



**Distribuciones con Colas Pesadas Adecuadas para  
Modelar Datos en Estadística Actuarial**

**Por**

**Neveka Magaly Olmos Carvajal**

Tesis doctoral

Universidad de Las Palmas de Gran Canaria

Escuela de Doctorado. Programa de Doctorado en Turismo, Economía y Gestión

Director : **Emilio Gómez-Déniz**

Enero, 2026

Las Palmas de Gran Canarias, España





Universidad de Las Palmas de Gran Canaria

Escuela de Doctorado. Programa de Doctorado en Turismo, Economía y Gestión

Título de la Tesis: **Distribuciones con Colas Pesadas Adecuadas para Modelar Datos en Estadística Actuarial**

Autora: Neveka Magaly Olmos Carvajal

Director: Emilio Gómez-Déniz

Las Palmas de Gran Canarias, España, Enero de 2026

El Director

La Doctoranda



# Agradecimientos

A mi Director, Prof. Dr. Emilio Gómez-Déniz por su gran colaboración en todo el desarrollo de mi tesis.

## **RESUMEN**

En esta tesis se estudia la distribución Gleser y algunas distribuciones relacionadas. Se muestra que tanto la distribución Gleser como las otras distribuciones son de cola derecha pesada, condición muy común en el campo de la Estadística Actuarial. Se hará uso de la metodología slash para generar nuevas distribuciones relacionadas a la distribución de Gleser. Se estudian sus propiedades básicas y la estimación de los parámetros se realiza por el método de máxima verosimilitud. Se llevan a cabo estudios de simulación para evaluar el comportamiento de los estimadores de máxima verosimilitud. Finalmente, se realizan aplicaciones a datos de ingresos y seguros, donde se observa que estas nuevas distribuciones son una alternativa a distribuciones conocidas, como por ejemplo, la distribución de Pareto.

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Formulación del Problema . . . . .	1
1.2. Definición de los objetivos . . . . .	3
1.3. Resumen de los contenidos . . . . .	4
1.4. Consideraciones finales . . . . .	5
<b>2. Distribución Gleser</b>	<b>7</b>
2.1. Introducción . . . . .	7
2.2. Distribución G . . . . .	10
2.2.1. Propiedades . . . . .	12
2.2.2. Medida actuarial . . . . .	18
2.2.3. Estadísticos de Orden . . . . .	20
2.2.4. Entropía . . . . .	22

2.3.	Cola de la distribución G . . . . .	25
2.4.	Inferencia . . . . .	30
2.4.1.	Estimación por MV . . . . .	31
2.4.2.	Estudio de Simulación . . . . .	32
2.4.3.	Matriz de información de Fisher . . . . .	33
2.5.	Aplicaciones . . . . .	35
2.5.1.	Aplicación numérica . . . . .	35
2.5.2.	Datos de ingresos per cápita . . . . .	37
2.6.	Discusión . . . . .	42
<b>3.</b>	<b>Mezcla de escala de Gleser</b>	<b>44</b>
3.1.	Introducción . . . . .	44
3.2.	Distribución MEG y sus propiedades . . . . .	46
3.2.1.	Densidad . . . . .	47
3.2.2.	Propiedades . . . . .	48
3.2.3.	Cola de la distribución . . . . .	53
3.3.	Inferencia . . . . .	54
3.3.1.	Estimación de MV . . . . .	54
3.3.2.	Estudio de simulación . . . . .	55
3.3.3.	Matriz de información de Fisher . . . . .	56

3.4. Aplicación a datos de índices de siniestralidad . . . . .	58
3.5. Conclusiones . . . . .	60
<b>4. Distribución Hipergeométrica Gaussiana Gleser</b>	<b>63</b>
4.1. Introducción . . . . .	63
4.2. Función Densidad y Propiedades . . . . .	65
4.2.1. Función densidad . . . . .	65
4.2.2. Propiedades . . . . .	66
4.2.3. Cola de la distribución . . . . .	74
4.3. Inferencia . . . . .	75
4.3.1. Estimación de MV . . . . .	76
4.3.2. Matriz de información de Fisher . . . . .	77
4.3.3. Estudio de simulación . . . . .	78
4.4. Aplicación con datos de ingresos por familia . . . . .	78
4.5. Conclusiones . . . . .	82
<b>5. Conclusiones y trabajo futuro</b>	<b>84</b>

# Capítulo 1

## Introducción

### 1.1. Formulación del Problema

En el modelado de datos actuariales, estos suelen ser unimodales, sesgados a la derecha, positivos y tienen una densidad con cola derecha pesada. Entonces existe una necesidad de desarrollar y proponer nuevas distribuciones más flexibles a las distribuciones conocidas para modelar conjuntos de datos de ingresos, seguros y gestión de riesgos, entre otros.

Con respecto a distribuciones de colas pesadas, Andrews and Mallows (1974) desarrolló un estudio de mezclas de escalas de la distribución normal, y su investigación proporciona una familia de distribuciones con colas pesadas (por ejemplo, la cono-

cida distribución Student- $t$ ). Estas familias de distribuciones se han utilizado en la inferencia robusta de datos simétricos. Otro caso particular de esta familia de distribuciones de mezcla de escala es la distribución slash (ver Jonhson et al., 1995), que se basa en la distribución del cociente de dos variables aleatorias independientes, una normal estándar y una con distribución Beta. La distribución de este cociente tiene las colas más pesadas que la distribución normal estándar. Más detalles de esta distribución se pueden encontrar en Rogers and Tukey (1972) y en Mosteller and Tukey (1977). Kafadar (1982) discute las estimaciones de máxima verosimilitud para la distribución slash con localización y escala, Wang and Genton (2006) introducen extensiones multivariadas simétricas y asimétricas de la distribución slash, Gómez et al. (2007) y Gómez and Venegas (2008) extienden la distribución slash e introducen la familia de distribuciones slash-elíptica, entre otros.

Por otro lado, una distribución clásica que tiene la cola derecha pesada es la distribución de Pareto (ver Arnold, 2015) y que se ha utilizado para modelar la cuantía del daño por siniestro. Algunos trabajos publicados en la literatura estadística son Gómez-Déniz and Calderín-Ojeda (2014,2015,2020), entre muchos otros.

El trabajo de Gleser (1989) considera una distribución, la llamamos distribución Gleser (G), que está involucrada en la representación de la distribución gamma como una mezcla de escala de distribuciones exponenciales. La distribución G tiene el so-

porte positivo y la cola derecha pesada, condiciones necesarias para el modelado de datos actuariales, y esta es la motivación principal del desarrollo de esta tesis.

## **1.2. Definición de los objetivos**

El objetivo general de esta tesis es construir y aplicar a datos actuariales nuevas distribuciones con soporte positivo y cola derecha pesada.

Los siguientes son los objetivos específicos de esta tesis.

1. Establecer una versión con soporte positivo de la distribución  $G$ , estudiar sus propiedades y aplicarla a datos actuariales.
2. Construir y aplicar a datos actuariales una mezcla de escala de la distribución  $G$ .
3. Usando la metodología slash unificar las distribuciones  $G$  y aplicarla a datos actuariales.

### 1.3. Resumen de los contenidos

Para cumplir los objetivos definidos anteriormente, esta tesis se organiza en cinco Capítulos, en la que los Capítulos 2, 3 y 4 incluyen los resultados de esta tesis. Así, los contenidos de estos Capítulos son los siguientes.

En el Capítulo 2 se muestra como se genera la distribución  $G$ , se estudian sus propiedades, función VaR, función de riesgo y el comportamiento de la cola derecha vía la familias subexponencial y de variación regular. La estimación de los parámetros se realiza por el método de máxima verosimilitud y se evalúa el rendimiento de los estimadores mediante un estudio de simulación. Se muestran dos aplicaciones, una con datos simulados y otra con datos de ingresos, para la que la distribución  $G$  ajusta mejor que la distribución de Pareto.

En el Capítulo 3 se introduce una nueva distribución con cola derecha pesada, esta distribución es una mezcla de escala entre las distribuciones  $G$  y un caso particular de la distribución Beta (ambas distribuciones comparten un mismo parámetro). Se estudian sus propiedades y se muestra que pertenece a la familia de variación regular y por lo tanto es de cola pesada. Se utiliza el método de máxima verosimilitud para la estimación de los parámetros y se realiza un estudio de simulación donde se observa un buen comportamiento de los estimadores de máxima verosimilitud. Se realiza una aplicación a datos de seguros donde la nueva distribución ajusta mejor que la

distribución de Pareto.

En el Capítulo 4 se introduce una extensión de la distribución G que unifica las distribuciones Gleser's. Esta extensión se genera en base a la metodología del tipo slash, es decir, similar a la distribución del Capítulo 3 donde la mezcla de escala entre las distribuciones G y un caso particular de la distribución Beta no comparten ningún parámetro. Se estudian sus propiedades y se muestra que esta extensión de la distribución G es de cola pesada. Se utiliza el método de máxima verosimilitud para la estimación de los parámetros y un estudio de simulación confirma que los estimadores de máxima verosimilitud son consistentes. Se lleva a cabo una aplicación con datos de ingresos por familia que muestra un buen ajuste en comparación con otras distribuciones conocidas.

En el Capítulo 5 se dan algunas conclusiones y futuras investigaciones.

## 1.4. Consideraciones finales

La tesis doctoral consta de tres manuscritos publicados en revistas de indexación WoS (Q1), con pequeñas modificaciones para ajustarse al formato de este documento.

**Capítulo 2:** Olmos, N. M., Gómez-Déniz, E., Venegas, O. (2022). The Heavy-Tailed Gleser Model: Properties, Estimation, and Applications. *Mathematics*, 10(23), 4577.

<https://doi.org/10.3390/math10234577>.

**Capítulo 3:** Olmos, N. M., Gómez-Déniz, E., Venegas, O. (2024). Scale Mixture of Gleser Distribution with an Application to Insurance Data. *Mathematics*, 12(9), 1397. <https://doi.org/10.3390/math12091397>. This article belongs to the Special Issue Probabilistic Models in Insurance and Finance.

**Capítulo 4:** Olmos, N. M., Gómez-Déniz, E., Venegas, O. (2025). The Gauss hypergeometric Gleser distribution with applications to flood peaks exceedance and income data. *AIMS Mathematics*, 8(6), 13575-13593. doi: 10.3934/math.2025611. This article belongs to the Special Issue Advances in Probability Distributions and Social Science Statistics.

# Capítulo 2

## Distribución Gleser

### 2.1. Introducción

Las distribuciones de cola pesada se han utilizado para modelar datos en diversas ciencias aplicadas, como las ciencias ambientales, las ciencias de la tierra y las ciencias económicas y actuariales. Los conjuntos de datos de seguros tienden a ser positivos y asimétricos hacia la derecha, con colas pesadas (ver Ibragimov and Prokhorov, 2017); por lo tanto, las distribuciones con estas características se utilizan para modelar datos de seguros. Algunos autores han extendido ciertas distribuciones asimétricas con soporte positivo mediante la metodología slash, para aumentar el peso de la cola derecha, por ejemplo, Gómez et al. (2009) lo aplican a la distribución Birnbaum-Saunders,

Olmos et al. (2012,2014) en los modelos half-normal y half-normal generalizada (ver Cooray and Ananda, 2008), Astorga et al. (2017) en el modelo exponencial generalizado (ver Mudholkar et al., 1995; Gupta and Kundu, 1999), y Gómez et al. (2019) en el modelo de Gumbel. Otros dos trabajos recientes son, por ejemplo, Bhati and Ravi (2018), en el modelo log-Moyal generalizado, y Afify et al. (2020), en el modelo exponencial de cola pesada. Se sabe que la distribución de Pareto y sus generalizaciones correspondientes son distribuciones de cola pesada; se han utilizado en diversas áreas del conocimiento, por ejemplo, Choulakian and Stephens (2001), Zhang (2007), Akinsete et al. (2008), Nassar and Nada (2011), Mahmoudi (2011) y Boumaraf et al. (2020), entre otros. Es importante estudiar distribuciones con estas características para modelar, por ejemplo, conjuntos de datos de ingresos, seguros y rendimientos financieros. En el presente Capítulo, estudiamos una distribución con una cola derecha pesada que se ajusta bien a los datos de ingresos familiares.

Una función que es muy necesaria en este documento de tesis doctoral es la función beta, denotada por  $B(a, b)$ , y que se puede expresar como:

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt = \int_0^\infty \frac{t^{a-1}}{(1+t)^{a+b}}dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (2.1)$$

donde  $a > 0$ ,  $b > 0$  y  $\Gamma(\cdot)$  es la función gamma. La función beta es la constante de normalización de la distribución Beta, es decir, decimos que la variable aleatoria  $Y$

tiene distribución Beta con parámetros  $a$  y  $b$  si su función de densidad de probabilidad (fdp) es dada por

$$f_Y(y; a, b) = \frac{1}{B(a, b)} y^{a-1} (1-y)^{b-1}, \quad (2.2)$$

donde  $0 < y < 1$ ,  $a > 0$  y  $b > 0$ .

Otra función especial que merece ser destacada es la función beta incompleta, denotada por  $B(y; a, b)$ , y que se puede expresar como:

$$B(y; a, b) = \int_0^y t^{a-1} (1-t)^{b-1} dt, \quad (2.3)$$

donde  $a > 0$ ,  $b > 0$  y  $0 < y < 1$ . Otra función relacionada es la función beta incompleta regularizada, denotada por  $I_y(a, b)$ , y se expresa como  $I_y(a, b) = \frac{B(y; a, b)}{B(a, b)}$ .

Finalmente, otra función que también se utiliza en esta tesis es la función hipergeométrica de Gauss (ver Abramowitz and Stegun, 1968), se denota por  ${}_2F_1$ , y es dada por

$${}_2F_1(a, b, c; x) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 v^{b-1} (1-v)^{c-b-1} (1-xv)^{-a} dv, \quad (2.4)$$

donde  $a, b, c > 0$  y  $c > b$ .

