5. Potencias de las pruebas Bootstrap paramétrico

En la sección anterior se revisó el tamaño de la prueba, para la prueba bootstrap paramétrica, obviamente el paso siguiente consiste en determinar la potencia de la prueba para el problema (4.1.2), determinar si los datos muestrales analizados pertenecen a la familia *LGG*. Por tratarse de una prueba bootstrap, se proporciona un programa en el Proyecto R que calcule la potencia de la prueba.

4.5.1. Metodología para determinar la potencia de las pruebas

Para determinar la potencia de la prueba con cualquiera de los 2 estimadores, se programó una función en el proyecto R que trabaja de la siguiente forma:

- 1).- Se genera una muestra $x_1, x_2, \dots, x_n \in \mathcal{H}$. En donde, \mathcal{H} es la distribución alternativa.
- 2).- Con la muestra generada en 1 calcular una estimación para el parámetro de forma y considerar su valor absoluto, denotándola por $\tilde{\kappa}_{j,0}$ de κ , con $j = 1, 2$.
- 3).- Con la muestra generada en 1 y $\tilde{\kappa}_{j,0}$ se calcula el coeficiente de correlación muestral, \hat{r}_0 .
- 4).- A partir de $\tilde{\kappa}_{j,0}$ iniciar un ciclo de bootstrap.
 - a).- Generar una muestra bootstrap de tamaño n de $LGG(0, 1, \tilde{\kappa}_{j,0})$.
 - b).- Con la muestra bootstrap del inciso 4a calcular una estimación para el parámetro de forma y considerar su valor absoluto, denotándola por $\tilde{\kappa}_{j,1}$.
 - c).- Con $\tilde{\kappa}_{j,1}$ del inciso 4b se calcula su coeficiente de correlación muestral, \tilde{r}_1 .
- 5).- El ciclo de bootstrap del inciso 4 se repite m veces para calcular $\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_m$, en donde m - cantidad de estimaciones de bootstrap para determinar el cuantil bootstrap, generalmente se recomienda $m \geq 1000$.
- 6).- Con el ciclo terminado $\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_m$ son ordenados en forma no decreciente, denotándolos con \hat{r}_i . Entonces, $\hat{r}_1 \leq \hat{r}_2 \leq \dots \leq \hat{r}_m$ y obtener los cuantiles requeridos. En este caso se tomaron los cuantiles 0.01, 0.025, 0.05, 0.075 y 0.10, que se denotan en general \hat{r}_{c_i} , para $c_i = 0.01, 0.025, \dots, 0.10$.
- 7).- Comparar \hat{r}_0 con los cuantiles del inciso 6.
 - a).- Si $\hat{r}_0 \leq \hat{r}_{c_i}$, se asigna 1 en el cuantil correspondiente.

- b).- Si $\tilde{r}_0 > \hat{r}_{c_i}$, se asigna 0 en el cuantil correspondiente.
- 8).- Repetir los incisos desde el inciso 1 hasta 7 M veces. Finalmente se obtienen los promedios por cuantil de estas repeticiones, dichos promedios son las potencias de las pruebas.

4.5.2. Distribuciones alternativas para el estudio de la potencia de la prueba

En general el problema de decidir qué distribuciones utilizar para calcular la potencia de la prueba, para el caso bootstrap paramétrico es sencillo, debido a la invarianza del coeficiente de asimetría con respecto a localidad y escala, y el hecho de que dos segmentos siempre son homotéticos, las densidades no necesariamente deben tener el mismo soporte, como era una primera idea al respecto.

Luego, las distribuciones que serán utilizadas en la potencia de la prueba bootstrap, se han clasificado en tres tipos: simétricas, sesgadas y las que pueden ser de alguno de los tipos anteriores, según se elijan sus parámetros. Se muestran sus funciones de densidad de cada una de ellas.

Distribuciones simétricas

- 1).- Normal $f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ con $x \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma > 0$.
- 2).- t-student $f(x; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}} (1 + \frac{x^2}{\nu})^{-(\nu+1)/2}$ con $x \in \mathbb{R}$, $\nu > 0$.
- 3).- Logística $f(x; \mu, \beta) = \frac{e^{-(x-\mu)/\beta}}{\beta(1+e^{-(x-\mu)/\beta})^2}$ con $x \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\beta > 0$.
- 4).- Laplace $f(x; \mu, \sigma) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}$ con $x \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma > 0$.
- 5).- Cauchy $f(x; \theta, \sigma) = \frac{1}{\pi\sigma [1 + (\frac{x-\theta}{\sigma})^2]}$ con $x \in \mathbb{R}$, $\theta \in \mathbb{R}$, $\sigma > 0$.
- 6).- t-no central $f(x; \nu, \delta) = \frac{U+\delta}{\sqrt{V/\nu}}$, $U \sim N(0, 1)$, $V \sim \chi_\nu^2$ independientes $\delta \in \mathbb{R}$, $\nu > 0$.

Distribución que puede ser asimétrica o simétrica

- 7).- Beta $f(x; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$ con $0 < x < 1$, $\alpha, \beta > 0$.

Distribuciones asimétricas

- 8).- Gumbel $f(x; \alpha, \beta) = \frac{1}{\beta} e^{-(x-\alpha)/\beta} \exp(-e^{-(x-\alpha)/\beta})$ con $x \in \mathbb{R}$, $\alpha \in \mathbb{R}$, $\beta > 0$.
- 9).- Gamma $f(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x\lambda}$ con $x > 0$, $\alpha, \lambda > 0$.
- 10).- Ji-cuadrada $f(x; \nu) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}$ con $x > 0$, $\nu > 0$.
- 11).- Weibull $f(x; \alpha, \beta) = \frac{\alpha x^{\alpha-1}}{\beta} e^{-x^\alpha/\beta}$ con $x > 0$, $\alpha, \beta > 0$.

- 12).- F-snedecor $f(x; \nu_1, \nu_2) = \frac{\Gamma(\frac{\nu_1+\nu_2}{2})\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2}}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}x^{\nu_1/2-1}\left(1 + \frac{\nu_1}{\nu_2}x\right)^{-(\nu_1+\nu_2)/2}$, $x > 0$, $\nu_1, \nu_2 > 0$.
- 13).- Fréchet $f(x; \alpha, \beta, \mu) = \frac{\alpha}{\beta\left(\frac{x-\mu}{\beta}\right)^{\alpha+1}} \exp\left(-\frac{1}{\left(\frac{x-\mu}{\beta}\right)^\alpha}\right)$ con $x > \mu$, $\mu \in \mathbb{R}$ $\alpha, \beta > 0$.
- 14).- Exponencial $f(x; \beta) = \frac{1}{\beta}e^{-x/\beta}$ con $x > 0$, $\beta > 0$.
- 15).- Pareto $f(x; \theta, \beta) = \frac{\beta\theta^\beta}{x^{\beta+1}}$ con $x > \theta$, $\theta, \beta > 0$.
- 16).- Log-normal $f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-(\log(x)-\mu)^2/(2\sigma^2)}$ con $x > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$.
- 17).- Log-Logística $f(x; \mu, \beta) = \frac{e^{-(\log(x)-\mu)/\beta}}{x\beta\left(1+e^{-(\log(x)-\mu)/\beta}\right)^2}$ con $x > 0$, $\mu \in \mathbb{R}$, $\beta > 0$.
- 18).- Log-Cauchy $f(x; \theta, \sigma) = \frac{1}{x\pi\sigma\left[1+\left(\frac{\log(x)-\theta}{\sigma}\right)^2\right]}$ con $x > 0$, $\theta \in \mathbb{R}$, $\sigma > 0$.
- 19).- Log t-student $f(x; \mu, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{x\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}}\left(1 + \frac{(\log(x)-\mu)^2}{\nu}\right)^{-(\nu+1)/2}$ con $x > 0$, $\mu \in \mathbb{R}$, $\nu > 0$.
- 20).- Log-Laplace $f(x; \mu, \sigma) = \frac{1}{2x\sigma}e^{-|\log(x)-\mu|/\sigma}$ con $x > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$.

4.5.3. Resultados de la Potencia de la prueba

En esta subsección se presentan algunas tablas que muestran la potencia de la prueba para ambos estimadores, $\hat{\kappa}_1$ y $\hat{\kappa}_2$, del parámetro de forma de la familia $LGG(0, 1, \kappa)$, con diferentes valores de significancia (0.01, 0.025, 0.05, 0.075 y 0.10) y tamaños de muestra (30, 40, 50, 75, 100, 150 y 200), aplicados a cada una de las distribuciones mostradas en la sección anterior. Se hace una diferencia entre las distribuciones que tienen soporte en todo \mathbb{R} y son simétricas con respecto a otras distribuciones que son asimétricas.

Las tablas están dadas de la siguiente forma, por columna se muestra el nivel de significancia y por filas las potencias de la prueba. Para cada nivel de significancia la columna izquierda corresponde a la prueba para el EMV $\hat{\kappa}_1$ y la columna derecha la prueba para el estimador de momentos $\hat{\kappa}_2$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.000	0.065	0.020	0.100	0.040	0.140	0.050	0.155	0.070	0.170
40	0.010	0.080	0.020	0.110	0.050	0.160	0.090	0.170	0.120	0.200
50	0.020	0.075	0.030	0.115	0.060	0.155	0.100	0.215	0.130	0.240
75	0.020	0.083	0.060	0.118	0.060	0.164	0.100	0.218	0.140	0.246
100	0.050	0.084	0.060	0.119	0.090	0.166	0.120	0.234	0.140	0.260
150	0.076	0.087	0.099	0.120	0.116	0.169	0.170	0.235	0.207	0.261
200	0.117	0.087	0.147	0.120	0.157	0.169	0.212	0.241	0.251	0.265

Cuadro 4.3: Potencias para la distribución alternativa normal con $\mu = 0$ y $\sigma = 0.2$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.130	0.200	0.190	0.265	0.260	0.325	0.270	0.380	0.300	0.390
40	0.220	0.200	0.280	0.260	0.320	0.330	0.400	0.380	0.420	0.405
50	0.320	0.255	0.400	0.335	0.460	0.395	0.500	0.425	0.540	0.495
75	0.290	0.296	0.350	0.389	0.440	0.446	0.500	0.459	0.560	0.569
100	0.367	0.359	0.442	0.473	0.514	0.523	0.581	0.510	0.636	0.682
150	0.450	0.444	0.532	0.584	0.609	0.627	0.689	0.579	0.755	0.833
200	0.502	0.511	0.591	0.673	0.680	0.710	0.750	0.634	0.830	0.953

Cuadro 4.4: Potencias para la distribución alternativa t-Student con $\nu = 5$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.075	0.080	0.135	0.160	0.185	0.210	0.250	0.250	0.265	0.300
40	0.080	0.096	0.150	0.192	0.210	0.252	0.270	0.300	0.305	0.360
50	0.145	0.170	0.200	0.208	0.255	0.273	0.280	0.325	0.305	0.390
75	0.160	0.174	0.255	0.217	0.325	0.284	0.370	0.338	0.405	0.406
100	0.181	0.236	0.266	0.248	0.336	0.326	0.372	0.388	0.410	0.465
150	0.224	0.294	0.320	0.277	0.397	0.364	0.425	0.433	0.469	0.520
200	0.263	0.361	0.369	0.305	0.450	0.400	0.476	0.476	0.522	0.572

Cuadro 4.5: Potencias para la distribución alternativa logística con $\mu = 0$ y $\beta = 2$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.260	0.250	0.350	0.330	0.400	0.400	0.410	0.435	0.470	0.470
40	0.340	0.310	0.450	0.400	0.490	0.485	0.560	0.575	0.590	0.620
50	0.390	0.340	0.460	0.435	0.520	0.528	0.580	0.645	0.610	0.695
75	0.510	0.375	0.560	0.476	0.670	0.577	0.760	0.726	0.800	0.782
100	0.532	0.410	0.584	0.517	0.688	0.627	0.793	0.809	0.822	0.871
150	0.578	0.447	0.622	0.560	0.738	0.679	0.856	0.895	0.883	0.963
200	0.622	0.485	0.657	0.604	0.784	0.733	0.915	0.983	0.939	1.000

Cuadro 4.6: Potencias para la distribución alternativa Laplace con $\mu = 0$ y $\sigma = 2$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.040	0.230	0.120	0.410	0.170	0.460	0.200	0.490	0.250	0.540
40	0.120	0.340	0.250	0.370	0.290	0.470	0.320	0.500	0.380	0.570
50	0.130	0.400	0.210	0.520	0.350	0.560	0.450	0.650	0.510	0.700
75	0.270	0.499	0.400	0.592	0.510	0.609	0.540	0.730	0.640	0.777
100	0.283	0.611	0.420	0.693	0.542	0.681	0.601	0.846	0.697	0.887
150	0.326	0.719	0.472	0.779	0.605	0.740	0.667	0.943	0.771	1.000
200	0.365	0.835	0.520	0.879	0.663	0.811	0.728	1.000	0.840	1.000

Cuadro 4.7: Potencias para la distribución alternativa Beta con $\alpha = 4$ y $\beta = 10$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.930	0.825	0.950	0.900	0.970	0.920	0.970	0.950	0.970	0.960
40	0.960	0.900	0.970	0.935	0.980	0.970	0.980	0.980	0.980	0.990
50	0.980	0.942	0.990	0.955	0.990	1.000	0.990	1.000	0.990	1.000
75	1.000	0.980	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Cuadro 4.8: Potencias para la distribución alternativa Cauchy con $\theta = 0$ y $\sigma = 2$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.340	0.605	0.390	0.740	0.460	0.785	0.530	0.795	0.580	0.795
40	0.460	0.740	0.600	0.800	0.640	0.870	0.680	0.885	0.720	0.905
50	0.610	0.808	0.690	0.830	0.740	0.912	0.790	0.930	0.800	0.960
75	0.790	0.886	0.870	0.865	0.890	0.962	0.920	0.982	0.950	1.000
100	0.842	0.966	0.937	0.900	0.953	1.000	0.978	1.000	0.994	1.000
150	0.928	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Cuadro 4.9: Potencias para la distribución alternativa Gumbel con $\alpha = 0$ y $\beta = 2$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.600	0.630	0.740	0.680	0.820	0.790	0.860	0.800	0.860	0.810
40	0.740	0.720	0.850	0.760	0.900	0.810	0.910	0.820	0.920	0.830
50	0.860	0.790	0.920	0.870	0.950	0.920	0.980	0.940	0.980	0.970
75	0.980	0.812	1.000	0.887	1.000	0.926	1.000	0.946	1.000	0.977
100	1.000	0.832	1.000	0.923	1.000	0.965	1.000	0.989	1.000	1.000

Cuadro 4.10: Potencias para la distribución alternativa gamma con $\alpha = 4$ y $\beta = 5$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.100	0.320	0.185	0.415	0.275	0.495	0.320	0.530	0.355	0.550
40	0.150	0.400	0.360	0.519	0.440	0.619	0.500	0.663	0.560	0.688
50	0.370	0.480	0.460	0.623	0.520	0.742	0.580	0.795	0.610	0.825
75	0.560	0.600	0.640	0.778	0.770	0.928	0.810	0.928	0.860	0.962
100	0.613	0.810	0.697	0.934	0.811	1.000	0.857	1.000	0.908	1.000
150	0.707	0.930	0.781	0.980	0.902	1.000	0.947	1.000	1.000	1.000
200	0.873	1.000	0.916	1.000	0.985	1.000	1.000	1.000	1.000	1.000

Cuadro 4.11: Potencias para la distribución alternativa Ji-cuadrada con $\nu = 20$.

<i>n</i>	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.040	0.370	0.160	0.500	0.230	0.530	0.260	0.620	0.390	0.680
40	0.150	0.510	0.220	0.560	0.310	0.660	0.420	0.730	0.470	0.770
50	0.150	0.620	0.300	0.730	0.380	0.780	0.510	0.830	0.520	0.850
75	0.420	0.756	0.610	0.840	0.760	0.912	0.810	0.941	0.880	0.940
100	0.514	0.830	0.723	0.941	0.772	1.000	0.847	1.000	0.890	1.000
150	0.557	0.940	0.776	0.970	0.876	1.000	0.949	1.000	0.986	1.000
200	0.819	1.000	0.912	1.000	0.975	1.000	1.000	1.000	1.000	1.000

Cuadro 4.12: Potencias para la distribución alternativa Weibull con $\alpha = 2$ y $\beta = 2$.

n	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.945	0.990	0.975	0.990	0.995	0.990	1.000	0.990	1.000	0.990
40	0.980	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Cuadro 4.13: Potencias para la distribución alternativa F con $\nu_1 = 4$ y $\nu_2 = 8$.

n	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.960	0.990	0.960	0.990	0.960	0.990	0.960	0.990	0.960	0.990
40	0.995	1.000	0.995	1.000	0.995	1.000	0.995	1.000	0.995	1.000

Cuadro 4.14: Potencias para la distribución alternativa Fréchet con $\mu = 0$, $\alpha = 2$ y $\beta = 5$.

n	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.910	0.980	0.960	0.990	0.990	1.000	0.990	1.000	1.000	1.000
40	0.980	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Cuadro 4.15: Potencias para la distribución alternativa Pareto con $\theta = 4$ y $\beta = 20$.

n	0.010	0.010	0.025	0.025	0.050	0.050	0.075	0.075	0.100	0.100
30	0.830	0.990	0.910	1.000	0.940	1.000	0.980	1.000	0.980	1.000
40	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Cuadro 4.16: Potencias para la distribución alternativa exponencial con $\lambda = 0.1$.

La potencia para las alternativas restantes:

- Log-normal con $\mu = 0$ y $\sigma = 2$.
- Log-logística con $\mu = 0$ y $\beta = 2$.
- Log-Cauchy con $\theta = 0$ y $\sigma = 2$.
- Log-T-student con $\mu = 0$ y $\nu = 5$.
- Log-Laplace con $\mu = 0$ y $\sigma = 2$.

Para tamaños de muestra $n \geq 30$ fue 1 en ambas pruebas.

NOTA 4.2. De las tablas para la potencia de la prueba se puede establecer:

- 1).- Las pruebas funcionan mejor mientras más pesada es la cola.
- 2).- Ambas pruebas tienen potencias elevadas en el caso de distribuciones sesgadas a la derecha.
- 3).- Las potencias para la prueba del estimador de momentos son mejores que para el EMV.
- 4).- Las pruebas son un poco bajas en el caso de distribuciones simétricas y tamaños de muestra pequeños.

4.6. Pruebas de bondad de ajuste para $LGG(\xi, \sigma, \kappa)$

En las secciones anteriores se revisó el problema de la prueba de bondad de ajuste para el problema 4.1.2, pero la pregunta obvia consiste en determinar qué paso con los parámetros ξ y σ . Como se mostró en el desarrollo del estudio, quedó establecida la prueba para determinar si los datos de una realización provienen o no de una distribución LGG , y que bajo la invarianza de los parámetros de localidad y escala se redujo a estudiar el problema 4.1.2. Para dar respuesta al problema 4.1.1 se utilizan los resultados de la distribución estándar, puesto que ahora se conoce el valor del estimador de κ , denotado por $\hat{\kappa}$, queda por conocer los otros dos valores que estiman a los parámetros ξ y σ .

Este problema quedó resuelto en la sección 2.2.1, en donde quedó expresada la respuesta para el caso cuando es conocido el valor del parámetro de forma, κ , se tiene que resolver primeramente la ecuación para el parámetro de escala, σ , dada por (2.2.7), este valor se sustituye en (2.2.6) para obtener el valor del parámetro de localidad ξ .

NOTA 4.3. Concluyendo el desarrollo de la sección, resulta que dada una realización x_1, \dots, x_n , para llevar a cabo la prueba de bondad de ajuste para el problema completo 4.1.1 se pueden seguir los siguientes pasos:

- 1).- Con la realización y un nivel de significancia α llevar a cabo la prueba bootstrap, propuesta en las secciones anteriores.
- 2).- Si la prueba bootstrap indica que los datos siguen una distribución $LGG(0, 1, \hat{\kappa})$, continuar en caso contrario se concluye que al nivel de significancia α los datos no siguen una distribución $LGG(0, 1, \hat{\kappa})$.
- 3).- Con el valor estimado $\hat{\kappa}$, resolver la ecuación para σ

$$\sum_{i=1}^n \exp\left(\frac{x_i - \bar{x}}{\sigma \hat{\kappa}}\right) - n \hat{\kappa} e^{-\psi(\hat{\kappa})} = 0$$

- 4).- Con el valor de $\hat{\sigma}$ calcular el valor de $\hat{\xi}$.

$$\xi = \bar{x} + \hat{\sigma} \sqrt{\hat{\kappa}} (\log(\hat{\kappa}) - \psi(\hat{\kappa})).$$

- 5).- Concluir: Se determinó que al nivel de significancia α los datos siguen una distribución $LGG(\hat{\xi}, \hat{\sigma}, \hat{\kappa})$.

4.6.1. Función en R para realizar la prueba de bondad de ajuste

Finalmente al haber revisado la prueba de bondad de ajuste para $LGG(0, 1, \kappa)$ y la parte complementaria sobre sus otros dos estimadores de los parámetros ξ y σ , es conveniente proporcionar una función que realice los cálculos para una muestra dada. Para esto primeramente se proporcionará la metodología para determinar, bajo cada una de las dos pruebas propuestas, los valores críticos. Por tratarse de una prueba bootstrap, que trabaja con la muestra de estudio,

se proporciona un programa en el Proyecto R, que dada una realización de la muestra, nivel de significancia, cantidad de repeticiones, error de aproximación y cantidad de iteraciones, indique si hay evidencias para rechazar o no la hipótesis nula en (4.1.2).

- 1).- Con el estimador (1 o 2) y las observaciones, calcular una estimación de κ , en caso de ser negativa se considera su valor absoluto, denótandolo por $\tilde{\kappa}_0$.
En el caso del estimador de momentos la estimación de κ no tiene problemas, porque está basada en el coeficiente de asimetría y éste es invariante en localidad y escala. Para el caso del EMV y debido a que la distribución está suponiéndose con los tres parámetros se utiliza la función `optim` del proyecto R, en lugar de `optimize`.
- 2).- Con las observaciones y $\tilde{\kappa}_0$ calcular una valor del estimador del coeficiente de correlación, sea \tilde{r}_0 .
- 3).- A partir de $\tilde{\kappa}_0$ iniciar un ciclo bootstrap.
 - a).- Generar una muestra bootstrap de tamaño n de $LGG(0, 1, \tilde{\kappa}_0)$.
 - b).- Con el estimador correspondiente (1 o 2) y la muestra bootstrap del inciso (3a) calcular una estimación para el parámetro de forma y considerar su valor absoluto, denotándola por $\tilde{\kappa}_1$.
 - c).- Con la muestra bootstrap del inciso (3a) y $\tilde{\kappa}_1$ calcular una estimación del coeficiente de correlación, sea \tilde{r}_1 .
 - d).- Guardar el estimador $\tilde{\kappa}_1$
- 4).- Con el ciclo bootstrap iniciado calcular $\tilde{r}_2, \tilde{r}_3, \dots, \tilde{r}_m$, en donde m es la cantidad de estimaciones bootstrap para determinar el cuantil bootstrap, generalmente se recomienda $m \geq 1000$.
- 5).- Con el ciclo terminado $\tilde{r}_1, \tilde{r}_2, \dots, \tilde{r}_m$, ordenarlos en forma no decreciente, denotándolos con \hat{r}_i . Luego, $\hat{r}_1 \leq \hat{r}_2 \leq \dots \leq \hat{r}_m$ y determinar la constante crítica del coeficiente de correlación al nivel α de significancia. Es decir, $\hat{r}_{[\alpha m]}$.
- 6).- Comparar \tilde{r}_0 con la constante crítica $\hat{r}_{[\alpha m]}$ y aplicar la regla de decisión.
 - a).- Si $\tilde{r}_0 < \hat{r}_{[\alpha m]}$, rechazar la hipótesis nula al α de significancia.
 - b).- Si $\tilde{r}_0 > \hat{r}_{[\alpha m]}$, no rechazar la hipótesis nula al α de significancia.
- 7).- Con los valores $\tilde{\kappa}_i$ guardados en el paso 3d obtener el estimador de κ por medio de su media. Con este valor y (2.2.6)- (2.2.7) calcular los estimadores para ξ y σ . En este punto al resolver (2.2.7) se recomienda encontrar con `optimize` el mínimo del valor absoluto de esta expresión.

4.7. Aplicaciones

El estudio de familias de distribuciones que tienen varios parámetros siempre resulta atractivo, debido a que con ellas puede darse respuesta a diferentes aplicaciones. La familia que se estudió en el trabajo no es la excepción puesto que con sus tres parámetros y transformaciones muy sencillas como $Y = -X$ o $Y = \log(X)$ representan distribuciones del tipo: Weibull, Gumbel, Exponencial, entre muchas otras. Por tales razones su uso resulta atractivo.

En Control Estadístico del Proceso juegan un papel muy importante las Cartas de Control, introducidas por Walter A. Shewhart en 1920, porque constituyen un procedimiento estadístico de control que ayudan a eliminar la variabilidad del proceso de producción o al menos disminuirla tanto como sea posible. Una de las limitantes de las cartas de control reside en que siguen utilizándose en procesos de producción con distribución normal, esto se debe a que resultan varias complicaciones para obtener metodologías que sirvan para determinar las cartas de control de un proceso de producción con distribución diferente a la normal.

En esta sección y con ayuda del material desarrollado en el trabajo se propondrá la metodología y las cartas de control para procesos de producción con distribución que sea de alguno de los tipos de distribuciones que puede abarcar la familia log-gamma generalizada.

4.7.1. Cartas de control \bar{X} y S bajo una política 3σ

Es bien conocido que las cartas de control \bar{X} y S bajo 3σ (3 desviaciones estándar) trabajan con los límites dados en las expresiones (4.7.1) y (4.7.2)

$$LIC_{\bar{X}} = \bar{\bar{X}} - A_2\bar{S} \qquad LIC_S = \bar{S}D_3 \qquad (4.7.1)$$

$$LSC_{\bar{X}} = \bar{\bar{X}} + A_2\bar{S} \qquad LSC_S = \bar{S}D_4 \qquad (4.7.2)$$

En donde, $A_2 = \frac{3}{d_2\sqrt{n}}$ y $d_2 = E(S)$; $D_3 = 1 - 3\frac{d_3}{d_2}$, $D_4 = 1 + 3\frac{d_3}{d_2}$ con $d_3 = \sqrt{V(S)}$.

Luego, para poder determinar los límites de control es necesario determinar la distribución de la estadística S_{n-1} , para esto se obtiene una distribución asintótica por medio del **método delta** para calcular el valor esperado y la desviación estándar de S_{n-1} .

Las distribuciones de S_{n-1}^2 y S_{n-1} están en Sen and Singer (1993), en el Ejemplo 3.4.1 páginas 126 y 127 se encuentra la distribución para S_{n-1}^2 y en el ejemplo 3.4.3 página 132 la distribución para S_{n-1} , los resultados se resumen en el teorema 4.2.