Gleser (1989) introduce una representación de la distribución gamma, que es producto de una mezcla de escala entre una distribución desconocida y la distribución exponencial. El objetivo principal de esta tesis es estudiar esta distribución desconocida, que es la distribución G, y aplicarla al modelamiento de datos en Estadística

Actuarial.

Este Capítulo se desarrolla de la siguiente manera. En la Sección 2.2 se muestra como se genera la distribución G y se estudian algunas de sus propiedades. En la Sección 2.3 se estudian el comportamiento de la cola de la distribución G. En la Sección 2.4 se realiza la estimación de los parámetros por el método de máxima verosimilitud (MV), además de un estudio de simulación y la convergencia asintótica de los estimadores de MV. En la Sección 2.5 se llevan a cabo dos aplicaciones, una con datos simulados y otra con datos de ingresos. En la Sección 2.6 se entregan algunas conclusiones de este Capítulo.

## 2.2. Distribución G

En esta Sección se muestran como se genera la distribución G y se estudian sus propiedades básicas.

Gleser (1989) introduce una representación de la distribución gamma, como producto de una mezcla de escala entre la distribución exponencial y una distribución desconocida, tal como lo muestra el siguiente Teorema.

**Teorema 2.1** *Sea  $h(\cdot)$  la distribución gamma con parámetro de escala  $\sigma$  y parámetro de forma  $\alpha$  ( $0 < \alpha < 1$ ). Es decir,*

$$h(t) = \frac{\sigma^\alpha}{\Gamma(\alpha)} t^{\alpha-1} \exp(-\sigma t), \quad t > 0, \quad (2.5)$$

donde  $\sigma > 0$  y  $0 < \alpha < 1$ . Entonces

$$h(t) = \int_{\sigma}^{\infty} y \exp(-yt) g(y) dy, \quad (2.6)$$

donde

$$g(y) = \frac{(y - \sigma)^{-\alpha} \sigma^\alpha}{y \Gamma(1 - \alpha) \Gamma(\alpha)}, \quad y > \sigma. \quad (2.7)$$

**Demostración.**

$$h(t) = \int_{\sigma}^{\infty} y \exp(-yt) g(y) dy = \frac{\sigma^\alpha}{\Gamma(1 - \alpha) \Gamma(\alpha)} \int_{\sigma}^{\infty} (y - \sigma)^{-\alpha} \exp(-yt) dy,$$

haciendo el siguiente cambio de variables  $u = (y - \sigma)t$  se obtiene el resultado.  $\square$

**Comentario 2.1** *La restricción en el espacio paramétrico con  $0 < \alpha < 1$  en el Teorema es necesaria, ya que toda mezcla de escalas de distribuciones exponenciales debe tener una tasa de riesgo decreciente (ver Barlow and Proschan, 1975; Proschan, 1963).*

Estudiar la distribución  $g$  desconocida es el objetivo principal de esta tesis. La siguiente proposición nos entrega la forma que seguiremos de esta distribución.

**Proposición 2.1** *Sea  $Y$  la variable aleatoria que tiene densidad dada en (2.7), entonces la distribución de la variable aleatoria  $X = Y - \sigma$  viene dada por*

$$f_X(x; \sigma, \alpha) = \frac{\sigma^\alpha x^{-\alpha}}{B(\bar{\alpha}, \alpha)(\sigma + x)}, \quad x > 0, \quad (2.8)$$

donde  $\sigma > 0$  es un parámetro de escala,  $0 < \alpha < 1$  es un parámetro de forma y  $\bar{\alpha} = 1 - \alpha$ . Cuando una variable aleatoria  $X$  siga la fdp (2.8) se notará mediante  $X \sim G(\sigma, \alpha)$ .

**Demostración** Utilizando la fdp dada en (2.7) y aplicando el cambio de variable oportuno se obtiene el resultado.  $\square$

La distribución G es una distribución alternativa a las distribuciones con dos parámetros que se utilizan para modelar datos de Estadística Actuarial, por ejemplo, como es la distribución de Pareto, entre otras. La Figura 2.1 muestra el gráfico de la distribución G para diferentes valores del parámetro  $\alpha$ .

Realizamos una breve comparación que ilustra que la cola derecha de la distribución G es más pesada cuando el parámetro  $\alpha$  disminuye. En el Cuadro 4.1 muestra  $P(X > x)$  en la distribución G, para diferentes valores de  $x$  y  $\alpha$ . Se puede observar que cuando el parámetro  $\alpha$  decrece la cola derecha de la distribución G es más pesada.

### 2.2.1. Propiedades

La siguiente proposición muestra una forma de representar la distribution G.

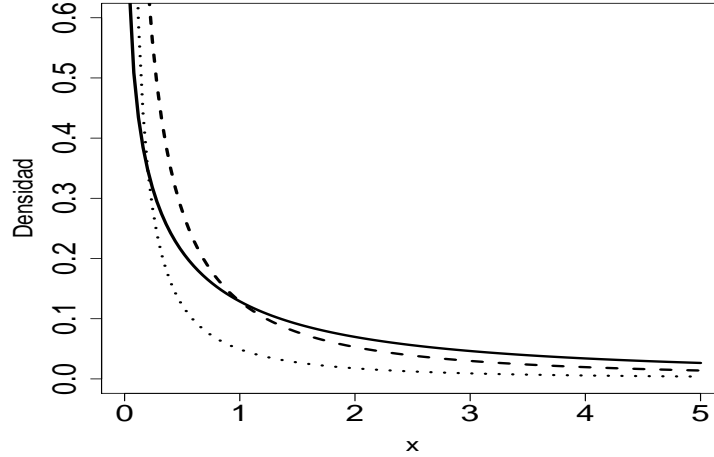


Figura 2.1: Ejemplos de la distribución  $G(1,0.3)$  (línea continua),  $G(1,0.7)$  (línea discontinua),  $G(1,0.9)$  (línea punteada)

**Proposición 2.2** Sea  $Y \sim \text{Beta}(\bar{\alpha}, \alpha)$ . Entonces  $X = \frac{\sigma Y}{1-Y} \sim G(\sigma, \alpha)$ .

**Demostración.** Utilizando el teorema de transformación de variables se obtiene el resultado.  $\square$

**Proposición 2.3** Sea  $X \sim G(\sigma, \alpha)$ . Entonces, la función de distribución acumulada (fda) de la variables aleatoria  $X$  viene dada por

$$F_X(x; \sigma, \alpha) = I_y(\bar{\alpha}, \alpha), \quad x > 0, \quad (2.9)$$

donde  $\sigma > 0$ ,  $0 < \alpha < 1$ ,  $y = \frac{x}{\sigma+x}$  y  $I_y(\cdot, \cdot)$  es la función beta incompleta regularizada.

Cuadro 2.1: Comparación de colas

Distribución	$P(X > 2)$	$P(X > 3)$	$P(X > 4)$
G(1,0.9)	0.04803	0.03537	0.02818
G(1,0.7)	0.19065	0.15106	0.12697
G(1,0.3)	0.63376	0.57717	0.53760

**Demostración.** Aplicando en forma directa la definición de fda se tiene

$$F_X(x; \sigma, \alpha) = \frac{\sigma^\alpha}{B(\bar{\alpha}, \alpha)} \int_0^x \frac{t^{-\alpha}}{\sigma + t} dt = \frac{1}{B(\bar{\alpha}, \alpha)} \int_0^{\frac{x}{\sigma}} \left( \frac{u}{1+u} \right) u^{-1-\alpha} du,$$

y haciendo el cambio de variable  $w = \frac{u}{1+u}$ , se obtiene el resultado  $\square$ .

La función de supervivencia  $s(t)$ , que es la probabilidad de que un artículo no falle antes del tiempo  $t$ , está definida por  $s(t) = \bar{F}(t) = 1 - F(t)$ . La función de supervivencia para la variable aleatoria  $T \sim G(\sigma, \alpha)$  viene dada por  $s(t) = 1 - I_{t/(\sigma+t)}(\bar{\alpha}, \alpha)$ . La función de riesgo,  $h(t)$ , definida por  $h(t) = \frac{f(t)}{s(t)}$ , para la variable aleatoria  $T \sim G(\sigma, \alpha)$  es dada por

$$h(t) = \frac{\sigma^\alpha t^{-\alpha}}{(\sigma + t) \left( B(\bar{\alpha}, \alpha) - B\left(\frac{t}{\sigma+t}, \bar{\alpha}, \alpha\right) \right)}, \quad t > 0.$$

La Figura 2.2 muestra la forma que tiene la función de riesgo para distintos valores de  $\alpha$ , considerando  $\sigma = 1$ .

**Proposición 2.4** *Sea  $T \sim G(\sigma, \alpha)$ . Entonces, la función de riesgo de  $T$  es decre-*

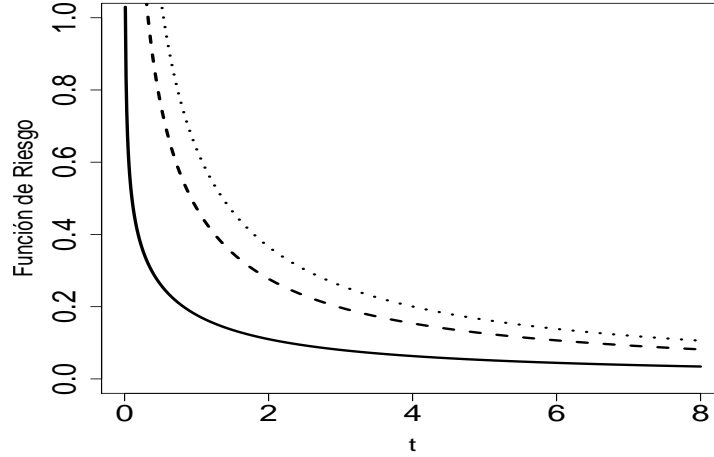


Figura 2.2: Ejemplos de la función de riesgo  $h(t)$ , para  $\alpha = 0.3$  (línea continua),  $\alpha = 0.7$  (línea discontinua),  $\alpha = 0.9$  (línea punteada)

*ciente para todo  $t > 0$ .*

**Demostración.** Utilizando el Teorema, ítem (b), dado en Glaser (ver Glaser, 1980) se tiene que

$$\eta(t) = -\frac{f'(t)}{f(t)} = \frac{\alpha}{t} + \frac{1}{\sigma + t},$$

donde  $f(t)$  es la fdp dada en (2.8), luego de la derivada de la función  $\eta(t)$  con respecto a  $t$

$$\eta'(t) = -\left(\frac{\alpha}{t^2} + \frac{1}{(\sigma + t)^2}\right) < 0, \quad \forall t > 0$$

se concluye el resultado  $\square$ .

**Proposición 2.5** *Los  $r$ -ésimos momentos,  $r = 1, 2, \dots$ , de la variable aleatoria  $X \sim G(1, \alpha)$  no existen.*

**Demostración.** Se tiene que por la integral valor- $p$ , la  $\int_0^\infty \frac{1}{t^p} dt$  converge si  $p > 1$ , por otro lado

$$\mathbb{E}(X^r) = \frac{1}{B(\alpha, \bar{\alpha})} \int_0^\infty \frac{x^{-\alpha+r}}{1+x} dx = \frac{1}{B(\alpha, \bar{\alpha})} \left\{ \int_0^1 \frac{x^{-\alpha+r}}{1+x} dx + \int_1^\infty \frac{x^{-\alpha+r}}{1+x} dx \right\},$$

El integrando de la primera integral  $h_1(x) = \frac{x^{-\alpha+r}}{1+x}$  tiene la misma magnitud que  $h_2(x) = \frac{1}{x^{\alpha-r}}$ , ya que  $\lim_{x \rightarrow \infty} \frac{h_1(x)}{h_2(x)} = \lim_{x \rightarrow 0} \frac{1}{1+x} = 1$ . Luego, se observa que debe tener la condición que  $r < \alpha < 1+r$  y sólo para  $r = 0$  se cumple la condición paramétrica y convergencia, luego para  $r = 1, 2, 3, \dots$  la integral diverge. El integrando de la segunda integral  $g_1(x) = \frac{x^{-\alpha+r}}{1+x}$  tiene la misma magnitud que  $g_2(x) = \frac{1}{x^{1+\alpha-r}}$ , ya que  $\lim_{x \rightarrow \infty} \frac{g_1(x)}{g_2(x)} = \lim_{x \rightarrow \infty} \frac{x}{1+x} = 1$ , aquí el valor  $p = 1 + \alpha - r$  y para  $r = 1, 2, 3, \dots$  el valor de  $p < 1$ , lo que indica que la integral diverge. Por lo tanto la integral completa diverge y esto significa que no existen los  $r$ -ésimos momentos de la variable aleatoria  $X \sim G(1, \alpha)$ .  $\square$

**Proposición 2.6** *Sea  $X \sim G(1, \alpha)$ . Entonces  $T = \frac{1}{1+X} \sim \text{Beta}(\alpha, \bar{\alpha})$ .*

**Demostración.** Utilizando transformación de variables aleatorias se obtiene el resultado.  $\square$

Por otro lado, los  $r$ -ésimos momentos de la variable aleatoria  $T$  son finitos y se pueden expresar como.

$$\mathbb{E}(T^r) = \frac{B(r + \alpha, \bar{\alpha})}{B(\alpha, \bar{\alpha})}.$$

Por ejemplo, el primer y segundo momentos de la variable aleatoria  $T$  son:

$$\mathbb{E}(T) = \alpha, \quad \mathbb{E}(T^2) = \frac{\alpha(1 + \alpha)}{2}.$$

**Proposición 2.7** *Sea  $X \sim G(\sigma, \alpha)$ . Entonces, la función cuantil ( $Q$ ) de la distribución  $G$  viene dada por*

$$Q(p) = \frac{\sigma I_p^{-1}(\bar{\alpha}, \alpha)}{1 - I_p^{-1}(\bar{\alpha}, \alpha)}, \quad 0 < p < 1, \quad (2.10)$$

donde  $I_p^{-1}$  es la función inversa de la función beta incompleta regularizada.

**Demostración.** Utilizando la fda dada en (2.9) se tiene que

$$p = F_X(x; \sigma, \alpha) = I_{x/(\sigma+x)}(\bar{\alpha}, \alpha).$$

Aplicando ahora la función inversa de la función beta incompleta regularizada

$$\frac{x}{\sigma + x} = I_p^{-1}(\bar{\alpha}, \alpha),$$

de donde despejando para  $x$  se obtiene el resultado  $\square$ .

Para estudiar los efectos del parámetro  $\alpha$  en los coeficientes de asimetría y kurtosis, utilizamos los coeficientes definidos por Galton (1883) y Moors (1988) respectivamente, y que están basados en la función cuantil y vienen dados en el siguiente resultado.

**Corolario 2.1** Sea  $X \sim G(\sigma, \alpha)$ , entonces los coeficientes de asimetría( $\sqrt{\beta_1}$ ) y curtosis( $\beta_2$ ) son respectivamente:

$$\sqrt{\beta_1} = \frac{Q\left(\frac{3}{4}\right) + Q\left(\frac{1}{4}\right) - 2Q\left(\frac{2}{4}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)},$$

$$\beta_2 = \frac{Q\left(\frac{3}{8}\right) - Q\left(\frac{1}{8}\right) + Q\left(\frac{7}{8}\right) - Q\left(\frac{5}{8}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)},$$

donde  $Q(\cdot)$  es la función cuantil dada en (2.10). La Figura 2.3 muestra las gráficas de los coeficientes de asimetría y curtosis en la distribución G.

### 2.2.2. Medida actuarial

Las distribuciones de probabilidad son utilizadas para describir una exposición al riesgo. Estos niveles de exposición al riesgo se muestran a través de indicadores que están basadas en propiedades de la distribución, como por ejemplo la función cuantil. Los actuarios a menudo usan estos indicadores para determinar el grado en que las empresas están sujetas a riesgo, como por ejemplo, precios de las acciones o tasas de interés, entre otras. Discutimos ahora la medida de riesgo VaR, también conocida como medida de riesgo cuantil o principio de prima cuantil, y que se especifica usualmente con un grado de confianza dado, que puede ser del 90%, 95% o 99%.

**Definición 2.1** El Valor en Riesgo de  $X$  al nivel de confianza  $100p$  % se denota por  $VaR_p$ , y corresponde al  $100p$  cuantil de la distribución del riesgo  $X$ .

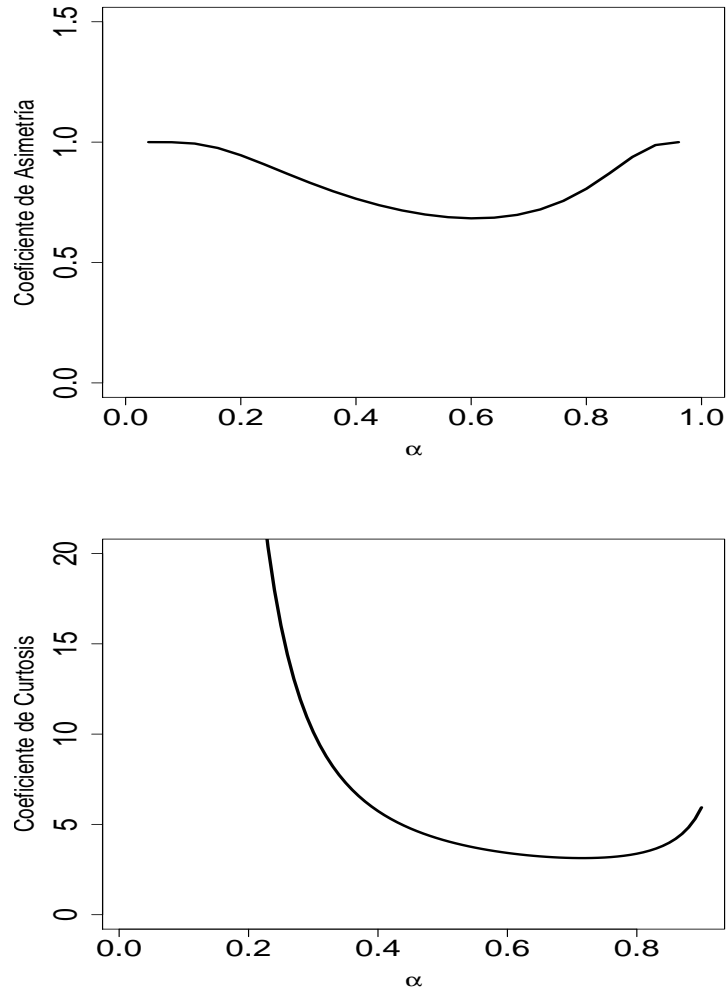


Figura 2.3: Gráficos de los coeficientes de asimetría y curtosis para la distribución G.

El VaR se utiliza para evaluar la exposición al riesgo, es decir, se puede utilizar para determinar la cantidad de capital necesaria para hacer frente resultados adversos. El VaR es pues una medida de riesgo pero que su utilización trasciende de la estadística

actuarial, haciéndose uso de ella fundamentalmente en economía financiera y cuantitativa. Usando la Definición 2.1 y (2.10), se tiene que el VaR de una variable aleatoria  $X \sim G(\sigma, \alpha)$  viene dada por (ver Artzner, 1999 and Artzner et al., 1999).

$$VaR_p = Q(p) = \frac{\sigma I_p^{-1}(\bar{\alpha}, \alpha)}{1 - I_p^{-1}(\bar{\alpha}, \alpha)}, \quad 0 < p < 1,$$

donde  $I_p^{-1}$  es la inversa de la función beta incompleta regularizada.

El Cuadro 2.2 proporciona algunos cálculos numéricos para la medida  $VaR_p$  de la distribución  $G(\sigma, \alpha)$  para diferentes valores de los parámetros.

La Figura 2.4 muestra gráficos de la medición de  $VaR_p$  de la distribución  $G(10, \alpha)$  para diferentes valores del parámetro  $\alpha$ . Observamos que a medida que el valor del parámetro  $\alpha$  decrece, el valor del  $VaR_p$  aumenta.

### 2.2.3. Estadísticos de Orden

Sean  $X_1, \dots, X_n$  una muestra aleatoria de la variable aleatoria  $X \sim G(\sigma, \alpha)$ , denotemos por  $X_{(j)}$  el  $j$ -ésimo estadístico de orden,  $j \in \{1, \dots, n\}$ .

**Proposición 2.8** *La fdp de  $X_{(j)}$  es*

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} \frac{\sigma^\alpha x^{-\alpha}}{B(\bar{\alpha}, \alpha)(\sigma+x)} [I_{x/(\sigma+x)}(\bar{\alpha}, \alpha)]^{j-1} [1 - I_{x/(\sigma+x)}(\bar{\alpha}, \alpha)]^{n-j}, \quad x > 0.$$

*En particular, la fdp del mínimo,  $X_{(1)}$ , es*

$$f_{X_{(1)}}(x) = \frac{n\sigma^\alpha x^{-\alpha}}{B(\bar{\alpha}, \alpha)(\sigma+x)} [1 - I_{x/(\sigma+x)}(\bar{\alpha}, \alpha)]^{n-1}, \quad x > 0.$$

Cuadro 2.2: Medida  $\text{VaR}_p$  y  $p$  es el nivel de significancia

Modelo	$p$	$\text{VaR}_p$	Modelo	$p$	$\text{VaR}_p$
G(1,0.2)	0.90	71645.52	G(10,0.2)	0.90	716455.2
G(1,0.2)	0.95	2292682	G(10,0.2)	0.95	22926824
G(1,0.2)	0.99	7164634401	G(10,0.2)	0.99	71646344013
G(1,0.5)	0.90	39.86346	G(10,0.5)	0.90	398.6346
G(1,0.5)	0.95	161.4476	G(10,0.5)	0.95	1614.476
G(1,0.5)	0.99	4052.181	G(10,0.5)	0.99	40521.81
G(1,0.8)	0.90	2.364536	G(10,0.8)	0.90	23.64536
G(1,0.8)	0.95	6.334845	G(10,0.8)	0.95	63.34845
G(1,0.8)	0.99	50.87682	G(10,0.8)	0.99	508.7682

y la fdp del máximo,  $X_{(n)}$ , es

$$f_{X_{(n)}}(x) = \frac{n\sigma^\alpha x^{-\alpha}}{B(\bar{\alpha}, \alpha)(\sigma + x)} [I_{x/(\sigma+x)}(\bar{\alpha}, \alpha)]^{n-1}, \quad x > 0$$

**Demostración.** Como es un modelo absolutamente continuo, la fdp del  $j$ -ésimo estadístico de orden se obtiene aplicando

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f(x) [F(x)]^{j-1} [1-F(x)]^{n-j}, \quad j \in \{1, \dots, n\},$$

donde  $F$  y  $f$  denotan la fda y la fdp de la distribución principal, que en este caso es

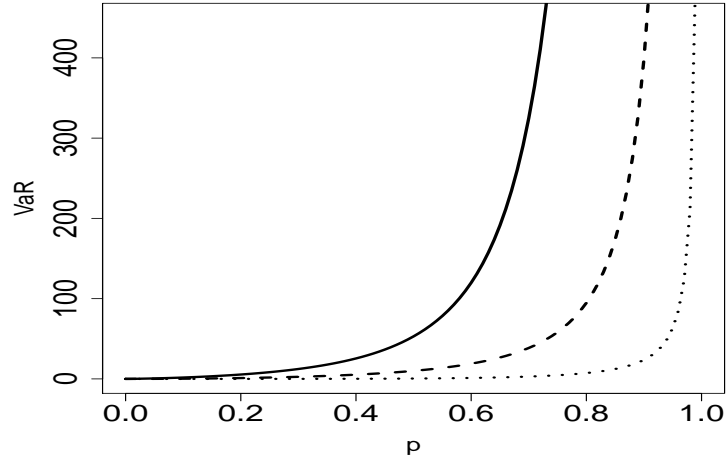


Figura 2.4: Gráficos del  $\text{VaR}_p$ , para  $G(1,0.3)$  (línea continua),  $G(1,0.5)$  (línea discontinua),  $G(1,0.8)$  (línea punteada)

$X \sim G(\sigma, \alpha)$ .  $\square$

## 2.2.4. Entropía

En esta Subsección se estudian las entropías de Shannon y Renyi.

**Definición 2.2** Sea  $X$  una variable aleatoria continua, entonces

1. la entropía de Shannon ( $S$ ) (ver Shannon, 1948) para una variable aleatoria  $X$ , es definida como

$$S(X) = -\mathbb{E}(\log f_X(x)). \quad (2.11)$$

2. una generalización de la entropía de Shannon es la entropía de Rényi ( $R_p$ ), que se define como

$$R_p(X) = \frac{1}{1-p} \log \left( \int_0^\infty [f(x)]^p dz \right). \quad (2.12)$$

El siguiente Lema es útil para calcular la entropía de Shannon en la distribución G.

**Lema 2.1** Sea  $X \sim G(\sigma, \alpha)$ , entonces se tienen los siguientes resultados.

1.  $\mathbb{E}(\log(X)) = \log(\sigma) + \psi(\bar{\alpha}) - \psi(\alpha)$ ,
2.  $\mathbb{E}(\log(\sigma + X)) = \log(\sigma) - \gamma - \psi(\alpha)$ ,

donde  $\psi(\cdot)$  es la función digamma y  $\gamma = -\psi(1)$  es la constante de Euler's.

**Demostración.** Ambos resultados se obtienen directamente utilizando la fdp que se muestra en (2.8).  $\square$

**Proposición 2.9** Sea  $X \sim G(\sigma, \alpha)$ .

1. Entonces la entropía de Shannon de  $X$  viene dada por

$$S(X) = \log(\sigma) + \log(B(\bar{\alpha}, \alpha)) - \gamma + \alpha\psi(\bar{\alpha}) - (1 + \alpha)\psi(\alpha). \quad (2.13)$$

2. Entonces, la entropía de Rényi de  $X$  viene dada por

$$R_p(X) = \frac{1}{1-p} [\log(B(\xi, p - \xi)) - (p - 1) \log(\sigma) - p \log(B(\bar{\alpha}, \alpha))]. \quad (2.14)$$

donde  $\xi = 1 - p\alpha > 0$ .

**Demostración.** En 1. aplicando (2.11) y el Lema 2.1 se obtiene el resultado, en 2. aplicando (2.12) y (2.1) se obtiene el resultado.  $\square$

La Figura 2.5 muestra la entropía de Shannon para el modelo G, fijando  $\sigma = 1$ .

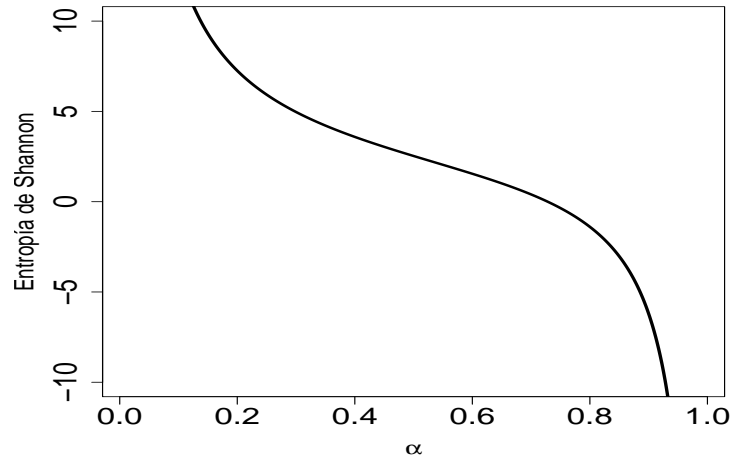


Figura 2.5: Entropía de Shannon para  $G(\alpha, \sigma = 1)$ .