Teorema 4.2. *Sea $\{X_n\}$ una sucesión de variables iid con media μ y varianza σ^2 , además $E(X_i^4) < \infty$ y denotando $E((X_i - \mu)^4) = \mu_4 < \infty$ y $Var((X_i - \mu)^2) = \mu_4 - \sigma^4 = \gamma^2$, entonces*

$$S_{n-1}^2 \xrightarrow{D} N\left(\sigma^2, \frac{\gamma^2}{n}\right), \qquad (4.7.3)$$

$$S_{n-1} \xrightarrow{D} N\left(\sigma, \frac{\gamma^2}{4\sigma^2 n}\right), \qquad (4.7.4)$$

Luego, del Teorema 4.2 y (4.7.4) resulta

$$E(S_{n-1}) = \sigma \text{ y } V(S_{n-1}) = \frac{\gamma^2}{4\sigma^2 n}. \qquad (4.7.5)$$

4.7.2. Cartas de control \bar{X} y S para una distribución log-gamma generalizada

En el caso de la distribución log-gamma generalizada en su forma estándar, ver proposiciones 1.6 y 1.8, se obtuvo

$$E(T - \mu)^2 = \kappa\psi'(\kappa) \quad (4.7.6)$$

$$E(T - \mu)^4 = \kappa^2 \left[\psi'''(\kappa) + 3(\psi'(\kappa))^2 \right]. \quad (4.7.7)$$

Luego, $\sigma^2 = \kappa\psi'(\kappa)$ y $\mu_4 = \kappa^2 \left[\psi'''(\kappa) + 3(\psi'(\kappa))^2 \right]$, de donde $\gamma^2 = \kappa^2 \left[\psi'''(\kappa) + 2(\psi'(\kappa))^2 \right]$. Se puede notar que las condiciones del teorema 4.2 se cumplen para $\kappa \in [\delta, \infty)$ con $\delta > 0$, ver la propiedad 10 de la función gamma. Por lo tanto, en el caso de la distribución log-gamma generalizada las distribuciones asintóticas de S_{n-1}^2 y S_{n-1} , ver (4.7.3) y (4.7.4), se obtienen sustituyendo los valores para σ^2 y γ^2

$$S_{n-1}^2 \xrightarrow{D} N \left(\kappa\psi'(\kappa), \frac{\kappa^2 \left[\psi'''(\kappa) + 2(\psi'(\kappa))^2 \right]}{n} \right), \quad (4.7.8)$$

$$S_{n-1} \xrightarrow{D} N \left(\sqrt{\kappa\psi'(\kappa)}, \frac{\kappa \left[\psi'''(\kappa) + 2(\psi'(\kappa))^2 \right]}{4\psi'(\kappa)n} \right). \quad (4.7.9)$$

Así, para los límites de las cartas de control dados por (4.7.1) y (4.7.2) para \bar{X} y S , en el caso de la distribución log-gamma generalizada

$$d_2 = E(S) = \sqrt{\kappa\psi'(\kappa)} \quad (4.7.10)$$

$$d_3 = \sqrt{V(S)} = \sqrt{\frac{\kappa \left[\psi'''(\kappa) + 2(\psi'(\kappa))^2 \right]}{4\psi'(\kappa)n}} \quad (4.7.11)$$

Mientras que $A_2 = \frac{3}{d_2\sqrt{n}}$, $D_3 = 1 - 3\frac{d_3}{d_2}$ y $D_4 = 1 + 3\frac{d_3}{d_2}$. En las tablas 4.18 - 4.27 se muestran los valores de estos términos para tamaños de muestra $n = 5, 6, \dots, 50$ y los valores respectivos de $\kappa = 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 1, 5, 10$ y 50 .

En la tabla 4.17 se muestran los niveles confianza para los valores de κ utilizados, así como la cantidad máxima de artículos defectuosos por cada 10,000 producidos.

κ		Nivel de confianza		Defectuosos por cada 10,000 artículos	
0.1	0.75	0.9819	0.9867	181	133
0.2	1	0.9825	0.9881	175	119
0.3	5	0.9833	0.9945	167	55
0.4	10	0.9841	0.9958	159	42
0.5	50	0.9849	0.9970	151	30

Cuadro 4.17: Niveles de confianza para la $LGG(0, 1, \kappa)$ con 3σ y cantidad de defectuosos por cada 10,000 producidos.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	1.9930	0.4213	0.0000	2.8773	28	0.8422	0.1780	0.2067	1.7933
6	1.8194	0.3846	0.0000	2.7138	29	0.8276	0.1749	0.2205	1.7795
7	1.6844	0.3560	0.0000	2.5866	30	0.8137	0.1720	0.2336	1.7664
8	1.5756	0.3330	0.0000	2.4842	31	0.8004	0.1692	0.2460	1.7540
9	1.4855	0.3140	0.0000	2.3993	32	0.7878	0.1665	0.2579	1.7421
10	1.4093	0.2979	0.0000	2.3275	33	0.7758	0.1640	0.2692	1.7308
11	1.3437	0.2840	0.0000	2.2657	34	0.7643	0.1615	0.2801	1.7199
12	1.2865	0.2719	0.0000	2.2118	35	0.7533	0.1592	0.2904	1.7096
13	1.2360	0.2613	0.0000	2.1643	36	0.7428	0.1570	0.3004	1.6996
14	1.1911	0.2517	0.0000	2.1219	37	0.7327	0.1549	0.3099	1.6901
15	1.1507	0.2432	0.0000	2.0839	38	0.7229	0.1528	0.3190	1.6810
16	1.1141	0.2355	0.0000	2.0495	39	0.7136	0.1508	0.3278	1.6722
17	1.0809	0.2285	0.0000	2.0181	40	0.7046	0.1489	0.3363	1.6637
18	1.0504	0.2220	0.0106	1.9894	41	0.6960	0.1471	0.3444	1.6556
19	1.0224	0.2161	0.0369	1.9631	42	0.6877	0.1453	0.3523	1.6477
20	0.9965	0.2106	0.0613	1.9387	43	0.6796	0.1436	0.3598	1.6402
21	0.9725	0.2056	0.0839	1.9161	44	0.6718	0.1420	0.3671	1.6329
22	0.9501	0.2008	0.1050	1.8950	45	0.6643	0.1404	0.3742	1.6258
23	0.9293	0.1964	0.1247	1.8753	46	0.6571	0.1389	0.3811	1.6189
24	0.9097	0.1923	0.1431	1.8569	47	0.6501	0.1374	0.3877	1.6123
25	0.8913	0.1884	0.1604	1.8396	48	0.6432	0.1360	0.3941	1.6059
26	0.8740	0.1847	0.1767	1.8233	49	0.6366	0.1346	0.4003	1.5997
27	0.8577	0.1813	0.1921	1.8079	50	0.6303	0.1332	0.4063	1.5937

Cuadro 4.18: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.1)$, $\kappa = 0.1$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	1.3979	0.5853	0.0000	2.8297	28	0.5907	0.2474	0.2268	1.7732
6	1.2761	0.5343	0.0000	2.6703	29	0.5805	0.2431	0.2403	1.7597
7	1.1815	0.4947	0.0000	2.5464	30	0.5707	0.2390	0.2530	1.7470
8	1.1052	0.4628	0.0000	2.4465	31	0.5614	0.2351	0.2652	1.7348
9	1.0420	0.4363	0.0000	2.3638	32	0.5526	0.2314	0.2767	1.7233
10	0.9885	0.4139	0.0000	2.2938	33	0.5441	0.2278	0.2878	1.7122
11	0.9425	0.3946	0.0000	2.2336	34	0.5361	0.2245	0.2983	1.7017
12	0.9024	0.3778	0.0000	2.1811	35	0.5284	0.2212	0.3084	1.6916
13	0.8670	0.3630	0.0000	2.1347	36	0.5210	0.2181	0.3181	1.6819
14	0.8354	0.3498	0.0000	2.0935	37	0.5139	0.2152	0.3274	1.6726
15	0.8071	0.3379	0.0000	2.0564	38	0.5071	0.2123	0.3363	1.6637
16	0.7815	0.3272	0.0000	2.0228	39	0.5005	0.2096	0.3449	1.6551
17	0.7581	0.3174	0.0077	1.9923	40	0.4942	0.2070	0.3531	1.6469
18	0.7368	0.3085	0.0357	1.9643	41	0.4882	0.2044	0.3610	1.6390
19	0.7171	0.3003	0.0614	1.9386	42	0.4823	0.2020	0.3687	1.6313
20	0.6990	0.2927	0.0851	1.9149	43	0.4767	0.1996	0.3761	1.6239
21	0.6821	0.2856	0.1072	1.8928	44	0.4712	0.1973	0.3832	1.6168
22	0.6664	0.2791	0.1277	1.8723	45	0.4660	0.1951	0.3901	1.6099
23	0.6518	0.2729	0.1469	1.8531	46	0.4609	0.1930	0.3968	1.6032
24	0.6381	0.2672	0.1649	1.8351	47	0.4560	0.1909	0.4032	1.5968
25	0.6252	0.2618	0.1817	1.8183	48	0.4512	0.1889	0.4095	1.5905
26	0.6130	0.2567	0.1976	1.8024	49	0.4466	0.1870	0.4155	1.5845
27	0.6016	0.2519	0.2126	1.7874	50	0.4421	0.1851	0.4214	1.5786

Cuadro 4.19: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.2)$, $\kappa = 0.2$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	1.1303	0.7000	0.0000	2.7692	28	0.4777	0.2958	0.2524	1.7476
6	1.0319	0.6390	0.0000	2.6151	29	0.4693	0.2907	0.2654	1.7346
7	0.9553	0.5916	0.0000	2.4953	30	0.4615	0.2858	0.2777	1.7223
8	0.8936	0.5534	0.0000	2.3987	31	0.4540	0.2811	0.2895	1.7105
9	0.8425	0.5217	0.0000	2.3187	32	0.4468	0.2767	0.3006	1.6994
10	0.7993	0.4950	0.0000	2.2510	33	0.4400	0.2725	0.3113	1.6887
11	0.7621	0.4719	0.0000	2.1928	34	0.4335	0.2684	0.3215	1.6785
12	0.7296	0.4518	0.0000	2.1420	35	0.4272	0.2646	0.3313	1.6687
13	0.7010	0.4341	0.0000	2.0972	36	0.4213	0.2609	0.3406	1.6594
14	0.6755	0.4183	0.0000	2.0573	37	0.4155	0.2573	0.3496	1.6504
15	0.6526	0.4041	0.0000	2.0215	38	0.4100	0.2539	0.3582	1.6418
16	0.6319	0.3913	0.0110	1.9890	39	0.4047	0.2506	0.3665	1.6335
17	0.6130	0.3796	0.0405	1.9595	40	0.3996	0.2475	0.3745	1.6255
18	0.5957	0.3689	0.0675	1.9325	41	0.3947	0.2444	0.3822	1.6178
19	0.5799	0.3591	0.0924	1.9076	42	0.3900	0.2415	0.3896	1.6104
20	0.5652	0.3500	0.1154	1.8846	43	0.3854	0.2387	0.3967	1.6033
21	0.5516	0.3416	0.1367	1.8633	44	0.3810	0.2360	0.4036	1.5964
22	0.5389	0.3337	0.1566	1.8434	45	0.3768	0.2333	0.4103	1.5897
23	0.5270	0.3264	0.1751	1.8249	46	0.3727	0.2308	0.4167	1.5833
24	0.5159	0.3195	0.1925	1.8075	47	0.3687	0.2283	0.4229	1.5771
25	0.5055	0.3130	0.2088	1.7912	48	0.3648	0.2259	0.4290	1.5710
26	0.4957	0.3070	0.2241	1.7759	49	0.3611	0.2236	0.4348	1.5652
27	0.4864	0.3012	0.2386	1.7614	50	0.3574	0.2214	0.4405	1.5595

Cuadro 4.20: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.3)$, $\kappa = 0.3$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	0.9697	0.7865	0.0000	2.7053	28	0.4098	0.3323	0.2794	1.7206
6	0.8852	0.7179	0.0000	2.5567	29	0.4026	0.3266	0.2919	1.7081
7	0.8195	0.6647	0.0000	2.4412	30	0.3959	0.3211	0.3038	1.6962
8	0.7666	0.6218	0.0000	2.3482	31	0.3894	0.3159	0.3151	1.6849
9	0.7228	0.5862	0.0000	2.2711	32	0.3833	0.3109	0.3259	1.6741
10	0.6857	0.5561	0.0000	2.2058	33	0.3775	0.3061	0.3362	1.6638
11	0.6538	0.5302	0.0000	2.1497	34	0.3719	0.3016	0.3460	1.6540
12	0.6259	0.5077	0.0000	2.1008	35	0.3665	0.2973	0.3555	1.6445
13	0.6014	0.4877	0.0000	2.0576	36	0.3614	0.2931	0.3645	1.6355
14	0.5795	0.4700	0.0000	2.0191	37	0.3565	0.2891	0.3731	1.6269
15	0.5599	0.4541	0.0154	1.9846	38	0.3517	0.2853	0.3814	1.6186
16	0.5421	0.4396	0.0467	1.9533	39	0.3472	0.2816	0.3894	1.6106
17	0.5259	0.4265	0.0752	1.9248	40	0.3428	0.2781	0.3971	1.6029
18	0.5111	0.4145	0.1012	1.8988	41	0.3386	0.2746	0.4045	1.5955
19	0.4974	0.4034	0.1252	1.8748	42	0.3346	0.2714	0.4116	1.5884
20	0.4848	0.3932	0.1473	1.8527	43	0.3307	0.2682	0.4185	1.5815
21	0.4732	0.3838	0.1679	1.8321	44	0.3269	0.2651	0.4251	1.5749
22	0.4623	0.3749	0.1870	1.8130	45	0.3232	0.2622	0.4316	1.5684
23	0.4521	0.3667	0.2049	1.7951	46	0.3197	0.2593	0.4378	1.5622
24	0.4426	0.3590	0.2216	1.7784	47	0.3163	0.2565	0.4438	1.5562
25	0.4337	0.3517	0.2374	1.7626	48	0.3130	0.2538	0.4496	1.5504
26	0.4252	0.3449	0.2522	1.7478	49	0.3098	0.2512	0.4553	1.5447
27	0.4173	0.3384	0.2662	1.7338	50	0.3066	0.2487	0.4607	1.5393