**Corolario 2.2** Sea  $X \sim G(\sigma, \alpha)$  con  $S(X)$  y  $R_p(X)$  las entropías de Shanon y Renyi, entonces tenemos

$$\lim_{p \rightarrow 1} R_p(X) = S(X). \quad (2.15)$$

**Demostración.** Aplicando la regla de L'Hopital se obtiene el resultado.  $\square$

## 2.3. Cola de la distribución G

En el ámbito de los seguros, se suele utilizar una variable aleatoria con soporte no negativo, como la distribución de Pareto clásica, para modelar la magnitud de las pérdidas. El tamaño de la cola de la distribución es fundamental si se desea que el modelo elegido capture cantidades suficientemente alejadas del inicio del soporte de la distribución, es decir, valores atípicos (extremos). El concepto de cola pesada es fundamental para este y otros escenarios financieros.

El uso de distribuciones con colas pesadas a la derecha es de vital importancia en los seguros generales o no vida. Las distribuciones de Pareto, lognormal y Weibull, entre otras, se han utilizado para modelar las pérdidas de seguros de responsabilidad civil de terceros para vehículos, reaseguros y seguros contra catástrofes, entre otros.

Sea  $\mathcal{S}$  la clase de distribuciones subexponenciales. Es decir,  $F \in \mathcal{S}$  definida en  $\mathbb{R}_+$  que satisface:

$$\lim_{x \rightarrow \infty} \frac{1 - F^{*2}(x)}{1 - F(x)} = 2,$$

donde  $F^{*j}$  es la  $j$ -convolución de  $F$ .

El siguiente Lema, que aparece, entre otros, en Rolski et al. (1999) es necesario en el siguiente teorema.

**Lema 2.2** Sean  $F(x)$  y  $H(x)$  dos fda's en  $\mathbb{R}_+$  y se supone que existe una constante

$c > 0$  tal que

$$\lim_{x \rightarrow \infty} \frac{\overline{H}(x)}{\overline{F}(x)} = c. \quad (2.16)$$

Entonces,  $F \in \mathcal{S}$  si y sólo si  $H \in \mathcal{S}$ .

**Teorema 2.2** *Sea la fda de  $G(\sigma, \alpha)$  dada en (2.9). Entonces,  $F_X(x; \sigma, \alpha) \in \mathcal{S}$ .*

**Demostración.** Sea  $\overline{H}_X(x; \sigma, \alpha) = (\sigma/(\sigma + x))^\alpha$  la función de supervivencia de la distribución de Pareto tipo II (distribución de Lomax), que junto con el complemento de la fda dada en (2.9) y mediante el cálculo (2.16) tras aplicar la regla de L'Hôpital, obtenemos que

$$\lim_{x \rightarrow \infty} \frac{\overline{H}(x)}{\overline{F}(x)} = \alpha B(\overline{\alpha}, \alpha) > 0.$$

Ahora bien, teniendo en cuenta que la distribución de Pareto II pertenece a la clase de distribuciones subexponenciales, se obtiene el resultado.  $\square$

Como consecuencia del Teorema anterior (ver Rolski et al., 1999), tenemos el siguiente corolario.

**Corolario 2.3** *Si  $X_1$  y  $X_2$  son variables aleatorias independientes e idénticamente distribuidas con fda's dadas en (2.9), entonces, cuando  $x \rightarrow \infty$*

$$\Pr(X_1 + X_2 > x) \sim \Pr(\max\{X_1, X_2\} > x).$$

**Demostración.** El resultado es consecuencia del hecho de que  $F_X(x; \sigma, \alpha) \in \mathcal{S}$ .  $\square$

**Proposición 2.10** *Se verifica que  $F_X(x; \sigma, \alpha)$  es una distribución de cola derecha pesada.*

**Demostración.** Dado que  $F_X(x; \sigma, \alpha) \in \mathcal{S}$ , el resultado es consecuencia de la aplicación del Teorema 2.5.2 en Rolski et al. (1999).  $\square$

Otra forma de mostrar que es  $F_X(x; \sigma, \alpha)$  es una distribución con colas derecha pesada, calculando el límite  $\lim_{x \rightarrow \infty} (-\log \bar{F}_X(x; \sigma, \alpha)/x)$ , lo cual resulta 0. Observe que  $-\log \bar{F}_X(x; \sigma, \alpha)/x$  es la función de riesgo de  $X$ .

Como consecuencia del resultado anterior, tenemos el siguiente Corolario:

**Corolario 2.4** *Se verifica que  $\limsup_{x \rightarrow \infty} e^{sx} \bar{F}_X(x; \sigma, \alpha) = \infty$ ,  $s > 0$ .*

En este caso, la distribución no posee ningún momento exponencial positivo, es decir,  $\int \exp(sx) dF_X(x; \sigma, \alpha) = \infty$  para todo  $s > 0$  (ver Foss et al., 2011). Las distribuciones de este tipo tienen función generadora de momentos  $M_{F_X(x; \sigma, \alpha)}(s) = \infty$ , para todo  $s > 0$ . Esto ocurre, por ejemplo, con la distribución log-normal.

Un aspecto importante en la teoría de valores extremos es la variación regular (ver Bingham, 1987 ; Konstantinides, 2018). Esta es una descripción flexible de la variación de algunas funciones según la forma polinómica del tipo  $x^{-\delta} + o(x^{-\delta})$ ,  $\delta > 0$ . Este concepto se formaliza en la siguiente definición.

**Definición 2.3** Una fda se denomina función de variación regular en el infinito con índice  $-\delta$  si se cumple:

$$\lim_{x \rightarrow \infty} \frac{\overline{F}(\tau x)}{\overline{F}(x)} = \tau^{-\delta},$$

donde  $\tau > 0$  y el parámetro  $\delta \geq 0$  se denomina índice de cola.

La siguiente proposición establece que la función de supervivencia de la distribución  $G$  es una distribución con variación regular.

**Proposición 2.11** La función de supervivencia de  $G(\sigma, \alpha)$  es de variación regular con el índice de cola  $\alpha$ .

**Demostración.** Aplicando la definición anterior y utilizando la regla de L'Hôpital, tenemos que

$$\lim_{x \rightarrow \infty} \frac{\overline{F}(tx)}{\overline{F}(x)} = t^{-\alpha+1} \lim_{x \rightarrow \infty} \left( \frac{\sigma + x}{\sigma + tx} \right)^{-\alpha} \left( \frac{\sigma + x}{\sigma + tx} \right)^{\alpha-1} \left( \frac{\sigma + x}{\sigma + tx} \right)^2 = t^{-\alpha+1} \lim_{x \rightarrow \infty} \frac{\sigma + x}{\sigma + tx},$$

Calculando el límite de la derecha, obtenemos el resultado.  $\square$

En el ámbito actuarial y en los modelos de riesgo individual y colectivo, el investigador suele estar interesado en la variable aleatoria.  $S_n = \sum_{i=1}^n X_i$  para  $n \geq 1$ . Aunque su fdp es difícil o imposible de calcular en la práctica, podemos aproximar sus probabilidades utilizando el siguiente Corolario. (ver Jessen and Mikosch, 2006).

**Corolario 2.5** Si  $X_1, \dots, X_n$  son variables aleatorias independientes e idénticamente distribuidas  $G(\sigma, \alpha)$  y  $S_n = \sum_{i=1}^n X_i$ ,  $n \geq 1$ , entonces

$$\Pr(S_n > x) \sim n \Pr(X_1 > x), \quad \text{cuando } x \rightarrow \infty. \quad (2.17)$$

Por lo tanto, si  $P_n = \max_{i=1, \dots, n} X_i$ ,  $n \geq 1$ , tenemos que

$$\Pr(S_n > x) \sim n \Pr(X > x) \sim \Pr(P_n > x).$$

Esto significa que para  $x$  grande el evento  $\{S_n > x\}$  se debe al evento  $\{P_n > x\}$ .

Por lo tanto, que la suma  $S_n$  supere umbrales altos se debe a que el valor más grande de la muestra supera dicho umbral.

Jessen and Mikosch (2006) señala que la expresión (2.17) puede considerarse como la definición de una distribución subexponencial. La clase de esas distribuciones es mayor que la clase de distribuciones que varían regularmente. El resultado, dado en el Corolario 2.5, sigue siendo válido para distribuciones subexponenciales debido a la subexponencialidad de  $S_n$  implica la subexponencialidad de  $X_1$ . Normalmente, esta propiedad se conoce como cierre de raíz de convolución de distribuciones subexponenciales. Se pueden consultar más detalles en Embrechts and Goldie (1980, 1982).

Por otro lado, sea la variable aleatoria  $X$ , cuyo soporte es  $0 < x < \infty$ , que representa un límite de póliza o un deducible de reaseguro (desde la perspectiva de una aseguradora); entonces la función de valor esperado limitado  $L$  de  $X$  con fda

$F(x)$ , es definida por (ver Hogg and Klugman, 1984):

$$L(x) = \mathbb{E}[\min(X, x)] = \int_0^x y dF(y) + x\bar{F}(x),$$

que es la esperanza de la fda  $F(x)$  truncada en este punto. En otras palabras, representa la cantidad esperada por reclamación que retiene el asegurado en una póliza con un deducible fijo de  $x$ . Observe que integramos según el intervalo  $(0, x)$ .

**Proposición 2.12** *Sea  $X$  una variable aleatoria que sigue la fdp (2.8). Entonces, el valor esperado limitado de  $X$  viene dado por*

$$L(z) = \frac{z^2}{\sigma B(\bar{\alpha}, \alpha)(2 - \alpha)} \left(\frac{\sigma}{z}\right)^\alpha {}_2F_1(2 - \alpha, 1, 3 - \alpha, -z/\sigma) + z\bar{F}(z), \quad \alpha < 1.$$

donde  ${}_2F_1$  representa la función hipergeométrica dada en (2.4).

**Demostración.** Al realizar el cambio de variable  $t = x/(z - x)$  en la integral

$$I = \frac{\sigma^\alpha}{B(\bar{\alpha}, \alpha)} \int_0^z \frac{dx}{x^{\alpha-1}(\sigma + x)},$$

se obtiene el resultado.  $\square$

## 2.4. Inferencia

En esta Sección realizamos estimación de los parámetros del modelo G por el método de MV, un estudio de simulación y estimación asintótica de los estimadores de MV.

### 2.4.1. Estimación por MV

Para una muestra aleatoria  $x_1, \dots, x_n$  derivada desde la distribución  $G(\sigma, \alpha)$ , la función logaritmo de la verosimilitud (log-verosimilitud) se puede escribir como

$$\ell(\sigma, \alpha) = n\alpha \log(\sigma) - \alpha \sum_{i=1}^n \log(x_i) - n \log(B(\bar{\alpha}, \alpha)) - \sum_{i=1}^n \log(\sigma - x_i), \quad (2.18)$$

Las ecuaciones de score vienen dadas por:

$$\frac{n\alpha}{\sigma} - \sum_{i=1}^n \frac{1}{\sigma + x_i} = 0, \quad (2.19)$$

$$\log(\sigma) - \overline{\log(x)} + \psi(1 - \alpha) - \psi(\alpha) = 0, \quad (2.20)$$

donde  $\psi(\cdot)$  es la función digamma y  $\overline{\log(x)} = \frac{1}{n} \sum_{i=1}^n \log(x_i)$ . De (2.19) obtenemos

$$\hat{\alpha} = \frac{\hat{\sigma}}{n} \sum_{i=1}^n \frac{1}{\hat{\sigma} + x_i}, \quad (2.21)$$

y el estimador de MV para  $\sigma$  ( $\hat{\sigma}$ ) se obtiene resolviendo numéricamente la siguiente ecuación

$$\log(\hat{\sigma}) = \overline{\log(x)} + \psi\left(\frac{\hat{\sigma}}{n} \sum_{i=1}^n \frac{1}{\hat{\sigma} + x_i}\right) - \psi\left(1 - \frac{\hat{\sigma}}{n} \sum_{i=1}^n \frac{1}{\hat{\sigma} + x_i}\right), \quad (2.22)$$

El estimador  $\hat{\sigma}$  es la solución de la ecuación (2.22), y sustituyéndolo en (2.21) obtenemos  $\hat{\alpha}$ . La solución para esta ecuación (2.22) se puede obtener mediante procedimientos numéricos como el algoritmo de Newton-Raphson. Alternativamente, estas

estimaciones se pueden encontrar maximizando directamente de la función de log-verosimilitud dada en (2.18) y utilizando la subrutina `optim` en el software R (ver R Core Team, 2021).

### 2.4.2. Estudio de Simulación

Para examinar el comportamiento de la estimación por MV, presentamos un estudio de simulación para evaluar su rendimiento, evaluando los parámetros  $\sigma$  y  $\alpha$  en el modelo G. El análisis de simulación se realiza generando 1000 muestras de tamaños  $n = 100, 200$  y  $500$  del modelo G. El objetivo de esta simulación es estudiar el comportamiento de los estimadores de MV para los parámetros del modelo G. Se pueden utilizar dos algoritmos para generar números aleatorios del modelo G como se muestra más abajo.

---

Algorithm 1 for simulating from the  $X \sim G(\sigma, \alpha)$  can proceed as follows.

1. Step 1: Generate  $Y \sim Uniform(0, 1)$ .
2. Step 2: Compute  $X = \frac{\sigma I_Y^{-1}(\bar{\alpha}, \alpha)}{1 - I_Y^{-1}(\bar{\alpha}, \alpha)}$

---

Donde  $\sigma > 0$ ,  $0 < \alpha < 1$  y  $I_p^{-1}$  es la función inversa de la función beta incompleta regularizada. Para cada muestra generada de la distribución G, las estimaciones de

---

Algorithm 2 for simulating from the  $X \sim G(\sigma, \alpha)$  can proceed as follows.