Cuadro 4.21: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.4)$, $\kappa = 0.4$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	0.8604	0.8541	0.0000	2.6432	28	0.3636	0.3609	0.3056	1.6944
6	0.7854	0.7797	0.0000	2.5000	29	0.3572	0.3547	0.3177	1.6823
7	0.7271	0.7219	0.0000	2.3887	30	0.3512	0.3487	0.3292	1.6708
8	0.6802	0.6752	0.0000	2.2990	31	0.3455	0.3430	0.3401	1.6599
9	0.6413	0.6366	0.0000	2.2247	32	0.3401	0.3376	0.3505	1.6495
10	0.6084	0.6040	0.0000	2.1619	33	0.3349	0.3325	0.3604	1.6396
11	0.5801	0.5758	0.0000	2.1078	34	0.3299	0.3275	0.3699	1.6301
12	0.5554	0.5513	0.0000	2.0607	35	0.3252	0.3228	0.3789	1.6211
13	0.5336	0.5297	0.0000	2.0190	36	0.3206	0.3183	0.3876	1.6124
14	0.5142	0.5104	0.0180	1.9820	37	0.3163	0.3140	0.3960	1.6040
15	0.4967	0.4931	0.0513	1.9487	38	0.3121	0.3098	0.4040	1.5960
16	0.4810	0.4775	0.0814	1.9186	39	0.3081	0.3058	0.4117	1.5883
17	0.4666	0.4632	0.1089	1.8911	40	0.3042	0.3020	0.4191	1.5809
18	0.4534	0.4502	0.1340	1.8660	41	0.3005	0.2983	0.4262	1.5738
19	0.4414	0.4382	0.1571	1.8429	42	0.2969	0.2947	0.4331	1.5669
20	0.4302	0.4271	0.1784	1.8216	43	0.2934	0.2913	0.4397	1.5603
21	0.4198	0.4168	0.1982	1.8018	44	0.2900	0.2879	0.4461	1.5539
22	0.4102	0.4072	0.2167	1.7833	45	0.2868	0.2847	0.4523	1.5477
23	0.4011	0.3982	0.2339	1.7661	46	0.2837	0.2816	0.4583	1.5417
24	0.3927	0.3898	0.2500	1.7500	47	0.2806	0.2786	0.4641	1.5359
25	0.3848	0.3820	0.2652	1.7348	48	0.2777	0.2757	0.4697	1.5303
26	0.3773	0.3746	0.2794	1.7206	49	0.2748	0.2728	0.4751	1.5249
27	0.3702	0.3676	0.2929	1.7071	50	0.2721	0.2701	0.4804	1.5196

Cuadro 4.22: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.5)$, $\kappa = 0.5$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	0.6944	0.9717	0.0000	2.5088	28	0.2934	0.4106	0.3624	1.6376
6	0.6339	0.8870	0.0000	2.3773	29	0.2883	0.4035	0.3735	1.6265
7	0.5869	0.8212	0.0000	2.2752	30	0.2835	0.3967	0.3840	1.6160
8	0.5490	0.7682	0.0000	2.1928	31	0.2789	0.3902	0.3941	1.6059
9	0.5176	0.7243	0.0000	2.1246	32	0.2745	0.3841	0.4036	1.5964
10	0.4910	0.6871	0.0000	2.0669	33	0.2703	0.3782	0.4127	1.5873
11	0.4682	0.6551	0.0000	2.0172	34	0.2663	0.3726	0.4214	1.5786
12	0.4482	0.6272	0.0261	1.9739	35	0.2625	0.3673	0.4297	1.5703
13	0.4307	0.6026	0.0643	1.9357	36	0.2588	0.3621	0.4377	1.5623
14	0.4150	0.5807	0.0983	1.9017	37	0.2553	0.3572	0.4454	1.5546
15	0.4009	0.5610	0.1289	1.8711	38	0.2519	0.3525	0.4527	1.5473
16	0.3882	0.5432	0.1566	1.8434	39	0.2486	0.3479	0.4598	1.5402
17	0.3766	0.5270	0.1817	1.8183	40	0.2455	0.3435	0.4666	1.5334
18	0.3660	0.5121	0.2048	1.7952	41	0.2425	0.3393	0.4731	1.5269
19	0.3562	0.4985	0.2260	1.7740	42	0.2396	0.3353	0.4794	1.5206
20	0.3472	0.4858	0.2456	1.7544	43	0.2368	0.3313	0.4855	1.5145
21	0.3388	0.4741	0.2638	1.7362	44	0.2341	0.3276	0.4914	1.5086
22	0.3310	0.4632	0.2807	1.7193	45	0.2315	0.3239	0.4971	1.5029
23	0.3238	0.4531	0.2965	1.7035	46	0.2289	0.3204	0.5026	1.4974
24	0.3170	0.4435	0.3113	1.6887	47	0.2265	0.3169	0.5079	1.4921
25	0.3106	0.4346	0.3252	1.6748	48	0.2241	0.3136	0.5130	1.4870
26	0.3045	0.4261	0.3384	1.6616	49	0.2218	0.3104	0.5180	1.4820
27	0.2988	0.4181	0.3507	1.6493	50	0.2196	0.3073	0.5229	1.4771

Cuadro 4.23: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 0.75)$, $\kappa = 0.75$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	0.6016	1.0461	-0.4071	2.4071	28	0.2542	0.4420	0.4054	1.5946
6	0.5492	0.9549	-0.2845	2.2845	29	0.2498	0.4344	0.4157	1.5843
7	0.5084	0.8841	-0.1892	2.1892	30	0.2456	0.4271	0.4255	1.5745
8	0.4756	0.8270	-0.1124	2.1124	31	0.2416	0.4201	0.4349	1.5651
9	0.4484	0.7797	-0.0488	2.0488	32	0.2378	0.4135	0.4438	1.5562
10	0.4254	0.7397	0.0050	1.9950	33	0.2342	0.4072	0.4523	1.5477
11	0.4056	0.7053	0.0513	1.9487	34	0.2307	0.4012	0.4604	1.5396
12	0.3883	0.6752	0.0917	1.9083	35	0.2274	0.3954	0.4682	1.5318
13	0.3731	0.6487	0.1273	1.8727	36	0.2242	0.3898	0.4756	1.5244
14	0.3595	0.6251	0.1591	1.8409	37	0.2211	0.3845	0.4827	1.5173
15	0.3473	0.6040	0.1876	1.8124	38	0.2182	0.3795	0.4896	1.5104
16	0.3363	0.5848	0.2134	1.7866	39	0.2154	0.3746	0.4962	1.5038
17	0.3262	0.5673	0.2369	1.7631	40	0.2127	0.3698	0.5025	1.4975
18	0.3171	0.5513	0.2584	1.7416	41	0.2101	0.3653	0.5086	1.4914
19	0.3086	0.5366	0.2782	1.7218	42	0.2076	0.3609	0.5145	1.4855
20	0.3008	0.5230	0.2964	1.7036	43	0.2051	0.3567	0.5202	1.4798
21	0.2935	0.5104	0.3134	1.6866	44	0.2028	0.3526	0.5257	1.4743
22	0.2868	0.4987	0.3292	1.6708	45	0.2005	0.3487	0.5310	1.4690
23	0.2805	0.4877	0.3439	1.6561	46	0.1983	0.3449	0.5361	1.4639
24	0.2746	0.4775	0.3577	1.6423	47	0.1962	0.3412	0.5410	1.4590
25	0.2690	0.4678	0.3707	1.6293	48	0.1942	0.3376	0.5459	1.4541
26	0.2638	0.4587	0.3829	1.6171	49	0.1922	0.3342	0.5505	1.4495
27	0.2589	0.4502	0.3945	1.6055	50	0.1902	0.3308	0.5550	1.4450

Cuadro 4.24: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 1)$, $\kappa = 1$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	0.3672	1.2754	0.0000	2.0473	28	0.1552	0.5389	0.5574	1.4426
6	0.3352	1.1643	0.0439	1.9561	29	0.1525	0.5296	0.5651	1.4349
7	0.3104	1.0779	0.1149	1.8851	30	0.1499	0.5207	0.5724	1.4276
8	0.2903	1.0083	0.1720	1.8280	31	0.1475	0.5122	0.5794	1.4206
9	0.2737	0.9506	0.2194	1.7806	32	0.1452	0.5041	0.5860	1.4140
10	0.2597	0.9018	0.2594	1.7406	33	0.1429	0.4964	0.5923	1.4077
11	0.2476	0.8599	0.2939	1.7061	34	0.1408	0.4891	0.5984	1.4016
12	0.2371	0.8233	0.3240	1.6760	35	0.1388	0.4820	0.6042	1.3958
13	0.2278	0.7910	0.3505	1.6495	36	0.1369	0.4753	0.6097	1.3903
14	0.2195	0.7622	0.3741	1.6259	37	0.1350	0.4688	0.6150	1.3850
15	0.2120	0.7363	0.3953	1.6047	38	0.1332	0.4626	0.6201	1.3799
16	0.2053	0.7130	0.4145	1.5855	39	0.1315	0.4567	0.6250	1.3750
17	0.1992	0.6917	0.4320	1.5680	40	0.1298	0.4509	0.6297	1.3703
18	0.1936	0.6722	0.4480	1.5520	41	0.1282	0.4454	0.6343	1.3657
19	0.1884	0.6543	0.4627	1.5373	42	0.1267	0.4400	0.6386	1.3614
20	0.1836	0.6377	0.4763	1.5237	43	0.1252	0.4349	0.6429	1.3571
21	0.1792	0.6223	0.4890	1.5110	44	0.1238	0.4299	0.6470	1.3530
22	0.1751	0.6080	0.5007	1.4993	45	0.1224	0.4251	0.6509	1.3491
23	0.1712	0.5946	0.5117	1.4883	46	0.1211	0.4205	0.6547	1.3453
24	0.1676	0.5821	0.5220	1.4780	47	0.1198	0.4160	0.6584	1.3416
25	0.1642	0.5704	0.5316	1.4684	48	0.1185	0.4116	0.6620	1.3380
26	0.1610	0.5593	0.5407	1.4593	49	0.1173	0.4074	0.6655	1.3346
27	0.1580	0.5488	0.5493	1.4507	50	0.1161	0.4033	0.6688	1.3312

Cuadro 4.25: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 5)$, $\kappa = 5$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	0.3409	1.3083	0.0028	1.9972	28	0.1440	0.5528	0.5786	1.4214
6	0.3112	1.1943	0.0897	1.9103	29	0.1415	0.5432	0.5859	1.4141
7	0.2881	1.1057	0.1572	1.8428	30	0.1392	0.5341	0.5929	1.4071
8	0.2695	1.0343	0.2117	1.7883	31	0.1369	0.5254	0.5995	1.4005
9	0.2541	0.9751	0.2567	1.7433	32	0.1347	0.5171	0.6058	1.3942
10	0.2410	0.9251	0.2949	1.7051	33	0.1327	0.5092	0.6118	1.3882
11	0.2298	0.8820	0.3277	1.6723	34	0.1307	0.5017	0.6176	1.3824
12	0.2200	0.8445	0.3563	1.6437	35	0.1288	0.4945	0.6231	1.3769
13	0.2114	0.8114	0.3816	1.6184	36	0.1270	0.4876	0.6284	1.3716
14	0.2037	0.7818	0.4041	1.5959	37	0.1253	0.4809	0.6334	1.3666
15	0.1968	0.7553	0.4243	1.5757	38	0.1236	0.4746	0.6383	1.3617
16	0.1906	0.7313	0.4426	1.5574	39	0.1221	0.4684	0.6429	1.3571
17	0.1849	0.7095	0.4592	1.5408	40	0.1205	0.4625	0.6474	1.3526
18	0.1797	0.6895	0.4744	1.5256	41	0.1190	0.4569	0.6518	1.3482
19	0.1749	0.6711	0.4885	1.5115	42	0.1176	0.4514	0.6559	1.3441
20	0.1704	0.6541	0.5014	1.4986	43	0.1162	0.4461	0.6600	1.3400
21	0.1663	0.6384	0.5134	1.4866	44	0.1149	0.4410	0.6638	1.3362
22	0.1625	0.6237	0.5246	1.4754	45	0.1136	0.4361	0.6676	1.3324
23	0.1589	0.6100	0.5351	1.4649	46	0.1124	0.4313	0.6712	1.3288
24	0.1556	0.5971	0.5448	1.4552	47	0.1112	0.4267	0.6748	1.3252
25	0.1524	0.5851	0.5540	1.4460	48	0.1100	0.4222	0.6782	1.3218
26	0.1495	0.5737	0.5627	1.4373	49	0.1089	0.4179	0.6815	1.3185
27	0.1467	0.5630	0.5709	1.4291	50	0.1078	0.4137	0.6847	1.3153

Cuadro 4.26: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 10)$, $\kappa = 10$.

n	d_3	A_2	D_3	D_4	n	d_3	A_2	D_3	D_4
5	0.3210	1.3349	0.0418	1.9582	28	0.1357	0.5641	0.5951	1.4049
6	0.2930	1.2186	0.1253	1.8747	29	0.1333	0.5543	0.6021	1.3979
7	0.2713	1.1282	0.1902	1.8098	30	0.1311	0.5450	0.6088	1.3912
8	0.2538	1.0554	0.2425	1.7575	31	0.1289	0.5361	0.6152	1.3848
9	0.2393	0.9950	0.2858	1.7142	32	0.1269	0.5277	0.6212	1.3788
10	0.2270	0.9439	0.3224	1.6776	33	0.1250	0.5196	0.6270	1.3730
11	0.2164	0.9000	0.3540	1.6460	34	0.1231	0.5119	0.6325	1.3675
12	0.2072	0.8617	0.3815	1.6185	35	0.1213	0.5046	0.6378	1.3622
13	0.1991	0.8279	0.4057	1.5943	36	0.1196	0.4975	0.6429	1.3571
14	0.1918	0.7978	0.4274	1.5726	37	0.1180	0.4907	0.6478	1.3522
15	0.1853	0.7707	0.4468	1.5532	38	0.1164	0.4842	0.6524	1.3476
16	0.1794	0.7463	0.4643	1.5357	39	0.1149	0.4780	0.6569	1.3431
17	0.1741	0.7240	0.4803	1.5197	40	0.1135	0.4720	0.6612	1.3388
18	0.1692	0.7036	0.4950	1.5050	41	0.1121	0.4662	0.6654	1.3346
19	0.1647	0.6848	0.5084	1.4916	42	0.1108	0.4606	0.6694	1.3306
20	0.1605	0.6675	0.5209	1.4791	43	0.1095	0.4552	0.6733	1.3267
21	0.1566	0.6514	0.5324	1.4676	44	0.1082	0.4500	0.6770	1.3230
22	0.1530	0.6364	0.5432	1.4568	45	0.1070	0.4450	0.6806	1.3194
23	0.1497	0.6224	0.5532	1.4468	46	0.1058	0.4401	0.6841	1.3159
24	0.1465	0.6093	0.5626	1.4374	47	0.1047	0.4354	0.6875	1.3125
25	0.1436	0.5970	0.5715	1.4285	48	0.1036	0.4308	0.6907	1.3093
26	0.1408	0.5854	0.5798	1.4202	49	0.1025	0.4264	0.6939	1.3061
27	0.1381	0.5745	0.5876	1.4124	50	0.1015	0.4221	0.6970	1.3030

Cuadro 4.27: Valores para los límites de las cartas de control \bar{X} y S para la $LGG(0, 1, 50)$, $\kappa = 50$.

En estas tablas se puede apreciar que cuando el valor del parámetro κ crece los valores de d_3 , A_2 , D_3 y D_4 se asemejan al caso de la normal. Cuando el valor de κ disminuye dichos valores aumentan, aumentando los límites de control.

NOTA 4.4. En las tablas 4.18 - 4.27 se muestran los resultados para los valores de $\kappa = 0.1, 0.2, 0.3, 0.4, 0.5, 0.75, 1, 5, 10$ y 50 , para otros valores se pueden interpolar los valores anteriores o se recomienda utilizar la función en el proyecto R, mostrada al final en el Anexo, en la que con base en los valores de κ y n da como respuesta d_3 , A_2 , D_3 y D_4 para estos parámetros.

4.7.3. Metodología para las cartas de control \bar{X} y S , de una distribución $LGG(\xi, \sigma, \kappa)$

Las aplicaciones de las cartas de control \bar{X} - S para una distribución log-gamma generalizada se pueden hacer, siguiendo la metodología.

- 1).- Con los datos x_1, x_2, \dots, x_n y las pruebas de bondad de ajuste propuestas en la sección 4.6, verificar si los datos provienen de una distribución log-gamma generalizada. En ocasiones puede ser necesaria una transformación:
 - Si los datos son sesgados a la derecha, se propone la transformación $Y = -X$.
 - En caso de ser del tipo gama generalizada, se propone la transformación $Y = \log(X)$.
- 2).- En caso de que los datos tengan un comportamiento log-gamma generalizada se estima su parámetro de forma κ .
- 3).- Se calculan los límites de control \bar{X} - S dados en (4.7.1) y (4.7.2), para los valores de A_2 , D_3 y D_4 se pueden utilizar las tablas propuestas 4.18-4.27, pero si el valor de κ o n no está en las tablas se puede interpolar o utilizar la función programada en el proyecto R que se encuentra al final del anexo.
- 4).- Se calculan las medias y desviaciones estándar por cada muestra tomada del proceso y trazan sus gráficas con los límites de control y el valor medio, para determinar si el proceso está bajo control.

4.7.4. Problema

Se consideran 10 muestras de tamaño 5 de un proceso de fabricación de cierto tipo de fibra de carbono, para estudiar su distribución de resistencia a la rotura. Este tipo de fibra de carbono es producido para la fabricación de materiales compuestos, que necesitan las fibras con una resistencia a la rotura mayor a 2.9992 (giga pascal) con una probabilidad del 99%. Se toman 10 muestras de $n = 5$ fibras, cada 50 mm de longitud y se mide su resistencia a la ruptura de cada fibra. Los resultados se muestran en la tabla 4.28, obtenga las cartas de control para este proceso de producción.

Los datos originales del problema están en el artículo de Padgett and Spurrier (1990) y el problema en el artículo de Pasquale, Giuliana and Sung (2008)

i -ésima muestra	Resistencia a la ruptura en Giga-Pascal				
1	3.70	2.74	2.73	2.50	3.60
2	3.11	3.27	2.87	1.47	3.11
3	4.42	2.41	3.19	3.22	1.69
4	3.28	3.09	1.87	3.15	4.90
5	3.75	2.43	2.95	2.97	3.39
6	2.96	2.53	2.67	2.93	3.22
7	3.39	2.81	4.20	3.33	2.55
8	3.31	3.31	2.85	2.56	3.56
9	3.15	2.35	2.55	2.59	2.38
10	2.81	2.77	2.17	2.83	1.92

Cuadro 4.28: Valores de las 10 muestras para medir la resistencia a la ruptura de las fibras de carbono.

Solución Se seguira la metodología propuesta.

- 1).- Al graficar los datos se observa que son sesgados a la derecha, luego se usa la transformación $Y = -X$, con los datos transformados se realiza la prueba de bondad de ajuste propuesta en la sección 4.6. Obteniendo al 5% de significancia que no existe evidencia para rechazar que los datos transformados tengan un comportamiento log-gamma generalizado, ver figura 4.1.

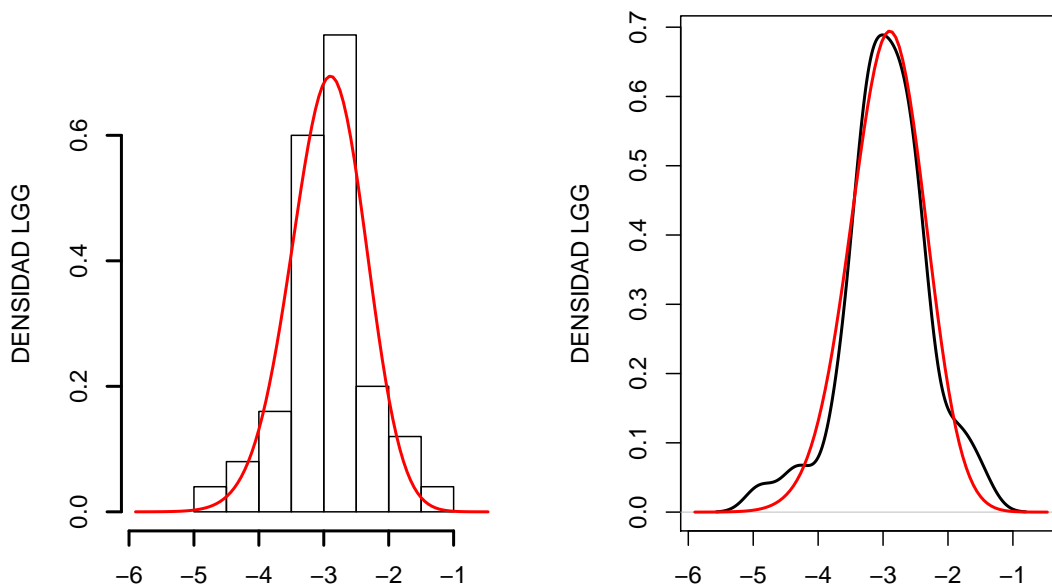


Figura 4.1: La figura de la izquierda muestra el histograma de los datos transformados y la distribución log-gamma que mejor ajusta; a la derecha se muestran en lugar del histograma su función de densidad empírica de los datos transformados.

- 2).- Con la prueba de bondad de ajuste propuesta se obtiene también un valor para estimar a κ , obteniendo su estimación puntual $\hat{\kappa} = 13.04$.
- 3).- Como este valor no está en la tablas propuestas, se puede hacer una aproximación o utilizar la función al final del anexo. Realizando los cálculos con la función para $n = 5$ y $\kappa = 13.04$ se obtiene: $A_2 = 1.316$, $D_3 = 0.0143$ y $D_4 = 1.9857$.

i -ésima	Resistencia a la ruptura en Giga-Pascal					\bar{x}_i	s_i
1	-3.70	-2.74	-2.73	-2.50	-3.60	-3.05	0.55
2	-3.11	-3.27	-2.87	-1.47	-3.11	-2.77	0.74
3	-4.42	-2.41	-3.19	-3.22	-1.69	-2.99	1.02
4	-3.28	-3.09	-1.87	-3.15	-4.90	-3.26	1.08
5	-3.75	-2.43	-2.95	-2.97	-3.39	-3.10	0.5
6	-2.96	-2.53	-2.67	-2.93	-3.22	-2.86	0.27
7	-3.39	-2.81	-4.20	-3.33	-2.55	-3.26	0.63
8	-3.31	-3.31	-2.85	-2.56	-3.56	-3.12	0.4
9	-3.15	-2.35	-2.55	-2.59	-2.38	-2.60	0.32
10	-2.81	-2.77	-2.17	-2.83	-1.92	-2.50	0.43

Cuadro 4.29: Valores de las 10 muestras transformadas, medias y desviaciones estándar de la tabla 4.28.

4).- Falta calcular las medias y desviaciones estándar para cada muestra, ver tabla 4.29.

Para trazar las gráficas de las cartas de control se calcula $\bar{\bar{x}} = -2.95$ y $\bar{s} = 0.59$, luego los límites de control para \bar{X} y S estarán dados por (4.7.12) y (4.7.13)

$$LIC_{\bar{X}} = \bar{\bar{X}} - A_2\bar{S} = -3.73 \quad LIC_S = \bar{S}D_3 = 0.0085 \quad (4.7.12)$$

$$LSC_{\bar{X}} = \bar{\bar{X}} + A_2\bar{S} = -2.17 \quad LSC_S = \bar{S}D_4 = 1.1807 \quad (4.7.13)$$

Las gráficas se muestran en las figuras 4.2 y 4.3. Finalmente, con los resultados obtenidos se concluye que el proceso está bajo control.

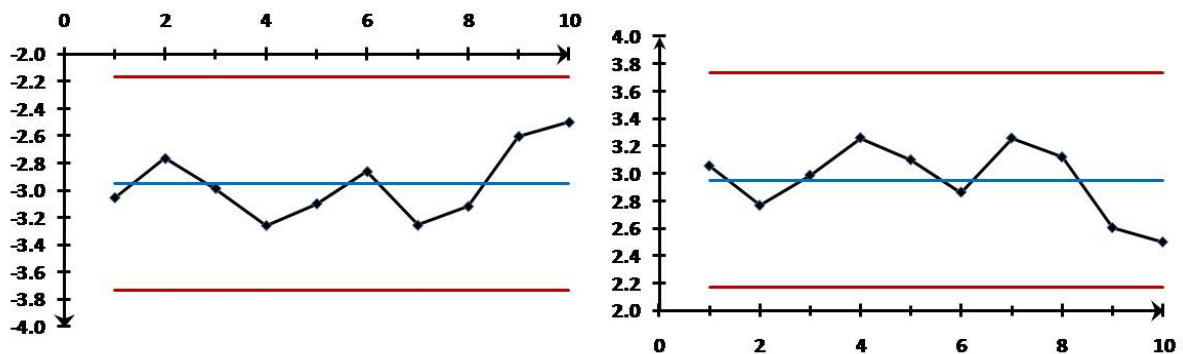


Figura 4.2: La figura de la izquierda muestra las cartas de control \bar{X} para los datos transformados y la gráfica de la derecha para los datos originales.

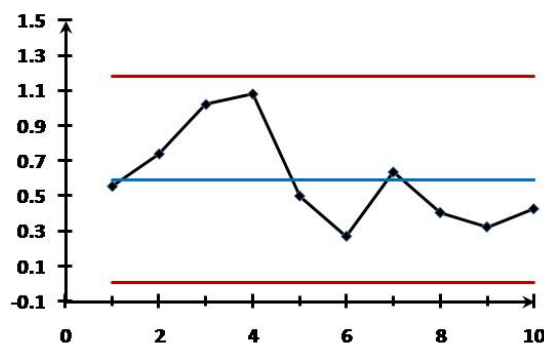


Figura 4.3: Carta de control S para los datos transformados u originales.

4.7.5. Ajuste de la curva

En el ajuste de los datos se usó la transformación $Y = -X$, con los datos transformados la prueba de bondad de ajuste bootstrap propuesta con el coeficiente de correlación como estadística de prueba se obtuvo al 5 % de significancia que no existe evidencia para rechazar que los datos transformados tengan un comportamiento log-gamma generalizado.

Luego, con la prueba de bondad de ajuste propuesta se obtuvo un valor para estimar a κ , resultando su estimación puntual $\hat{\kappa} = 13.040208$. Con este valor y las ecuaciones (2.2.6) y (2.2.7)

$$\xi = \bar{x} + \sigma\sqrt{\kappa_0}(\log(\kappa_0) - \psi(\kappa_0)).$$

$$\sum_{i=1}^n \exp\left(\frac{x_i - \bar{x}}{\sigma\kappa_0}\right) - n\kappa_0 e^{-\psi(\kappa_0)} = 0.$$

se obtienen los valores para $\hat{\xi} = -2.863490$ y $\hat{\sigma} = 0.618434$.

Como los datos se transformaron con $Y = -X$, entonces la densidad de los datos originales tiene los parámetros: forma $\hat{\kappa} = 13.040208$, localidad $\hat{\xi} = 2.863490$ y escala $\hat{\sigma} = 0.618434$, con función de densidad (4.7.14)

$$\begin{aligned} f(x) &= \frac{13.04^{12.54}}{0.6184\Gamma(13.04)} \exp\left(-\left(\frac{x - 2.8635}{0.6184}\right) \sqrt{13.04} - 13.04 e^{-\left(\frac{x - 2.8635}{0.6184}\right)/\sqrt{13.04}}\right) \\ &= 295165.6 \exp\left(-\left(\frac{x - 2.8635}{0.1713}\right) - 13.04 e^{-\left(\frac{x - 2.8635}{2.2331}\right)}\right), \quad -\infty < x < \infty. \end{aligned} \quad (4.7.14)$$

El ajuste para los datos transformados se puede ver en la figura 4.1.

Conclusiones

C.1. Desarrollo analítico

Después de analizar los resultados analíticos del trabajo se puede concluir que la revisión de las propiedades, tanto de la función gamma, como de las distribuciones gamma y log-gamma generalizada serán de bastante ayuda para el análisis de este tipo de distribuciones, puesto que en ellas están contenidas las principales propiedades para el análisis de resultados extremos de las distribuciones, gama y log-gamma generalizadas estándar en ambos sentidos, cuando el parámetro de forma tiende a $+\infty$ y en el caso contrario cuando tiende a 0^+ .