1. Step 1: Generate  $Y \sim \text{Beta}(\bar{\alpha}, \alpha)$ .

2. Step 2: Compute  $X = \frac{\sigma Y}{1-Y}$

---

MV se obtienen aplicando el algoritmo de Newton-Raphson. El Cuadro 2.3 muestra el sesgo empírico (SE), la media de los errores estándar (EE), la raíz del error cuadrático medio empírico (ECME) y la probabilidad de cobertura (PC) del 95 % basada en la distribución asintótica para los estimadores de MV para los parámetros. Como se puede ver en el Cuadro 2.3, el rendimiento de las estimaciones mejora cuando  $n$  aumenta.

### 2.4.3. Matriz de información de Fisher

Consideremos ahora  $X \sim G(\sigma, \alpha)$ . Para una sola observación  $x$  de  $X$ , la función de log-verosimilitud para  $\boldsymbol{\theta} = (\sigma, \alpha)$  es dado por

$$\ell(\boldsymbol{\theta}) = \alpha \log(\sigma) - \alpha \log(x) - \log(B(\bar{\alpha}, \alpha)) - \log(\sigma + x).$$

La primera y segunda derivadas parciales correspondientes de la función de log-verosimilitud se entregan en el Apéndice A. Se puede demostrar que la matriz de información de Fisher, que denotamos por  $I_F(\cdot)$ , para la distribución G viene dada

Cuadro 2.3: SE, EE, ECME y 95 % PC para los estimadores de MV de  $\sigma$  y  $\alpha$  en la distribución G con diferentes combinaciones en los valores de los parámetros.

valor verdadero		$n = 100$				$n = 200$				$n = 500$				
$\sigma$	$\alpha$ par.	SE	EE	ECME	PC	SE	EE	ECME	PC	SE	EE	ECME	PC	
1	0.2	$\sigma$	0.1934	0.5709	0.6543	0.9402	0.1113	0.3727	0.4083	0.9474	0.0494	0.2219	0.2292	0.9540
		$\alpha$	0.0054	0.0260	0.0270	0.9618	0.0039	0.0181	0.0186	0.9562	0.0022	0.0113	0.0116	0.9542
	0.5	$\sigma$	0.2191	0.7544	0.8668	0.9004	0.1006	0.4920	0.5425	0.9156	0.0468	0.3010	0.3131	0.9370
		$\alpha$	0.0004	0.0679	0.0667	0.9258	-0.0012	0.0496	0.0491	0.9308	0.0005	0.0322	0.0320	0.9424
	0.8	$\sigma$	0.0727	0.5090	0.5498	0.8988	0.0348	0.3463	0.3602	0.9216	0.0118	0.2137	0.2143	0.9432
		$\alpha$	-0.0044	0.0258	0.0271	0.9550	-0.0022	0.0179	0.0183	0.9496	-0.0010	0.0112	0.0114	0.9510
10	0.2	$\sigma$	1.9401	5.7013	6.4320	0.9394	1.1103	3.7292	4.0676	0.9486	0.4651	2.2140	2.2556	0.9528
		$\alpha$	0.0052	0.0260	0.0268	0.9562	0.0037	0.0181	0.0189	0.9492	0.0020	0.0113	0.0112	0.9560
	0.5	$\sigma$	1.9611	7.4418	8.4714	0.8872	1.0610	4.9654	5.2373	0.9176	0.4148	2.9938	3.0764	0.9360
		$\alpha$	-0.0014	0.0681	0.0668	0.9216	0.0002	0.0497	0.0490	0.9344	0.0000	0.0322	0.0315	0.9484
	0.8	$\sigma$	0.6832	5.0843	5.3895	0.8952	0.3724	3.4679	3.5625	0.9236	0.1692	2.1465	2.1721	0.9406
		$\alpha$	-0.0047	0.0259	0.0274	0.9510	-0.0020	0.0179	0.0181	0.9544	-0.0008	0.0112	0.0113	0.9534

por

$$I_F(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\bar{\alpha}\alpha}{2\sigma^2} & -\frac{1}{\sigma} \\ -\frac{1}{\sigma} & \psi'(\alpha) + \psi'(\bar{\alpha}) \end{pmatrix}$$

donde  $\psi'(\cdot)$  es la función trigamma.

Por lo tanto, para muestras grandes, los estimadores de MV,  $\hat{\boldsymbol{\theta}}$ , de  $\boldsymbol{\theta}$  es asintóticamente normal bivariada, es decir,

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathcal{L}} N_2(\mathbf{0}, I_F(\boldsymbol{\theta})^{-1}),$$

lo que resulta en la varianza asintótica de los estimadores MV ( $\hat{\theta}$ ), y es la inversa de la matriz de información de Fisher  $I_F(\theta)$ . Dado que los parámetros son desconocidos, generalmente se considera la matriz de información observada, donde los parámetros desconocidos se estiman mediante MV.

## 2.5. Aplicaciones

En esta Sección analizamos dos aplicaciones: la primera con datos simulados y la segunda con datos reales. Para comparar los ajustes de los modelos, utilizamos los criterios de información de Akaike AIC (ver Akaike, 1974) y el de información bayesiano BIC (ver Schwarz, 1978). Se sabe que  $AIC = 2k - 2 \log lik$  y  $BIC = k \log(n) - 2 \log lik$ , dónde  $k$  es el número de parámetros del modelo ajustado,  $n$  es el tamaño de la muestra y  $\log lik$  es el valor maximizado de la función de verosimilitud.

### 2.5.1. Aplicación numérica

En esta aplicación numérica analizamos 200 datos simulados generados a partir del modelo  $G(1,0.3)$  usando el Algoritmo 2 dado en la Sección 2.4.2. El objetivo de este ejemplo numérico es utilizar el método de MV dado en la Sección 2.4.1 para determinar si, al ajustar el modelo  $G$ , este tiene una forma propia. Como consideramos

un valor de  $\alpha = 0.3$ , esto implica que los datos tienen una alta curtosis, esto implica que modelos con cola derecha pesada podrían ser buenos modelos para ajustar estos datos. Por esto consideramos el modelo clásico de Pareto para este ajuste. Decimos que una variable aleatoria  $X$  tiene distribución Pareto si su fdp es dada por

$$f_X(x; \alpha, \beta) = \frac{\beta \alpha^\beta}{x^{\beta+1}}, \quad x \geq \alpha, \quad (2.23)$$

donde  $\alpha > 0$  parámetro de escala y  $\beta > 0$  parámetro de forma. Se denotará mediante  $X \sim \text{Pareto}(\alpha, \beta)$ .

El Cuadro 2.4 muestra las estimativas de MV para los parámetros de los modelos G y Pareto, así como los valores AIC y BIC para cada modelo.

Cuadro 2.4: 200 datos simulados: Modelo, estimativas de MV, valores de AIC y BIC.

Modelo	Estimativas de MV	AIC	BIC
$G(\sigma, \alpha)$	$\hat{\sigma} = 0.890, \hat{\alpha} = 0.295$	<b>1967.556</b>	<b>1974.152</b>
$\text{Pareto}(\alpha, \beta)$	$\hat{\alpha} = 0.002, \hat{\beta} = 0.121$	2144.958	2151.555

Observamos que los valores de los criterios AIC y BIC del modelo G son significativamente más bajos que los valores del modelo Pareto, lo que significa que este modelo se ajusta mejor a los datos que el modelo Pareto. Esto es de esperar, ya que los datos se simularon a partir del modelo G. Los valores anteriores para las medidas

indican que el modelo G tiene una forma propia y que puede ser difícil reemplazarlo por cualquier otro modelo conocido de dos parámetros.

### 2.5.2. Datos de ingresos per cápita

Este conjunto de datos proviene de la Encuesta de Finanzas del Consumidor, una muestra representativa a nivel nacional que contiene amplia información sobre activos, pasivos, ingresos y características demográficas de los encuestados (clientes potenciales estadounidenses). Contiene una muestra aleatoria de 500 hogares con ingresos positivos entrevistados en la encuesta de 2004. La variable de interés es el ingreso anual de la familia en miles de dólares estadounidenses dividido entre el número de miembros del hogar. Las estadísticas descriptivas de estos datos se encuentran en el Cuadro 2.5, donde CS es el coeficiente de asimetría muestral y CK es el coeficiente de kurtosis muestral. Se puede observar un coeficiente de kurtosis alto, de 1.655, y esto significa que la cola derecha de los datos es muy pesada.

Cuadro 2.5: Estadísticas descriptivas para datos de ingresos

$n$	Mediana	Media	Varianza	CS	CK
500	21.125	216.709	11270001	0.435	1.655

En la Figura 2.6 se presentan dos gráficos de cajas y bigotes (boxplots), donde el

boxplot de arriba muestra un dato muy extremo y en el boxplot de abajo muestra el conjunto de datos quitando el dato extremo con el objetivo de mostrar los otros datos extremos que en el boxplot de la arriba no se pueden ver. Estos datos atípicos hacen que la cola derecha sea más pesada. Se puede notar que la mayoría de las observaciones están en torno a 21125 dólares per cápita por familia y existe un valor muy atípico que es un ingreso de 75 millones de dólares.

El Cuadro 2.6 muestra las estimativas de MV para los parámetros de los modelos G y Pareto, también se muestran los valores de los criterios de selección de modelos AIC y BIC para cada modelo.

Cuadro 2.6: Estimativas de MV para los datos de ingresos con sus correspondientes errores estándar (entre paréntesis), valores AIC y BIC.

Modelo	Estimativas de MV	AIC	BIC
G( $\sigma, \alpha$ )	$\hat{\sigma} = 21.555$ (6.264), $\hat{\alpha} = 0.497$ (0.033)	<b>5139.176</b>	<b>5147.605</b>
Pareto( $\alpha, \beta$ )	$\hat{\alpha} = 0.065$ (0.001), $\hat{\beta} = 0.171$ (0.008)	5867.910	5876.339

Se puede observar que los valores más pequeños de los criterios AIC y BIC corresponden al modelo G, esto significa que el modelo G ajusta mejor estos datos que el modelo de Pareto. Los errores estándar de las estimativas de MV del modelo G fueron calculados usando la subsección 2.4.3.

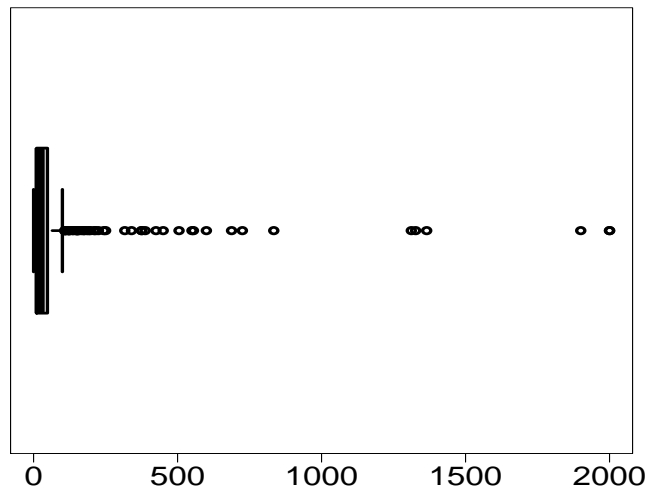
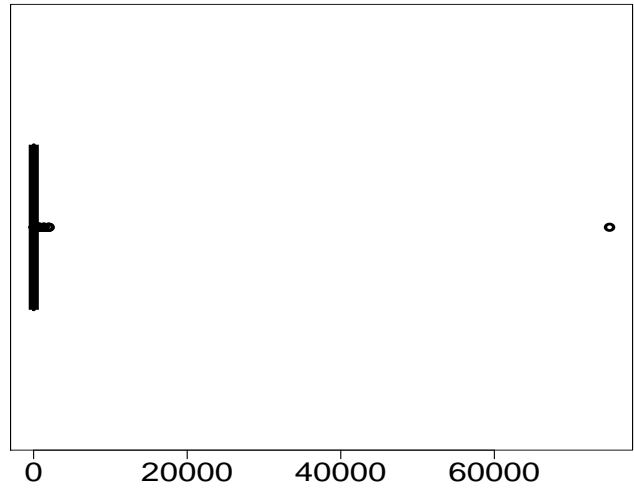


Figura 2.6: Boxplot para los datos de ingresos (arriba) y boxplot para los datos de ingresos sin el valor extremo (abajo).

Cuadro 2.7: Comparación de los valores estimados del VaR de los modelos G y Pareto para los datos de ingresos

Modelo \ Significancia	0.50	0.60	0.70	0.80	0.90	0.95
G( $\hat{\sigma}, \hat{\alpha}$ )	22.034	41.764	85.060	209.891	889.770	3632.538
Pareto ( $\hat{\alpha}, \hat{\beta}$ )	3.744	13.805	74.248	795.165	45800.120	2638007

El Cuadro 2.7 presenta estimaciones del VaR para los modelos G y Pareto al nivel de significación de: 0.50, 0.60, 0.70, 0.80, 0.90 y 0.95. Se sabe que los modelos con los valores de VaR más altos presentan las colas más pesadas. Por lo tanto, con base en esta característica, podemos observar que el modelo G proporciona valores de VaR más altos hasta el nivel de 0.70 que el modelo Pareto, ya que el VaR del modelo Pareto se disparan en cuantiles altos. En base a los criterios de selección de modelos AIC y BIC, el modelo G se ajusta mejor a los datos de ingresos. La Figura 2.7 muestra el comportamiento de los valores de VaR de los dos modelos con respecto al VaR empírico, se observa que el VaR del modelo G se encuentra más cerca al VaR empírico. La Figura 2.8 muestra la fda empírica con las fda estimadas de los modelos G y Pareto, lo que también muestra la excelente concordancia entre el modelo G y los datos de ingresos.

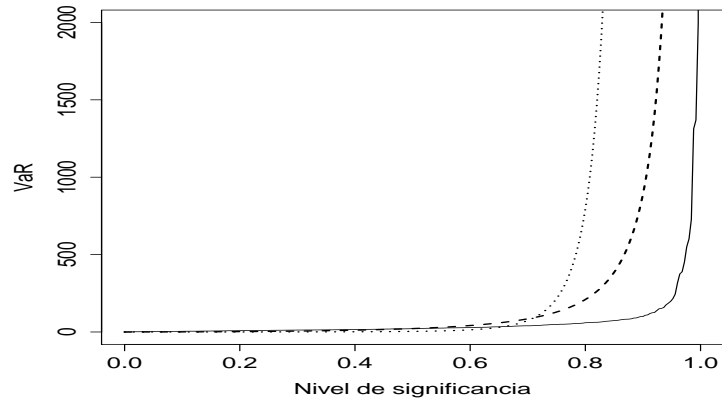


Figura 2.7: Gráficos del VaR empírica (línea continua), con las VaR estimados de los modelos G (línea discontinua) y Pareto (línea de punteada).

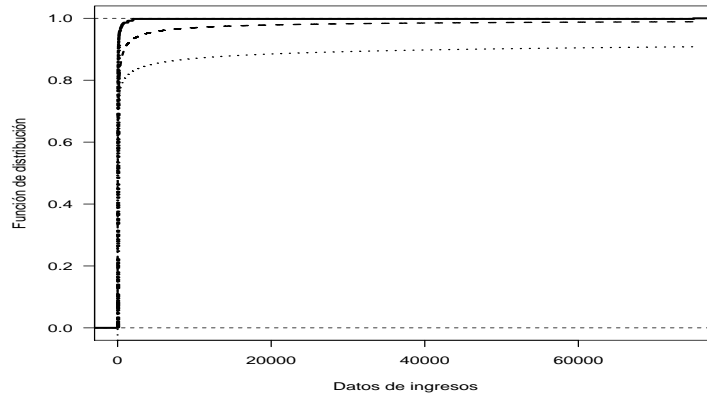


Figura 2.8: Gráficos de la fda empírica (línea continua) con las fda estimadas de los modelos G (línea discontinua) y Pareto (línea de punteada).

## 2.6. Discusión

Este Capítulo presenta un estudio del modelo G, que se obtiene a partir de una caracterización de la distribución gamma. El modelo G tiene dos parámetros, lo que lo convierte en una alternativa atractiva frente a diversos modelos de dos parámetros utilizados en estadística actuarial, como la distribución gamma, lognormal, Weibull y la distribución inversa Gaussiana. El modelo G se presenta como una opción viable para ajustar datos con observaciones extremas. Algunas características más del modelo G son:

- El modelo G presenta su cola derecha más pesadas a medida que el parámetro  $\alpha$  sea más pequeño.
- El modelo G tiene una representación explícita dada en la Proposición 2.2.
- Fdp, fda, función de riesgo y función cuantil son explícitas y están representadas por funciones conocidas.
- La medida de riesgo VaR es explícita y es utilizada para mostrar que la cola derecha del modelo G es pesada.
- Se muestra que el modelo G pertenece a la familia subexponencial, como también es de variación regular y por lo tanto es un modelo de cola derecha pesada.

- En las aplicaciones se muestra que el modelo G tiene su propia característica en comparación con el modelo de Pareto, y que es un buen candidato para modelar datos de ingresos con cola derecha pesada.

# Capítulo 3

## Mezcla de escala de Gleser

### 3.1. Introducción

Andrews and Mallows (1974) desarrolló un estudio de mezclas de escalas de la distribución normal, y su investigación proporcionó una familia de distribuciones con colas pesadas (por ejemplo, la conocida distribución  $t$  de Student). Estas familias de distribuciones se han utilizado en la inferencia robusta de datos simétricos. Una distribución de mezcla de escalas se puede obtener mezclando una densidad base con una distribución de escalas. La distribución resultante se puede expresar como sigue:

$$f(x) = \int g(x|\kappa(w))dH(w),$$

donde  $g(x|\kappa(w))$  es la distribución condicionada de la variable aleatoria  $X$ , dado que  $\kappa(w)$  es una función positiva de una variable aleatoria  $W$  con fda  $H(w)$ . Por ejemplo, en el caso de la distribución  $t$  de Student,  $W \sim \text{Gamma}(\nu/2, \nu/2)$  y  $\kappa(w) = 1/w$ . Otro caso que se ha estudiado varias veces en los últimos años es el de la distribución slash, donde  $W \sim \text{Uniforme}(0, 1)$  y  $\kappa(w) = w^{1/q}$ .

Como hemos indicado en el Capítulo 2, la distribución de Pareto (ver Pareto, 1897), que lleva el nombre de Vilfredo Pareto, es un modelo ampliamente utilizado en diversas ciencias aplicadas, como las ciencias actuariales, la economía, las finanzas, los estudios de vida, la climatología y la distribución de la renta, donde la ocurrencia se describe generalmente en observaciones extremas. Algunos investigadores que han utilizado la distribución de Pareto son Beirlant et al. (1996,2004) y Resnick (1997), entre otros.

Gómez-Déniz and Calderín-Ojeda (2014), estudia una distribución con dos parámetros que ofrece una alternativa a la distribución de Pareto; esta distribución se considera un caso particular de la distribución log-gamma (LG) y en este Capítulo la denotamos por LG2. Decimos que una variable aleatoria  $X$  tiene una distribución LG2 si su fdp es dada por

$$f_X(x; \alpha, \beta) = \frac{\beta^2 \log(x + 1 - \alpha)}{(x + 1 - \alpha)^{\beta+1}}, \quad x \geq \alpha,$$

donde  $\alpha \geq 0$  y  $\beta \geq 0$ . La denotamos por  $X \sim \text{LG2}(\alpha, \beta)$ .

El objetivo de este Capítulo es estudiar la distribución de mezcla de escala de Gleser (MEG), que está relacionada con la distribución G y es el resultado de una mezcla de escalas entre las distribuciones G y Beta. La distribución MEG tiene dos parámetros y una cola derecha pesada, lo que la convierte en una posible alternativa a las distribuciones conocidas de dos parámetros con cola pesada, para modelar datos con observaciones atípicas.

Este Capítulo se organiza de la siguiente manera: en la Sección 3.2, se estudian las principales propiedades de la distribución MEG; en la Sección 3.3, se estiman los parámetros mediante el método de MV, un estudio de simulación y la matriz de información de Fisher, que será útil para calcular los errores estándar asintóticos de las estimaciones de los parámetros, en la Sección 3.4, se muestra una aplicación a datos reales, y en la Sección 3.5, se presentan algunas conclusiones.

## **3.2. Distribución MEG y sus propiedades**

En esta Sección se presenta la distribución MEG y se estudian algunas de sus propiedades.

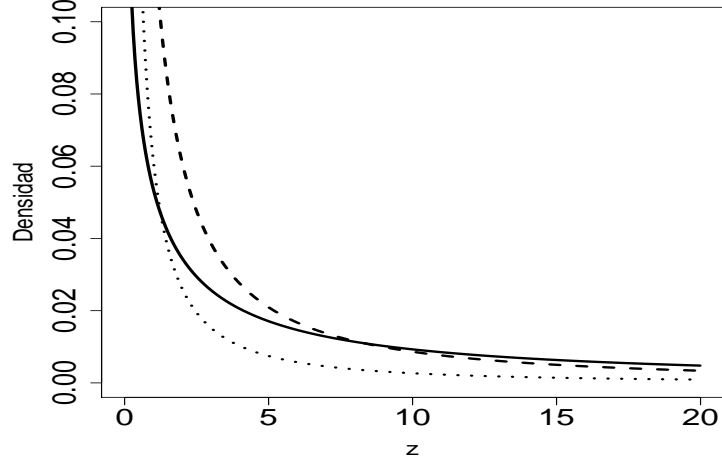


Figura 3.1: Ejemplos de la distribución MEG(1, 0, 3) (línea continua), MEG(1, 0, 7) (línea discontinua) y MEG(1, 0, 9) (línea punteada)

### 3.2.1. Densidad

**Definición 3.1** Sea  $Z \sim MEG(\sigma, \alpha)$ , la fdp de  $Z$  viene dada por

$$f_Z(z; \sigma, \alpha) = \frac{\alpha \sigma^\alpha}{B(\bar{\alpha}, \alpha)} z^{-(\alpha+1)} \log \left( 1 + \frac{z}{\sigma} \right), \quad (3.1)$$

donde  $z > 0$ ,  $\sigma > 0$  es un parámetro de escala y  $0 < \alpha < 1$  es un parámetro de forma.

La Figura 3.1 muestra algunos gráficos de la distribución MEG considerando diferentes valores del parámetro  $\alpha$ .

El Cuadro 3.1 muestra  $P(Z > z)$  para diferentes valores de  $z$  en la distribución MEG, ilustrando que las colas son más pesadas a medida que disminuye el parámetro  $\alpha$ .

Cuadro 3.1: Comparación de colas

Distribución	$P(Z > 5)$	$P(Z > 10)$	$P(Z > 15)$
MEG(1,0.9)	0.06493	0.04284	0.03310
MEG(1,0.7)	0.26006	0.19377	0.16107
MEG(1,0.3)	0.79235	0.73030	0.69215

### 3.2.2. Propiedades

Las siguientes proposiciones muestran la forma de representar la distribución MEG. La primera proposición muestra que la distribución MEG es el producto de una mezcla de escalas entre las distribuciones G y Beta.

**Proposición 3.1** *Sea  $Z|U = u \sim G(\sigma u^{-1}, \alpha)$  y  $U \sim Beta(\alpha, 1)$  entonces  $Z \sim MEG(\sigma, \alpha)$*

**Demostración.** Calculando la integral directamente, tenemos

$$f_Z(z; \sigma, \alpha) = \int_{Rec(u)} f(z|u) f_U(u) du = \int_0^1 \frac{(\sigma u^{-1})^\alpha z^{-\alpha}}{B(1-\alpha, \alpha)(\sigma u^{-1} + z)} \alpha u^{\alpha-1} du,$$

de donde el resultado se obtiene haciendo el siguiente cambio de variable  $t = \sigma + uz$ .

□

**Proposición 3.2** *Sea  $X \sim G(\sigma, \alpha)$  y  $Y \sim Beta(\alpha, 1)$  sean variables aleatorias independientes. Entonces  $Z = \frac{X}{Y} \sim MEG(\sigma, \alpha)$ .*

**Demostración.** Utilizando el método del jacobiano tenemos

$$\left. \begin{array}{l} Z = XY^{-1} \\ V = Y \end{array} \right\} \Rightarrow \left. \begin{array}{l} X = ZV \\ Y = V \end{array} \right\} \Rightarrow J = \begin{vmatrix} \frac{\partial X}{\partial Z} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial Z} & \frac{\partial Y}{\partial V} \end{vmatrix} = \begin{vmatrix} v & z \\ 0 & 1 \end{vmatrix} = v$$

$$f_{Z,V}(z, v) = |J|f_{X,Y}(zv, v)$$

$$f_{Z,V}(z, v) = vf_X(zv)f_Y(v), \quad z > 0, \quad 0 < v < 1,$$

luego, marginalizando con respecto a la variable  $V$ , obtenemos la función de densidad asociada con  $Z$ .

$$\begin{aligned} f_Z(z; \sigma, \alpha) &= \frac{\alpha}{\sigma B(\bar{\alpha}, \alpha)} \int_0^1 \frac{\left(\frac{zv}{\sigma}\right)^{-\alpha} v^\alpha}{1 + \frac{zv}{\sigma}} dv, \\ &= \frac{\alpha \sigma^{\alpha-1}}{B(\bar{\alpha}, \alpha) z^\alpha} \int_0^1 \frac{1}{1 + \frac{zv}{\sigma}} dv, \end{aligned}$$

realizando el siguiente cambio de variables  $u = 1 + \frac{zv}{\sigma}$ , se obtiene el resultado.  $\square$

Se pueden utilizar dos algoritmos para generar números aleatorios a partir del modelo MEG, como se muestra a continuación. El Algoritmo 1 utiliza la Proposición 3.1 y el método de composición (ver Tanner, 1996); El Algoritmo 2 utiliza la representación dada en la Proposición 3.2.

---

**Algorithm 1** for simulating  $Z \sim MEG(\sigma, \alpha)$  can proceed as follows.

---

1: Generate  $U \sim Beta(\alpha, 1)$ .

2: Compute  $Z|U = u \sim G(\sigma u^{-1}, \alpha)$ .

---

**Proposición 3.3** Sea  $Z \sim MEG(\sigma, \alpha)$ . Entonces, la fda de  $Z$  es dada por

$$F_Z(z; \sigma, \alpha) = I_y(\bar{\alpha}, \alpha) - \frac{\sigma^\alpha z^{-\alpha}}{B(\bar{\alpha}, \alpha)} \log\left(1 + \frac{z}{\sigma}\right), \quad z > 0, \quad (3.2)$$

---

**Algorithm 2** for simulating  $Z \sim MEG(\sigma, \alpha)$  can proceed as follows.

---

- 1: Generate  $V \sim Beta(1 - \alpha, \alpha)$ .
  - 2: Compute  $X = \frac{\sigma V}{1-V}$ .
  - 3: Generate  $Y \sim Beta(\alpha, 1)$ .
  - 4: Compute  $Z = \frac{X}{Y}$ .
- 

donde  $\sigma > 0$ ,  $0 < \alpha < 1$ ,  $y = \frac{z}{\sigma+z}$  y  $I_y(\cdot, \cdot)$  es la función beta incompleta regularizada.