En la primera situación cuando el parámetro $\kappa \rightarrow +\infty$, se tiene el resultado conocido de que la distribución gamma generalizada tiene un comportamiento log-normal y similarmente la distribución log-gamma generalizada tiene un comportamiento normal para $\kappa \rightarrow +\infty$. Esto se puede ver reflejado en las potencias de la prueba cuando se contrastan con la distribución normal, puesto que sus potencias en general resultaron pequeñas, indicando que la prueba no detecta diferencias entre la distribución normal y la log-gamma generalizada.

Mientras que en el otro caso extremo, cuando el parámetro $\kappa \rightarrow 0^+$, la distribución toma una forma muy similar a una recta, pero con valores de la función obviamente muy pequeños, lo que dificulta la búsqueda de cuantiles. Luego, por estas razones se explica el porqué las potencias resultaron en algunos casos inesperadas. Por ejemplo en una distribución sesgada a la derecha se esperarían potencias de 1, pero sin embargo en algunos casos con tamaños de muestra pequeños las potencias eran inferiores a 1.

Por otro lado, los resultados analíticos obtenidos sobre los momentos de una distribución log-gamma generalizada, aunque en apariencia siguen siendo laboriosos, el trabajo de obtención de éstos no es nada comparable que sin dicho resultado. Similarmente el resultado para la distribución de un cociente o de cualquier otra función de dos momentos centrales muestrales, simplifica en mucho los estudios asintóticos, en particular de sumas o promedios de variables aleatorias.

Con los resultados asintóticos para los estimadores y la acotación de los momentos fue posible demostrar que ambos estimadores utilizados son asintóticamente insesgados y consistentes en error cuadrado medio. Resultado que posteriormente fue comprobado por simulación.

Finalmente de los resultados analíticos se concluye que no era posible obtener una prueba

para cualquier valor del parámetro κ , porque en ellos se muestra que existe dependencia con respecto al parámetro, luego no era posible establecer una prueba para cualquier valor de κ , situación que orilló a desarrollar el trabajo de simulación con respecto a una prueba bootstrap.

C.2. Desarrollo por simulación

Después de analizar los resultados obtenidos por simulación se concluye que cualquiera de los dos estimadores, el de máxima verosimilitud o el de momentos, son buenos estimadores para el parámetro de forma κ , con ambos se obtuvieron buenas estimaciones y tamaños de prueba bastante aceptables, en general por abajo del valor α . En el caso de la potencia la prueba basada en el estimador de momentos dio mejores resultados que la prueba basada en el EMV, en ambos casos las potencias fueron muy grandes para las alternativas asimétricas y en el caso de alternativas simétricas para tamaños de muestra mayor a 50 las potencias fueron también grandes, como era de esperarse las potencias fueron pequeñas para el caso de la distribución normal como alternativa.

Bibliografía

B.1. Referencias

- [1] Bain J. Lee, Weeks L. David. (1965) Tolerance Limits for the Generalized Gamma Distribution. *Journal of the American Statistical Association*, Vol. 60, No. 312, pp. 1142-1152.
- [2] Block, H. W., and Rao, B. Raja (1974). Some generalized distributions based on Stacy's generalized gamma distribution. *Scandinavian Actuarial Journal*, 185-189.
- [3] Boerrigter, D. G. (1998). Parameter estimation of the mixed generalized gamma distribution using maximum likelihood estimation and minimum distance estimation. Thesis. Faculty of the Graduate School of Engineering of the Air Force Institute of Technology, USA. 123p.
- [4] Bondesson, L. (1992). *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*, New York: Springer-Verlag.
- [5] Bradley M. Bell. (1988). Generalized gamma parameter estimation and moment evaluation. *Communications in Statistics - Theory and Methods*, Volume 17, Issue 2 , pages 507 - 517
- [6] Cantú S. M., Villaseñor A.J.A. and Barry C.A. (2001). Modeling the Lifetime of longitudinal elements. *Communications in Statistics-Simulation and Computation*, 34 (3): 717-741.
- [7] Coale A. J. and McNeil D.R. (1972). The Distribution by Age of the Frequency of the First Marriage in a Female Cohort. *Journal of the American Statistical Association*, 67 (340): 743-749.
- [8] DiCiccio, T.J. (1987). Approximate Inference for the Generalized Gamma Distribution. *Technometrics*, 29 (1): 33-40.

-
- [9] Efthymios, G. Tsionas. (2005). Exact inference in four-parameter generalized gamma distributions. *Communications in Statistics-Theory and Methods*, Publisher Taylor Francis, 30-4:747-756.
- [10] Gavin E. Crooks (2007). The Amoroso Distribution. *Tech. Note* 1 (3), 1-6
- [11] Gomes, O., Combes, C. Dussauchoy A. (2008). Parameter estimation of the generalized gamma distribution. *Mathematics and Computers in Simulation* 79: 955-963
- [12] Hager, H. W. (1970). *Statistical Inference for the Generalized Gamma Distribution*, Ph.D. dissertation, Department of Statistics, Rolla: University of Missouri.
- [13] Hager, H.W.; Bain J. Lee. (1970). Inferential procedures for the generalized gamma distribution. *Journal of the American Statistical Association*, 65, 1601-1609.
- [14] Hager, H.W.; Bain J. Lee., Antle, C.E. (1971). Reliability estimation for the generalized gamma distribution and robustness of the Weibull model. *Technometrics*, 13:547-557.
- [15] Harter, H.L. (1967). Maximum likelihood estimation of the parameters of a fourparameter generalized gamma population from complete and censored samples. *Technometrics*. 9:159-165.
- [16] Jakuszenkova, H. (1973). On some property of the generalized gamma-distribution. *Prace Matematyczne*, 17, 237-243.
- [17] Khan, A. H., and Khan, R. U. (1983). Recurrence relation between the moments of order statistics from generalized gamma distribution. *Journal of Statistical Research (Bangladesh)*, 17, 75-82.
- [18] Klebaner, F. C. (1989). Stochastic difference equations and generalized gamma distribution. *Annals of Probability*, 17, 178-188.
- [19] Kleiber C., Kotz S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley-Interscience, ISBN 0-471-15064-9.
- [20] Lawless, Jerald F. (2003). *Statistical Models and methods for lifetime data 2nd*. Ed. John Wiley Sons, Inc., Hoboken, New Jersey. 637p.
- [21] Lawless, Jerald F. (1980). Inference in the generalized gamma and log-gamma distribution. *Technometrics*. 22: 409-419.
- [22] Lee, M. -L. T., and Gross, A. J. (1989). Properties of conditionally independent generalized gamma distributions. *Probability in the Engineering and Informational Sciences*, 3, 289-297.
- [23] Lefante, J. J., Jr., and Turner, M. E., Jr. (1985). The average likelihood and a fiducial approximation: One parameter members of the generalized gamma distributions. *Communications in Statistics-Theory and Methods*, 14, 419-436.

-
- [24] Lienhard, H. J., and Meyer, P. L. (1967). A physical basis for the generalized gamma distribution. *Quarterly of Applied Mathematics*, 25,330-334.
- [25] Malik, H. J. (1968a). Exact distribution of the product of independent generalized gamma variables with the same shape parameter. *Annals of Mathematical Statistics*, 39, 1751-1752.
- [26] Malik, H. J. (1968b). A characterization of the generalized gamma distribution. *Metron*, 27(3-4), 41-45.
- [27] Malik, H. J. (1976). Distribution of a linear function and the ratio of two independent linear functions of independent generalized gamma variables. *Naval Research Logistics Quarterly*, 23, 339-343.
- [28] Meeker Q.W. and Escobar A.L. (1998). *Statistical Methods for Reliability* Ed. John Wiley Sons, Inc., USA-Canada 680p.
- [29] Mees, E., and Gerard, G. (1984). The generalized gamma distribution: A model for seed germination curves. *Biometrie-Praximetrie*, 24, 101-115.
- [30] Montgomery, Douglas (2006). *Introduction to stadistical quality control 3a ed.* Ed. John Wiley Sons, Inc., New York, Chichester, Brisbane, Singapore, Toronto and Weinheim. 797p.
- [31] Padgett, W.J. and Spurrier J.D. (1990). Shewhart-type charts for percentiles of strength distribution. *J.Quality Technol.* 22(4), 283-288.
- [32] Pasquale Erto, Giuliana Pallotta and Sung H. Park. (2008). An example of data technology product: A control chart for weibull processes. *International Statistical Review*, 76, 2, 157-166.
- [33] Pereyra, Bragancia B. (1998). Structural Inference for the generalized gamma distribution. *Mat. Estat.*, Sao Paulo, 16:81-89.
- [34] Pérez Rodríguez Paulino. (2008). *Algunos aspectos de inferencia estadística en la distribución normal asimétrica*, Tesis, Colegio de Postgraduados, México, 122p.
- [35] Podolski, H. (1972). The distribution of a product of n independent random variables with generalized gamma distribution. *Demonstratio Mathematica*, 4, 119-123.
- [36] Rao, A.V.; Kantam, R.R; Narasimham, V.L. (1991). Linear estimation of location and scale parameters in the generalized gamma distribution. *Communications in Statistics-Theory and Methods*. 20:3823-3848.
- [37] Roy, D. (1984). A characterization of the generalized gamma distribution, *Calcutta Statistical Association Bulletin*, 33, 137-141.

-
- [38] Ryuichi Kaneko (2003). Elaborate of the Coale-McNeil Nuptiality Model as the Generalized Log Gamma Distribution: A New Identity and Empirical Enhancements, *Demographic Research-Max-Planck-Gesellschaft*, 9 (Article 10), 137-141.
- [39] Sen, P. K. y Singer, J. M. (1993). *Large Sample Methods in Statistics: An Introduction with Applications*. Chapman Hall, New York.
- [40] Sheldon M. Ross. (2006). *Simulation*, Fourth Edition. ELSEVIER.
- [41] Sreekumar, N.V., Thomas, P.Y. (2007). Estimation of the parameters of log-gamma distribution using order statistics. *Springer-Verlang, Metrika*. 66:115-127.
- [42] Stacy, E.W. (1962). A generalization of the gamma distribution. *Annals of Mathematical Statistics*, 33: 1187-1192
- [43] Stacy, E.W.; Mihran, G.A. (1965). Parameter estimation for a generalized gamma distributions. *Technometrics*, vol. 7, No. 3. 347-358
- [44] Taguchi, T. (1980). On an interpretation and an estimation of shape parameters of the generalized gamma distribution. *Metron*, 38, 27-40
- [45] Young DH, Bakir ST. (1987). Bias correction for a generalized log-gamma regression model. *Technometrics* 29:183-191.
- [46] Wingo, D. R. (1987a). Computing globally optimal maximum likelihood estimates of generalized gamma distribution parameters-Some new numerical approaches and analytical results. *Nineteenth Symposium on Interface of Computer Science and Statistics*.
- [47] Wingo, D. R. (1987b). Computing maximum-likelihood parameter estimates of the generalized gamma distribution by numerical root isolation. *IEEE Transactions on Reliability*, R-36, 586-590.
- [48] Wong J. Y. (1993). Simultaneously estimating the three parameters of the generalized gamma distribution. *Microelectronics and reliability* (Elsevier, Oxford) ISSN 0026-2714 CODEN MCRLAS. 33, n15, pp. 2225-2232.
- [49] Yehia, A. Y., and Ahmed, A. N. (1987). A characterization of the generalized gamma distribution using residual moments. *Department of Mathematics*, Cairo University.

Anexos

En esta parte se revisarán los programas para la elaboración de las gráficas y cálculos de simulación desarrollados durante el trabajo.

A.1. Gráficas del Capítulo 1

En esta sección se mostrarán los programas para trazar las gráficas del capítulo 1.

```
##### PARA LA GRÁFICA 1.2 #####
densidad1 <- function()
{
  x <- rep(0,411); y <- (-10:400)/1000
  plot(x,y,type = "l",col =1, lty = 1,lwd = 1,main=" ", xlab="", ylab=" ",xlim=c(-10,3),
        ylim=c(-0.01,0.40))
  auxx <- function(x) { 0*x }
  plot(auxx,type = "l",col =1, lty = 1,lwd = 1,main=" ", xlab="", ylab=" ",xlim=c(-10,3),
        ylim=c(-0.01,0.40),add=TRUE)
  aux <- function(x,k){ k^(k-0.5)*exp(x*sqrt(k)-k*exp(x/sqrt(k)))/gamma(k)}
  auxk1 <- (1:10)/100;
  for(i in 1:10)
  {
    auxiliar1 <- function(x){aux(x,k=auxk1[i])}
    plot(auxiliar1,type = "l",col = i,lty = 1,lwd = 1.5,main=" ", xlab="", ylab=" ",
          xlim=c(-10,3), ylim=c(-0.01,0.40),add=TRUE)
  }
}
densidad1()
##### PARA LA GRÁFICA 1.3 #####
densidad2 <- function()
{
  x <- rep(0,411); y <- (-10:400)/1000
  plot(x,y,type = "l",col =1, lty = 1,lwd = 1,main=" ", xlab="", ylab=" ",xlim=c(-10,3),
        ylim=c(-0.01,0.40))
  auxx <- function(x) { 0*x }
  plot(auxx,type = "l",col =1, lty = 1,lwd = 1,main=" ", xlab="", ylab=" ",xlim=c(-10,3),
        ylim=c(-0.01,0.40),add=TRUE)
  aux <- function(x,k){ k^(k-0.5)*exp(x*sqrt(k)-k*exp(x/sqrt(k)))/gamma(k)}
  auxk1 <- (1:10)/100; auxk2 <- (1:10)/10
  for(i in 1:10)
  {
    auxiliar1 <- function(x){aux(x,k=auxk1[i])}
```



```

plot(auxiliar1,type = "l",col = i,lty = 1,lwd = 1.5,main=" ", xlab="", ylab=" ",
      xlim=c(-10,3), ylim=c(-0.01,0.40),add=TRUE)
auxiliar2 <- function(x){aux(x,k=auxk2[i])}
plot(auxiliar2,type = "l",col = i,lty = 1,lwd = 1.5,main=" ", xlab="", ylab=" ",
      xlim=c(-10,3), ylim=c(-0.01,0.40),add=TRUE)
}
}
densidad2()
##### PARA LA GRÁFICA 1.4 #####
densidad3 <- function()
{
  x <- rep(0,411); y <- (-10:400)/1000
  plot(x,y,type = "l",col =1, lty = 1,lwd = 1,main=" ", xlab="", ylab=" ",xlim=c(-10,3),
        ylim=c(-0.01,0.40))
  auxx <- function(x) { 0*x }
  plot(auxx,type = "l",col =1, lty = 1,lwd = 1,main=" ", xlab="", ylab=" ",xlim=c(-10,3),
        ylim=c(-0.01,0.40),add=TRUE)
  aux <- function(x,k){ k^(k-0.5)*exp(x*sqrt(k)-k*exp(x/sqrt(k)))/gamma(k)}
  auxk1 <- (1:10)/100; auxk2 <- (1:10)/10; auxk3 <- (1:10)*10
  for(i in 1:10)
  {
    auxiliar1 <- function(x){aux(x,k=auxk1[i])}
    plot(auxiliar1,type = "l",col = i,lty = 1,lwd = 1.5,main=" ", xlab="", ylab=" ",
          xlim=c(-10,3), ylim=c(-0.01,0.40),add=TRUE)
    auxiliar2 <- function(x){aux(x,k=auxk2[i])}
    plot(auxiliar2,type = "l",col = i,lty = 1,lwd = 1.5,main=" ", xlab="", ylab=" ",
          xlim=c(-10,3), ylim=c(-0.01,0.40),add=TRUE)
    auxiliar3 <- function(x){aux(x,k=auxk3[i])}
    plot(auxiliar3,type = "l",col = i,lty = 1,lwd = 1.5,main=" ", xlab="", ylab=" ",
          xlim=c(-10,3), ylim=c(-0.01,0.40),add=TRUE)
  }
}
densidad3()
##### PARA LA GRÁFICA 1.5 #####
par(mfrow = c(1, 2))
gammagen1 <- function()
{
  x <- rep(0,1011); y <- (-10:1000)/1000
  plot(x,y,type = "l",col =1, lty = 1,lwd = 1,main=" ", xlab="", ylab=" ",xlim=c(-0.5,6),
        ylim=c(-0.01,1))
  auxx <- function(x) { 0*x }
  plot(auxx,type = "l",col =1, lty = 1,lwd = 1,main=" ", xlab="", ylab=" ",xlim=c(-0.5,6),
        ylim=c(-0.01,1),add=TRUE)
  aux <- function(x,k){ k^(k-0.5)*x^(sqrt(k)-1)*exp(-k*x^(1/sqrt(k)))/gamma(k)}
  auxk1 <- (2:11) ##(1:10)/10  #(2:11)#10
  for(i in 1:10)
  {
    auxiliar1 <- function(x){aux(x,k=auxk1[i])}
    plot(auxiliar1,type = "l",col = i,lty = 1,lwd = 1.5,main=" ", xlab="", ylab=" ",
          xlim=c(-0.5,6), ylim=c(-0.01,1),add=TRUE)
  }
}

```

```

}
gammagen1()

```

A.2. Gráficas del Capítulo 2

En esta sección se mostrarán los programas para trazar las gráficas del capítulo 2.

```

##### PARA LA GRÁFICA 2.1 #####
densidad <- function(x,k)
{
  k^(k-0.5)*exp(x*sqrt(k)-k*exp(x/sqrt(k)))/gamma(k)
}
histograma <- function(n,k)
{
  x <- sqrt(k)*log(rgamma(n,k,1)/k)
  par(mfrow = c(1, 3))
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(a)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
  x <- sqrt(k)*log(rgamma(4*n,k,1)/k)
  x <- sqrt(k)*log(rgamma(2*n,k,1)/k)
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(b)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
  x <- sqrt(k)*log(rgamma(20*n,k,1)/k)
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(c)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
}
histograma(500,0.5)

##### PARA LA GRÁFICA 2.3 #####
## HACIENDO LAS MODIFICACIONES RESPECTIVAS SE PUEDEN TRAZAR LAS GRÁFICAS 2.2, 2,4 Y 2.5 ##
#####
mui <- function(k,m)
{
  if(m==2){valor <- psigamma(k,1)}
  else
  {
    if(m==3){valor <- psigamma(k,2)}
    else
    {
      if(m==4){valor <- (psigamma(k,3)+3*(psigamma(k,1))^2)}
      else
      {
        if(m==5){valor <- (psigamma(k,4)+10*psigamma(k,1)*psigamma(k,2))}
        else
        {
          if(m==6){valor <- (psigamma(k,5)+15*psigamma(k,1)*psigamma(k,3)+10*
            (psigamma(k,2))^2+15*(psigamma(k,1))^3)}
          else
          {

```

```

        if(m==7){valor <- (psigamma(k,6)+15*psigamma(k,1)*psigamma(k,4)+35*psigamma(k,2)*
psigamma(k,3)+ 105*(psigamma(k,1))^2*psigamma(k,2)+6*psigamma(k,1)*psigamma(k,4))}
      }
    }
  }
}
valor
}
parametroreal <- function(k,m)
{
  c1 <- log(k)-psigamma(k,0)-1/(2*k)+1
  mu1 <- sqrt(k)*(psigamma(k,0)-log(k))
numerador <- (c1-1)*mu1/sqrt(k)+mu1^2/(2*k)-2*c1
  factor <- mu1/(2*sqrt(k))-1
  comun <- exp(psigamma(k,0))/k
  aux1 <- (mu1/(2*sqrt(k))+c1)^2+numerador*comun+(factor*comun)^2+mui(k,2)/4
  sumaux <- 0
for(j in 0:m)
{
  aux2 <- 1/(2*factorial(j))-1/(2*k*factorial(j+1))+numerador/factorial(j+2)
  aux3 <- 1/(4*factorial(j))+2*factor/factorial(j+1)+4*factor^2/factorial(j+2)
sumaux <- sumaux + ( aux2*comun + aux3*(comun^2)*2^j )*mui(k,j+2)
}
informacion <- aux1 + sumaux
abs(1/informacion)
}
estimador1 <- function(datos)
{
  vero <- function(x)
  {
    (x-0.5)*log(x)+mean(datos)*sqrt(x)-log(gamma(x))-x*mean(exp(datos/sqrt(x)))
  }
  optimize(vero,c(0.001,200),tol = 0.00001,maximum=TRUE)$maximum
}
histo <- function(k,n,m,momentos)
{
  aux <- rep(NA,m)
  for(i in 1:m)
  {
    datos <- function(n,k){sqrt(k)*log(rgamma(n,k,1)/k)}
    aux[i] <- estimador1(datos(n,k))
  }
  parametro <- parametroreal(k,momentos)
  desviacion <- sqrt(parametro/n)
  media1 <- mean(aux)
desviacion1 <- sd(aux)
  print(c(k, (desviacion^2))) ##Por la estadística
  print(c(media1,(desviacion1^2))) ## Por los datos
  plot(density(aux),lty = 1,lwd = 2,col = 1,type="l",ylim=c(0,20))
  curve(dnorm(x,k,desviacion), col = 2, lty = 1, lwd = 2, add = TRUE,bg=2)
}

```

```

    curve(dnorm(x,medial,desviacion1), col = 4, lty = 1, lwd = 2, add = TRUE,bg=2)
  }
histo1 <- function(k,n,m)
{
  au <- c(1,2,3,5)
  par(mfrow = c(2, 3))
  for(j in 0:5)
  {
    momentos <- j
    aux <- rep(NA,m)

    for(i in 1:m)
    {
      datos <- function(n,k){sqrt(k)*log(rgamma(n,k,1)/k)}
      aux[i] <- estimador1(datos(n,k))
    }
    parametro <- parametrorreal(k,momentos)
    desviacion <- sqrt(parametro/n)
    medial <- mean(aux)
    desviacion1 <- sd(aux)
    print(c(k, (desviacion^2))) ##Por la estadística
    print(c(medial,(desviacion1^2))) ## Por los datos
    aa <- j+2
    plot(density(aux),lty=1,lwd=2,col=1,type="l",ylim=c(0,20),xlab=(aa),ylab="",main="")
    curve(dnorm(x,k,desviacion), col = 2, lty = 1, lwd = 2, add = TRUE,bg=2)
    curve(dnorm(x,medial,desviacion1), col = 4, lty = 1, lwd = 2, add = TRUE,bg=2)
  }
}
histo(1,100000,10000,5)

##### PARA LA GRÁFICA 2.6 #####
mui <- function(k,m)
{
  if(m==2){valor <- k*psigamma(k,1)}
  else
  {
    if(m==3){valor <- k^(1.5)*psigamma(k,2)}
    else
    {
      if(m==4){valor <- k^2*(psigamma(k,3)+3*(psigamma(k,1))^2)}
      else
      {
        if(m==5){valor <- k^(2.5)*(psigamma(k,4)+10*psigamma(k,1)*psigamma(k,2))}
        else
        {
          if(m==6){valor <- k^3*(psigamma(k,5)+15*psigamma(k,1)*psigamma(k,3)+10*
            (psigamma(k,2))^2+15*(psigamma(k,1))^3)}
        }
      }
    }
  }
}
}

```

```

valor
}
parametrosreal <- function(k)
{
  if(k <= 1.0722)
  {
    c1 <- -1.0535
    c2 <- -1.5
    c3 <- -1
    media <- c1*(mui(k,2))^(c2)/(-mui(k,3))^(c3)
    var <- media^2*( c3^2*(mui(k,6))/(mui(k,3))^2 + c2^2*(mui(k,4))/(mui(k,2))^2- 2*c2*c3*
      (mui(k,5))/(mui(k,2)*mui(k,3)) -(c2-c3)^2 )
    parametros <- c(media+2.1446,var)
  }
  else
  {
    {
      if(k <= 10.6352)
      {
        c1 <- 1.2101
        c2 <- 2.808
        c3 <- 1.872
        media <- c1*(mui(k,2))^(c2)/(-mui(k,3))^(c3)
        var <- media^2*( c3^2*(mui(k,6))/(mui(k,3))^2 + c2^2*(mui(k,4))/(mui(k,2))^2-
          2*c2*c3*(mui(k,5))/(mui(k,2)*mui(k,3)) -(c2-c3)^2 )
        parametros <- c(media,var)
      }
      else
      {
        c1 <- 1.0078
        c2 <- 2.997
        c3 <- 1.998
        media <- c1*(mui(k,2))^(c2)/(-mui(k,3))^(c3)
        var <- media^2*( c3^2*(mui(k,6))/(mui(k,3))^2 + c2^2*(mui(k,4))/(mui(k,2))^2-
          2*c2*c3*(mui(k,5))/(mui(k,2)*mui(k,3)) -(c2-c3)^2 )
        parametros <- c(media,var)
      }
    }
  }
  parametros
}
parametrosreal1 <- function(k)
{
  c1 <- 1.0517
  c2 <- 3.0495
  c3 <- 2.033
  media <- c1*(mui(k,2))^(c2)/(-mui(k,3))^(c3)
  var <- media^2*( c3^2*(mui(k,6))/(mui(k,3))^2 + c2^2*(mui(k,4))/(mui(k,2))^2- 2*c2*c3*
    (mui(k,5))/(mui(k,2)*mui(k,3)) -(c2-c3)^2 )
  parametros <- c(media,var)
  parametros
}
estimador2<-function(datos)

```

```

{
  asimetria <- mean((datos-mean(datos))^3)/(var(datos))^(3/2)
  if(asimetria <= -2){asimetria <- -1.999}
  if(asimetria < -1.102){ k <- -1.0535*abs(asimetria)+2.1446}
  else
    {if(asimetria < -0.315){ k <- 1.2101/(abs(asimetria))^1.872}
      else{ k <- 1.0078/(abs(asimetria))^1.998}}
  k
}
histo <- function(k,n,m)
{
  aux <- rep(NA,m)
  for(i in 1:m)
  {
    datos <- function(n,k){sqrt(k)*log(rgamma(n,k,1)/k)}
    aux[i] <- estimador2(datos(n,k))
  }
  parametros <- parametrosreal(k)
  media <- parametros[1]
  desviacion <- sqrt(parametros[2]/n)
  media1 <- mean(aux)
  desviacion1 <- sd(aux)
  print(c(media, (desviacion^2))) ##Por la estadística
  print(c(media1, (desviacion1^2))) ## Por los datos
  plot(density(aux),lty = 1,lwd = 2,col = 1,type="l")
  curve(dnorm(x,media,desviacion), col = 2, lty = 1, lwd = 2, add = TRUE,bg=2)
  curve(dnorm(x,media,desviacion1), col = 4, lty = 1, lwd = 2, add = TRUE,bg=2)
}
histo(1,100000,10000)

```

A.3. Gráficas y cálculos por simulación del Capítulo 3

En esta sección se mostrarán los programas para trazar las gráficas y efectuar los cálculos de la simulación realizada en el capítulo 3.

```

##### PARA LA GRÁFICA 3.2 #####
##### GENERACIÓN DE LOG-GAMMA GENERALIZADA EN SU FORMA ESTÁNDAR #####
densidad <- function(x,k)
{
  k^(k-0.5)*exp(x*sqrt(k)-k*exp(x/sqrt(k)))/gamma(k)
}
histograma <- function(n,k)
{
  par(mfrow = c(1, 3))
  x <- sqrt(k)*log(rgamma(n,k,1)/k)
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(a)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
  x <- sqrt(k)*log(rgamma(4*n,k,1)/k)
  x <- sqrt(k)*log(rgamma(2*n,k,1)/k)
}