**Demostración.** Aplicando directamente la definición de la fda, tenemos

$$F_Z(z; \sigma, \alpha) = \frac{\sigma^\alpha}{B(\bar{\alpha}, \alpha)} \int_0^z t^{-(\alpha+1)} \log \left( 1 + \frac{t}{\sigma} \right) dt,$$

y aplicando el método de integración por partes y considerando la transformación  $u = \log \left( 1 + \frac{t}{\sigma} \right)$ , se obtiene el resultado.  $\square$

La función de supervivencia para una variable aleatoria MEG viene dada por

$$\bar{F}_Z(t; \sigma, \alpha) = 1 - I_y(\bar{\alpha}, \alpha) + \frac{\sigma^\alpha t^{-\alpha}}{B(\bar{\alpha}, \alpha)} \log \left( 1 + \frac{t}{\sigma} \right),$$

donde  $y = \frac{t}{\sigma+t}$ . La función de riesgo para una variable aleatoria MEG es dada por

$$h(t; \sigma, \alpha) = \frac{\alpha \sigma^\alpha t^{-(\alpha+1)} \log \left( 1 + \frac{t}{\sigma} \right)}{B(\bar{\alpha}, \alpha) - B\left(\frac{t}{\sigma+t}; \bar{\alpha}, \alpha\right) + \sigma^\alpha t^{-\alpha} \log \left( 1 + \frac{t}{\sigma} \right)}, \quad t > 0.$$

La Figura 4.3 muestra la forma de la función de riesgo para diferentes valores de  $\alpha$ , considerando  $\sigma = 1$ .

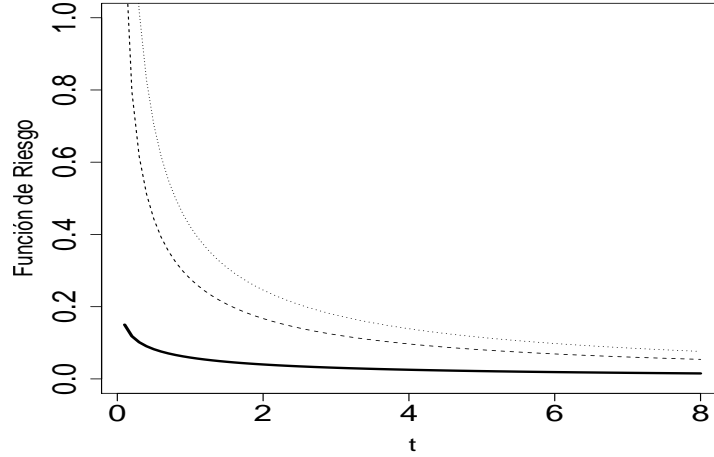


Figura 3.2: Ejemplos de la función  $h(t)$ , para MEG(1,0.3) (línea continua), MEG(1,0.7) (línea discontinua) y MEG(1,0.9) (línea punteada)

**Proposición 3.4** Sea  $T \sim \text{MEG}(\sigma, \alpha)$ . Entonces, la función de riesgo de  $T$  es decreciente para todo  $t > 0$ .

**Demostración.** Usando el Teorema, ítem (b), dada en Glaser (1980) tenemos que

$$\eta(t) = -\frac{f'(t)}{f(t)} = \frac{\alpha + 1}{t} - \frac{1}{(\sigma + t) \log\left(1 + \frac{t}{\sigma}\right)},$$

donde  $f(t)$  es la fdp dada en (3.1). Luego, derivando la función  $\eta(t)$  con respecto a  $t$ , obtenemos

$$\eta'(t) = -\left(\frac{\alpha + 1}{t^2} - \frac{1}{(\sigma + t)^2 \log\left(1 + \frac{t}{\sigma}\right)} - \frac{1}{(\sigma + t)^2 \log^2\left(1 + \frac{t}{\sigma}\right)}\right) < 0, \quad \forall t > 0,$$

de donde se concluye el resultado.  $\square$

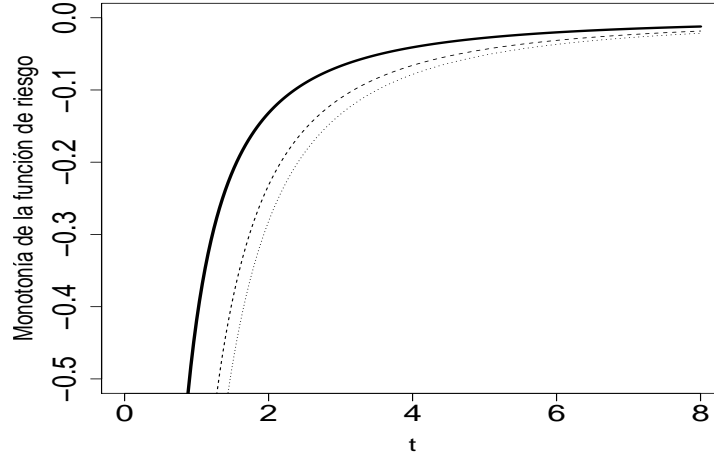


Figura 3.3: Ejemplos de  $\eta'(t)$ , para  $\alpha = 0.3$  (línea continua),  $\alpha = 0.7$  (línea discontinua),  $\alpha = 0.9$  (línea punteada).

La Figura 3.3 muestra, fijando  $\sigma = 1$ , que la función  $\eta'(t)$  es siempre negativa. Para otros valores diferentes de  $\sigma$  y  $\alpha$  el resultado es el mismo.

**Proposición 3.5** *Los momentos  $r$ -ésimos de la variable aleatoria  $Z \sim MEG(1, \alpha)$  no existen para  $r \geq \alpha$ .*

**Demostración.** Utilizando integración por partes, con  $u = \log(1 + x)$ , tenemos

$$\begin{aligned} \mathbb{E}(Z^r) &= \frac{\alpha\sigma^\alpha}{B(\alpha, \bar{\alpha})} \int_0^\infty z^{r-\alpha-1} \log(1+z) dz \\ &= \frac{\alpha\sigma^\alpha z^{r-\alpha} \log(1+z)}{(r-\alpha)B(\bar{\alpha}, \alpha)} \Big|_0^\infty - \frac{\alpha\sigma^\alpha}{(r-\alpha)B(\bar{\alpha}, \alpha)} \int_0^\infty \frac{z^{r-\alpha}}{1+z} dz. \end{aligned}$$

La expresión del lado izquierdo es igual a cero si  $r < \alpha$  y no existe si  $r \geq \alpha$ ; la integral del lado derecho se analizó en la Proposición 1 (e) del trabajo de Olmos et al. (2022) y no converge. Por lo tanto, la integral completa diverge y esto significa que no existen los  $r$ -ésimos momentos para la variable aleatoria  $Z \sim MEG(1, \alpha)$ . Por lo tanto, la integral completa diverge y esto significa que no existen  $r$ -ésimos momentos para la variable aleatoria  $Z \sim MEG(1, \alpha)$ .  $\square$

### 3.2.3. Cola de la distribución

Como hemos indicado en el Capítulo 2, el modelo Pareto es una distribución con una cola derecha pesada y por esta característica se utiliza en seguros para modelar datos de cuantía o de daños; por lo tanto, el tamaño de la cola de una distribución es fundamental para modelar valores extremos. Así, el concepto de cola pesada es fundamental para este y otros escenarios financieros. Se sabe que cualquier distribución de probabilidad, especificada por su fda  $F(x)$ , tiene una cola derecha pesada (ver Rolski et al., 1999) si  $\limsup_{x \rightarrow \infty} (-\log(\bar{F}(x)))/x = 0$ .

Un tema importante en la teoría de valores extremos es la variación regular (ver Bingham, 1987). La siguiente proposición establece que la función de supervivencia de la distribución MEG es una distribución de variación regular.

**Proposición 3.6** *La función de supervivencia de la variable aleatoria  $X \sim MEG(\sigma, \alpha)$*

es una función de supervivencia con colas de variación regular.

**Demostración.** Aplicando la Definición 2.3 y utilizando la regla de L'Hôpital, tenemos que

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(x)} = \lim_{x \rightarrow \infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\alpha} \lim_{x \rightarrow \infty} \frac{\log\left(1 + \frac{tx}{\sigma}\right)}{\log\left(1 + \frac{x}{\sigma}\right)} = t^{-\alpha+1} \lim_{x \rightarrow \infty} \frac{\sigma + x}{\sigma + tx},$$

de donde calculando este último límite se obtiene el resultado.  $\square$  Como la distribución MEG es de colas de variación regular, entonces es de cola pesada (ver Feller, 1971).

### 3.3. Inferencia

En esta Sección estimamos los parámetros de la distribución MEG utilizando el método de MV, se discute un estudio de simulación y la estimación asintótica de los estimadores de MV.

#### 3.3.1. Estimación de MV

Para una muestra aleatoria  $z_1, \dots, z_n$  derivado de la distribución  $\text{MEG}(\sigma, \alpha)$ , la función log-verosimilitud se puede escribir como

$$\ell(\sigma, \alpha) = c(\sigma, \alpha) - (\alpha + 1) \sum_{i=1}^n \log(z_i) + \sum_{i=1}^n \log(\log(1 + z_i/\sigma)), \quad (3.3)$$

donde  $c(\sigma, \alpha) = n \log(\alpha) + n\alpha \log(\sigma) - n \log(B(\bar{\alpha}, \alpha))$ .

Al igualar a cero las primeras derivadas con respecto a cada parámetro, obtenemos

$$\sum_{i=1}^n \frac{z_i}{(\sigma + z_i) \log(1 + z_i/\sigma)} = n\alpha, \quad (3.4)$$

$$\sum_{i=1}^n \log(z_i) + n\psi(\alpha) - n\psi(1 - \alpha) = \frac{n}{\alpha} + n \log(\sigma).$$

De (3.4) obtenemos

$$\hat{\alpha}(\hat{\sigma}) = \frac{1}{n} \sum_{i=1}^n \frac{z_i}{(\hat{\sigma} + z_i) \log(1 + z_i/\hat{\sigma})}, \quad (3.5)$$

y el estimador de MV para  $\sigma$  ( $\hat{\sigma}$ ) se obtiene resolviendo numéricamente la siguiente ecuación:

$$\frac{1}{n} \sum_{i=1}^n \log(z_i) - \psi(1 - \hat{\alpha}(\hat{\sigma})) + \psi(\hat{\alpha}(\hat{\sigma})) = \frac{n}{\hat{\alpha}(\hat{\sigma})} + n \log(\hat{\sigma}). \quad (3.6)$$

El estimador  $\hat{\sigma}$  es la solución de la ecuación (3.6), y al sustituirlo en (3.5) obtenemos  $\hat{\alpha}$ . La solución de esta ecuación (3.6) se puede obtener mediante procedimientos numéricos como el algoritmo de Newton-Raphson. Alternativamente, estas estimaciones se pueden encontrar maximizando directamente la función log-verosimilitud, dada en (3.3), utilizando la subrutina `optim` del software R (ver R Core Team, 2021).

### 3.3.2. Estudio de simulación

Para evaluar la efectividad del enfoque propuesto, realizamos un estudio de simulación para analizar el procedimiento de estimación de los parámetros  $\alpha$  y  $\sigma$  de la

distribución MEG. El estudio consistió en simular 1000 muestras del modelo MEG con tres tamaños de muestra diferentes:  $n = 100, 200$  y  $500$ . El Cuadro 3.2 muestra el SE, EE, ECME y PC. Del Cuadro 3.2, concluimos que las estimaciones de MV son estables. El sesgo se reduce al aumentar el tamaño de la muestra. Los PC también se aproxima al 95 % nominal a medida que aumenta  $n$ .

### 3.3.3. Matriz de información de Fisher

Consideremos ahora  $Z \sim MEG(\sigma, \alpha)$ , para una sola observación  $z$  de  $Z$ , la función log-verosimilitud para  $\boldsymbol{\theta} = (\sigma, \alpha)$  es dada por

$$\ell(\boldsymbol{\theta}) = \log(\alpha) + \alpha \log(\sigma) - \log(B(\bar{\alpha}, \alpha)) - (\alpha + 1) \log(z) + \log(\log(1 + z/\sigma)).$$

Las primera y segunda derivadas parciales de la función log-verosimilitud se encuentran en el Apéndice B. Se puede mostrar que la matriz de información de Fisher, para la distribución MEG viene dada por

$$I_F(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\alpha a_\alpha}{\sigma^2 B(\bar{\alpha}, \alpha)} & -\frac{1}{\sigma} \\ -\frac{1}{\sigma} & \psi'(\alpha) + \psi'(\bar{\alpha}) + \frac{1}{\alpha^2} \end{pmatrix},$$

donde  $a_\alpha = \int_0^\infty \frac{u^{1-\alpha}}{(1+u)^2 \log(1+u)} du$ .