```

```

plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(b)",ylab="")
curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
x <- sqrt(k)*log(rgamma(20*n,k,1)/k)
plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(c)",ylab="")
curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
}
histograma(500,0.5) ## Para 1000 y 10,000 se cambia el valor

##### GENERACIÓN DE LOG-GAMMA GENERALIZADA EN SU FORMA NO ESTÁNDAR #####
densidad <- function(x,k,u,b)
{
  k^(k-0.5)*exp(((x-u)/b)*sqrt(k)-k*exp((x-u)/(b*sqrt(k))))/(b*gamma(k))
}

densidadgen <- function(n,k,u,b)
{
  par(mfrow = c(1, 3))
  x <- b*sqrt(k)*log(rgamma(n,k,1)/k)+u
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(a)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
  x <- b*sqrt(k)*log(rgamma(2*n,k,1)/k)+u
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(b)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
  x <- b*sqrt(k)*log(rgamma(20*n,k,1)/k)+u
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(c)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
}

##### PARA LA TABLA 3.2 #####
##### ESTIMADOR 1 EMV #####
estimador1 <- function(n,k)
{
  datos <- log(rgamma(n,1,1)) ##### o también ##### datos <- sqrt(k)*log(rgamma(n,k,1)/k)
  vero <- function(x,n)
  {
    exp((x-0.5)*log(x)+mean(datos)*sqrt(x)-log(gamma(x))-x*mean(exp(datos/sqrt(x))))
    ##### SE PUEDE PONER LA LOG-VER A UNA FUNCIÓN #####
  }
  optimize(vero,c(0.001, 5),tol = 0.000001,maximum=TRUE)$maximum*k
}

##### ESTIMADOR 2 MOMENTOS #####
invertir2 <- function(g0,error1)
{
  error <- 1
  k0 <- 1 ### La semilla
  while(error > error1)
  {
    numerador <- 2*psigamma(k0,1)*(psigamma(k0,2)-g0*(psigamma(k0,1))^(1.5))
    denominador <- 2*psigamma(k0,1)*psigamma(k0,3)-3*(psigamma(k0,2))^2
    kn <- k0 - numerador/denominador
    if(kn == 0){ error <- abs(kn-k0)}else{error <- abs((kn-k0)/kn)}
  }
}

```

```

    k0 <- kn
  }
  kn
}
Asimetria<-function(n,k)
{
  datos <- sqrt(k)*log(rgamma(n,k,1)/k)
  mean((datos-mean(datos))^3)/(var(datos))^(3/2)
}
estimador2 <- function(n,k,error1)
{
  invertir2(Asimetria(n,1),error1)*k
}
##### ESTIMADOR 3 MOMENTOS #####
estimador3 <- function(n,k,error1)
{
  error <- 1
  datos <- log(rgamma(n,1,1)) ###sqrt(k)*log(rgamma(n,k,1)/k)
  k0 <- 1 ### La semilla
  while(error > error1)
  {
    kn <- k0 + k0^(3/2)*(mean(exp(datos/sqrt(k0)))-1)/mean(datos*exp(datos/sqrt(k0)))
    if(kn == 0){ error <- abs(kn-k0)}else{error <- abs((kn-k0)/kn)}
    k0 <- kn
  }
  kn*k
}
##### PARA LA TABLA 3.3 #####
comparacion <- function(n,error1)
{
  auxk <- c(0.1,0.2,0.4,0.6,0.8,1,2,3,4,5,10,15,20)
  kontak <- length(auxk)
  matrizresultados <- matrix(NA,kontak,3)
  for(i in 1:kontak)
  {
    matrizresultados[i,] <- c(estimador1(n,auxk[i]),estimador2(n,auxk[i],error1),
                             estimador3(n,auxk[i],error1))
  }
  matrizresultados
}
comparacion(10000,0.00001)
##### PARA LA TABLA 3.4 #####
ECM1 <- function(n,k,m)
{
  vectoraux<-rep(NA,m)
  for(i in 1:m){vectoraux[i] <- estimador1(n,k)}
  c(var(vectoraux)/k, mean((vectoraux-k)^2)/k)
}
##### PARA LA TABLA 3.5 #####
ECM2 <- function(n,k,error1,m)
{

```



```

vectoraux<-rep(NA,m)
  for(i in 1:m){vectoraux[i] <- estimador2(n,k,error1)}
  c(var(vectoraux)/k, mean((vectoraux-k)^2)/k)
}
##### PARA LA TABLA 3.6 #####
ECM3 <- function(n,k,error1,m)
{
  vectoraux<-rep(NA,m)
  for(i in 1:m){vectoraux[i] <- estimador3(n,k,error1)}
  c(var(vectoraux)/k, mean((vectoraux-k)^2)/k)
}

```

A.4. Gráficas y cálculos por simulación del Capítulo 4

En esta sección se mostrarán los programas para el tamaño de la prueba y la potencia de la prueba utilizados en el capítulo 4.

```

#####
##### PROGRAMA PARA EL TAMAÑO Y POTENCIA DE LA PRUEBA #####
#####
densidad = function(y,k)
{
  if(k <= 130){k^(k-0.5)*exp(y*sqrt(k)-k*exp(y/sqrt(k)))/gamma(k)}
  else{exp(y*sqrt(k)-k*(exp(y/sqrt(k))-1))/sqrt(2*pi)}
}
cuantil<-function(c0,k,error1,iteraciones)
{
  numerador <- function(y0,k,iteraciones,c0)
  {
    kaux <- k
    auxiliar <- function(x){densidad(x,kaux)}
    integrate(auxiliar, -Inf, y0, subdivisions=iteraciones)$value-c0
  }
  buscar <- TRUE
  a <- -2
  b <- -1
  contador <- 1
  while(buscar && (contador < 500))
  {
    if(numerador(a,k,iteraciones,c0)*numerador(b,k,iteraciones,c0) < 0)
    {y0 <- (a+b)/2; buscar <- FALSE}
    else
    { if(numerador(a,k,iteraciones,c0) < 0){a <- b; b <- b+1}
      else{b <- a; a <- a-1}
    }
    contador <- contador + 1
  }
  error <- 1
  y0 ### La semilla

```

```

contador1 <- 1
while(error > error1 && (contador1 < 50))
{
  if( numerador(a,k,iteraciones,c0)*numerador(y0,k,iteraciones,c0) < 0)
    {b <- y0; y0 <- (a+b)/2; error <- abs(a-b)}
  else{a <- y0; y0 <- (a+b)/2; error <- abs(a-b)}
  contador1 <- contador1 + 1
}
y0
}
generador = function(n,k){sqrt(k)*log(rgamma(n,k,1)/k)}
correlacion<-function(vectorX,k,error1,iteraciones)
{
  n <- length(vectorX)
  vectorU <- rep(NA,n)
  rangodatos <- rank(vectorX)
  for(j in 1:n)
  {
    c1 <- (rangodatos[j]-0.5)/n
    vectorU[j] <- cuantil(c1,k,error1,iteraciones)
  }
  cor(vectorX,vectorU)
}

estimador2 = function(datos)
{
  asimetria <- -abs(mean((datos-mean(datos))^3)/(var(datos))^(3/2))
  if(asimetria <= -2){asimetria <- -1.999}

  if(asimetria < -1.5){ k <- -0.9367*abs(asimetria)+1.9411}
  else
    if(asimetria < -1.102){ k <- 1.352/(abs(asimetria))^2.231}
    else
      {if(asimetria < -0.315){ k <- 1.2764/(abs(asimetria))^1.809}
       else{ k <- 1.0438/(abs(asimetria))^1.987}}
  k
}

estimador1 <- function(datos)
{
  vero <- function(x)
  {
    (x-0.5)*log(x)+mean(datos)*sqrt(x)-log(gamma(x))-x*mean(exp(datos/sqrt(x)))
  }
  optimize(vero,c(-200,200),tol = 0.00001,maximum=TRUE)$maximum
}

unaRn = function(muestra,m,error1,iteraciones)
{
  vectorauxR = rep(NA,m)
  vectorauxK = rep(NA,m)

```

```

decision = rep(NA,12)
kmuestra = abs(estimador1(muestra)) ### aquí se cambia para el estimador 2
  n = length(muestra)
contadoraux1 = 0
while(contadoraux1 < m)
{
  muestrabootstrap = generador(n,kmuestra)
  contadoraux1 = contadoraux1 + 1
  kbootstrap = abs(estimador1(muestrabootstrap))
  ### aquí se cambia para el estimador 2
  vectorauxK[contadoraux1] = kbootstrap
  vectorauxR[contadoraux1] = correlacion(muestrabootstrap,kbootstrap,error1,iteraciones)
}
Rmuestral = correlacion(muestra,kmuestra,error1,iteraciones)

ordenados = sort(vectorauxR)
estimadorK = mean(vectorauxK)
valorcritico = ordenados[c((m*0.01),(m*0.02),(m*0.025),(m*0.03),(m*0.04),(m*0.05),
(m*0.06),(m*0.07),(m*0.075),(m*0.08),(m*0.09),(m*0.10))]
## ES EL ESTADÍSTICO DE PRUEBA ###
for( i in 1:12)
  if(valorcritico[i] > Rmuestral){decision[i] = 1}else{decision[i] = 0}
  ## 0-NO RECHAZAR, 1-RECHAZAR ###
  else{decision = rep(1,12)}
decision
}
potenciaprueba <- function(n,m,M,a,b,error1,iteraciones)
{
  salida = matrix(NA,M,12)
  potencia = rep(NA,12)
  for(i in 1:M)
  {
    muestras = rpareto(n,a,b)
    ##### AQUÍ VA LA DISTRIBUCIÓN PARA COMPARAR LA POTENCIA n Y PARÁMETROS #####
    salida[i,] = unaRn(muestras,m,error1,iteraciones)
    # print(i)
  }
# print(salida)
for(i in 1:12)
{
  potencia[i] = mean(salida[,i])
}
potencia
}

#####
##### GENERACIÓN DE NÚMEROS ALEATORIOS FALTANTES #####
#####
##### LAPLAS POR EL MÉTODO DE COMPOSICIÓN #####
rlaplace = function(n,a,b)
{

```

```

aux <- rep(NA,n)
contador <- 0
contador1 <- 0
for( i in 1:n)
{
  u <- runif(1)
  if(u < 0.5){contador <- contador +1; aux[contador] <- a+b*log(2*u)}else
  {aux[n-contador1] <- a-b*log(2*(1-u)); contador1 <- contador1+1 }
}
aux
}
#####          GUMBEL POR EL MÉTODO DE LA INVERSA #####
rgumbel = function(n,a,b){ a- b*log(-log(runif(n)))}
#####          FRECHET POR EL MÉTODO DE LA INVERSA #####
rfrechet = function(n,a,b,mu){mu+ b/(-log(runif(n)))^(1/a)}
#####          PARETO POR TRANSFORMACIÓN #####
rpareto = function(n,a,b){a/(1-runif(n))^(1/b)}
#####          LOG-LOGISTICA POR TRANSFORMACIÓN #####
rllogis = function(n,a,b){exp(rlogis(n,a,b))}
#####          LOG-CAUCHY POR TRANSFORMACIÓN #####
rlcauchy = function(n,a,b){exp(rcauchy(n,a,b))}
#####          LOG-STUDENT POR TRANSFORMACIÓN #####
rlt = function(n,a,b){exp(rt(n,a,b))}
#####          LOG-LAPLACE POR TRANSFORMACIÓN #####
rllaplace = function(n,a,b){exp(rlaplace(n,a,b))}

#####          GENERACIÓN DE LOG-GAMMA GENERALIZADA EN SU FORMA NO ESTÁNDAR #####
densidad <- function(x,k,u,b)
{
  k^(k-0.5)*exp(((x-u)/b)*sqrt(k)-k*exp((x-u)/(b*sqrt(k))))/(b*gamma(k))
}

densidadgen <- function(n,k,u,b)
{
  par(mfrow = c(1, 3))
  x <- b*sqrt(k)*log(rgamma(n,k,1)/k)+u
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(a)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
  x <- b*sqrt(k)*log(rgamma(2*n,k,1)/k)+u
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(b)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
  x <- b*sqrt(k)*log(rgamma(20*n,k,1)/k)+u
  plot(density(x),col=1,ylim=c(0,0.35),xlim=c(-10,2),lty=1,lwd=2,main="",xlab="(c)",ylab="")
  curve(densidad(x, k), col = 2, lty = 1, lwd = 2, ylim = c(0, 1), add = TRUE,bg=2)
}

#####          CARTAS DE CONTROL ###          CARTAS DE CONTROL ###          CARTAS DE CONTROL #####
#####          d2 PARA LA LOG-GAMMA GENERALIZADA
d2LGG=function(k)

```

```

{
  sqrt(k*psigamma(k,1))
}
### d3 PARA LA LOG-GAMMA GENERALIZADA
d3LGG=function(k,n)
{
  sqrt(k*(psigamma(k,3)+2*(psigamma(k,1))^2)/(4*psigamma(k,1)*n))
}
### TABLA DE VALORES PARA LA LOG-GAMMA GENERALIZADA #####
valoresLGG = function(m,k)
{
  matriz <- matrix(NA,(m-4),5)
  for(i in 5:m)
  {
    d2 <- d2LGG(k)
    d3 <- d3LGG(k,i)
    A2 <- 3/(d2*sqrt(i))
    D3 <- 1-3*d3/d2
    D4 <- 1+3*d3/d2
    matriz[(i-4),] <- c(i,d3,A2,D3,D4)
  }
  matriz
}
### FUNCIÓN DE VALORES PARA LA LOG-GAMMA GENERALIZADA #####
valorLGG=function(k,n)
{
  d2 <- d2LGG(k)
  d3 <- d3LGG(k,n)
  A2 <- 3/(d2*sqrt(n))
  D3 <- 1-3*d3/d2
  D4 <- 1+3*d3/d2
  c(n,d3,A2,D3,D4)
}
## CÁLCULOS PARA NIVEL DE CONFIANZA #####
distribucion = function(x,k)
{
  k^(k-0.5)*exp(x*sqrt(k)-k*exp(x/sqrt(k)))/gamma(k)
}
calculos <- function(k)
{
  inferior <- sqrt(k)*(psigamma(k,0)-log(k)-3*sqrt(psigamma(k,1)))
  superior <- sqrt(k)*(psigamma(k,0)-log(k)+3*sqrt(psigamma(k,1)))
  aux = function(x){distribucion(x,k)}
  integrate(aux,inferior,superior)
}

```