Por lo tanto, para muestras grandes, el estimador de MV,  $\hat{\boldsymbol{\theta}}$ , de  $\boldsymbol{\theta}$  es asintóticamente

Cuadro 3.2: SE, EE, ECME y 95 % PC para los estimadores de MV de  $\sigma$  y  $\alpha$  en la distribución MEG con diferentes combinaciones en los valores de los parámetros.

valor verdadero		$n = 100$				$n = 200$				$n = 500$					
$\sigma$	$\alpha$ par.	SE	EE	ECME	PC	SE	EE	ECME	PC	SE	EE	ECME	PC		
1	0.1	$\alpha$	0.0068	0.0093	0.0116	0.9190	0.0071	0.0066	0.0097	0.8460	0.0062	0.0053	0.0083	0.7990	
		$\sigma$	0.4032	2.2812	4.0426	0.9350	0.3730	1.2368	1.6858	0.9640	0.3860	0.8955	1.1496	0.9710	
	0.25	$\alpha$	0.0071	0.0093	0.0120	0.9250	0.0065	0.0065	0.0093	0.8620	0.0065	0.0053	0.0083	0.7960	
		$\sigma$	4.0440	22.4201	34.7855	0.9390	3.5141	11.7503	15.3452	0.9600	3.5002	8.8431	11.3256	0.9780	
	0.5	$\alpha$	0.0030	0.0265	0.0273	0.9630	0.0012	0.0185	0.0194	0.9470	-0.0002	0.0149	0.0153	0.9420	
		$\sigma$	0.0919	0.9833	1.2023	0.9110	0.0365	0.5801	0.6274	0.9300	0.0287	0.4429	0.4736	0.9320	
	0.75	$\alpha$	0.0017	0.0266	0.0285	0.9580	0.0018	0.0184	0.0185	0.9570	0.0011	0.0150	0.0158	0.9550	
		$\sigma$	1.4353	10.1717	12.4871	0.9090	0.4046	5.7823	6.4739	0.9360	0.2705	4.5401	5.3019	0.9390	
	0.9	$\alpha$	0.0000	0.0645	0.0663	0.9190	0.0006	0.0470	0.0486	0.9290	-0.0009	0.0389	0.0378	0.9390	
		$\sigma$	0.0090	1.3658	2.1311	0.8800	0.0027	0.7859	0.9812	0.8930	-0.0025	0.5698	0.6251	0.9090	
	10	0.1	$\alpha$	0.0031	0.0643	0.0629	0.9210	0.0015	0.0471	0.0488	0.9270	0.0013	0.0388	0.0387	0.9390
			$\sigma$	0.2496	12.7439	24.3474	0.8950	0.4281	8.0812	10.7496	0.9060	0.3605	5.9643	6.8460	0.9230
0.25		$\alpha$	-0.0026	0.0381	0.0423	0.9560	-0.0023	0.0262	0.0271	0.9540	-0.0017	0.0213	0.0224	0.9590	
		$\sigma$	-0.0281	0.6856	0.7240	0.8880	-0.0512	0.4548	0.4745	0.9130	-0.0301	0.3663	0.3837	0.9210	
0.5		$\alpha$	-0.0033	0.0381	0.0420	0.9450	-0.0016	0.0261	0.0265	0.9470	-0.0011	0.0211	0.0217	0.9530	
		$\sigma$	-0.7347	6.8163	7.3233	0.8840	-0.3695	4.6192	4.8007	0.9290	-0.2962	3.6364	3.8096	0.9060	
0.75		$\alpha$	0.0001	0.0116	0.0122	0.9550	-0.0005	0.0081	0.0082	0.9470	-0.0005	0.0066	0.0069	0.9470	
		$\sigma$	-0.0263	0.7099	0.7945	0.8810	-0.0256	0.4557	0.4643	0.9150	-0.0227	0.3672	0.3941	0.9230	
0.9		$\alpha$	-0.0016	0.0117	0.0120	0.9600	-0.0002	0.0081	0.0083	0.9400	-0.0004	0.0066	0.0066	0.9480	
		$\sigma$	-0.7632	6.7365	7.6937	0.8640	-0.4858	4.5617	4.9562	0.9070	-0.1317	3.6848	3.8046	0.9260	

normal bivariado, es decir,

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{\mathcal{L}} N_2(\mathbf{0}, I_F(\boldsymbol{\theta})^{-1}),$$

la varianza asintótica del estimador MV  $\hat{\theta}$  es, por lo tanto, la inversa de la matriz de información de Fisher. Dado que los parámetros son desconocidos, generalmente se considera la matriz de información observada, donde los parámetros desconocidos se estiman mediante MV.

### 3.4. Aplicación a datos de índices de siniestralidad

En esta Sección, analizamos una aplicación a un conjunto de datos reales, utilizando los modelos de Pareto y LG2 para compararlos con el modelo MEG. Para comparar los modelos, utilizamos los criterios de información AIC y BIC. El conjunto de datos representa los índices de accidentes (datos anuales en miles de millones de dólares) para seguros contra terremotos en California desde 1971 hasta 1993 para valores mayores que cero. Los datos se proporcionan en Embrechts et al. (1999). Los estadísticos descriptivos de estos datos se muestran en el Cuadro 3.3. El boxplot de la Figura 3.4 muestra la existencia de dos datos extremos. Estos datos atípicos hacen que la cola derecha sea más pesada.

Cuadro 3.3: Estadísticas descriptivas de los datos de índices de siniestralidad

$n$	Mediana	Media	Varianza	CS	CK
19	5.0	15.72	889.582	3.204	12.648

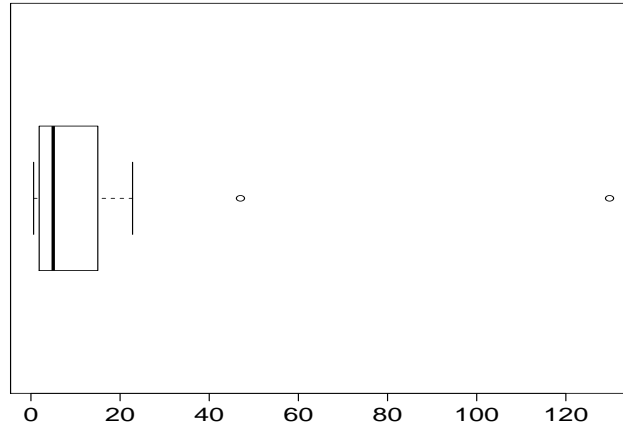


Figura 3.4: Boxplot para los datos de índices de siniestralidad

El Cuadro 3.4 muestra las estimativas de MV junto con los errores estándar para los parámetros de los modelos MEG, Pareto y LG2, así como los valores de los criterios AIC y BIC para cada modelo. Para los modelos Pareto y LG2, se utilizó el valor  $\alpha = 0.1$ .

Se observa que los valores más bajos de los criterios AIC y BIC corresponden al modelo MEG, lo que significa que el modelo MEG se ajusta mejor a los datos que los modelos Pareto y LG2. Por esta razón, el modelo MEG puede considerarse una alternativa al modelo de Pareto. La Figura 3.5 muestra el gráfico QQ-plot de los residuos cuantiles (ver Dunn and Smyth, 1996) del modelo MEG, donde muestra que el modelo MEG es apropiado para ajustar los datos de índices de siniestralidad, ya

Cuadro 3.4: Estimativas de MV para los datos de índices de siniestralidad con sus correspondientes errores estándar, valores AIC y BIC.

Modelos	Estimativas de MV	AIC	BIC
MEG( $\sigma, \alpha$ )	$\hat{\sigma} = 2.0190(0.793), \hat{\alpha} = 0.573(0.069)$	<b>151.815</b>	<b>153.704</b>
Pareto( $\alpha, \beta$ )	$\hat{\beta} = 0.249(0.057)$	157.878	158.822
LG2( $\alpha, \beta$ )	$\hat{\beta} = 0.688(0.112)$	176.332	177.277

que los cuantiles muestrales son muy próximos a los teóricos.

### 3.5. Conclusiones

Este Capítulo presenta un estudio de la distribución MEG; se analizan algunas de sus propiedades, se estiman los parámetros mediante el método MV y se presenta una aplicación a datos reales. La distribución MEG es el producto de una mezcla de escala de las distribuciones G y Beta. Esta distribución tiene dos parámetros, lo que la convierte en una alternativa interesante a otros modelos de dos parámetros utilizados, por ejemplo, en estadística actuarial. El modelo MEG se presenta como una opción viable para ajustar datos con observaciones atípicas. Otras características del modelo MEG son:

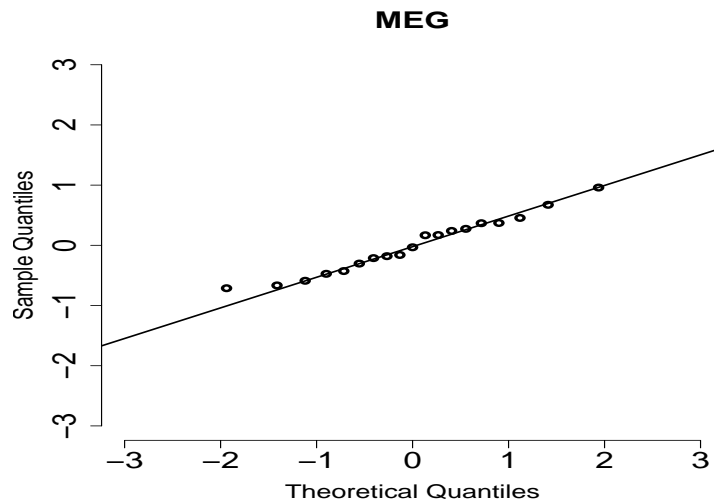


Figura 3.5: QQ-plots de los residuos cuantiles del modelo MEG.

- La cola derecha del modelo MEG es más pesada para valores pequeños del parámetro  $\alpha$ .
- Existen dos representaciones del modelo MEG, dadas en las Proposiciones 3.1 y 3.2.
- La distribución MEG tiene la cola derecha pesada, ya que es de variación regular.
- La fdp, la fda y la función de riesgo son explícitas y están representadas por funciones conocidas.
- El análisis de la aplicación muestra que la distribución MEG es una buena

candidata para modelar datos de índices de siniestralidad, con un rendimiento mejor que dos distribuciones bien conocidas, Pareto y LG2.

# Capítulo 4

## Distribución Hipergeométrica

## Gaussiana Gleser

### 4.1. Introducción

Como hemos comentado en los Capítulos anteriores, las distribuciones de cola derecha pesada se han utilizado en el modelado de datos en distintas áreas del conocimiento. Usualmente en el modelamiento de estos datos se utilizan distribuciones con soporte positivo y que tengan éstas características, Aban et al. (2006) contiene una lista de áreas donde se consideran aplicables las distribuciones de cola pesada. Como indicamos en el Capítulo 2, Andrews and Mallows (1974) desarrolló un estudio

de mezclas de escalas de la distribución normal, obteniendo algunas distribuciones con colas pesadas. Entonces podemos realizar una mezcla de escala con dos distribuciones con soportes positivo. La distribución resultante se puede expresar de la siguiente manera

$$f_Z(z) = \int_0^\infty f_{Z|U=u}(z; \kappa(u)) f_U(u) du,$$

donde  $f_{Z|U=u}$  es la distribución condicional de la variable aleatoria  $Z$  dado  $U = u$  con soporte positivo, donde  $\kappa(u)$  es una función positiva de la variable aleatoria  $U$  con fdp  $f_U$  y que influye en la escala de la variable aleatoria  $Z$ .

Una función necesaria en este Capítulo es la función trascendente de Lerch ( $\Phi$ ), y que se puede expresar como:

$$\Phi(-z, s, a) = \frac{1}{\Gamma(s)} \int_0^1 \frac{x^{a-1} \log^{s-1} \left(\frac{1}{x}\right)}{1 + zx} dx, \quad (4.1)$$

donde  $a > 0$ ,  $s > 0$  y  $z > -1$ .

El Objetivo de este Capítulo es extender las distribuciones G y MEG, y que se genera en base a una mezcla de escala entre las distribuciones G y Beta (con parámetros distintos), similar a la distribución MEG, y que contenga como casos particulares a las distribuciones G y MEG. Estudiar algunas de sus propiedades y mostrar que puede ser una alternativa para el modelado de datos que provienen de las ciencias actuariales, como por ejemplo, de ingresos por familia.

El Capítulo se desarrolla de la siguiente manera. En la Sección 4.2 se entrega la fdp y sus propiedades básicas de la nueva distribución. En la Sección 4.3 se realiza inferencia por el método de MV, un estudio de simulación y su matriz de información de Fisher. En la Sección 4.4 se muestra una aplicación a datos de ingresos por familia y en la Sección 4.5 se entregan algunas conclusiones.

## 4.2. Función Densidad y Propiedades

En esta Sección introducimos la distribución hipergeométrica Gaussiana Gleser (HGG), y estudiamos algunas de sus propiedades.

### 4.2.1. Función densidad

La siguiente definición muestra la fdp de la distribución HGG, como algunos casos particulares.

**Definición 4.1** *Sea  $Z \sim HGG(\sigma, \alpha, \beta)$ . Entonces, la fdp de  $Z$  viene dada por*

$$f_Z(z; \sigma, \alpha, \beta) = \frac{\beta B(\bar{\alpha} + \beta, 1)}{\sigma^{\bar{\alpha}} B(\bar{\alpha}, \alpha)} z^{-\alpha} {}_2F_1\left(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -\frac{z}{\sigma}\right), \quad z > 0, \quad (4.2)$$

donde  $\sigma > 0$  es un parámetro de escala,  $0 < \alpha < 1$  y  $\beta > 0$  son parámetros de formas y  ${}_2F_1$  es la función hipergeométrica de Gauss dada en (2.4).

**Comentario 4.1** 1. Si  $\beta < \alpha$ , entonces la distribución HGG se puede representar

como

$$f_Z(z; \sigma, \alpha, \beta) = \frac{\beta \sigma^\beta z^{-(\beta+1)}}{B(\bar{\alpha}, \alpha)} B\left(\frac{z}{\sigma+z}, \bar{\alpha} + \beta, \alpha - \beta\right), \quad z > 0, \quad (4.3)$$

2. Si  $\beta = \alpha$ , entonces se obtiene la distribución MEG( $\sigma, \alpha$ ), que fue estudiada en el Capítulo anterior (ver también Olmos et al., 2024).

3. Si  $\frac{z}{\sigma} < 1$  y utilizando la función trascendental de Lerch dada en (4.1), se obtiene

$$f_Z(z; \sigma, \alpha, \beta) = \frac{\beta \sigma^{-\bar{\alpha}} z^{-\alpha}}{B(\bar{\alpha}, \alpha)} \Phi\left(-\frac{z}{\sigma}, 1, \bar{\alpha} + \beta\right), \quad z > 0, \quad (4.4)$$

La Figura 4.1 muestra los gráficos de la distribución HGG para  $\sigma = 1$ ,  $\alpha = 0.5$  y dos valores del parámetro  $\beta$ .

El Cuadro 4.1 muestra  $P(Z > z)$  para diferentes valores de  $z$  de la distribución HGG. Se puede ver que la distribución HGG tiene la cola derecha más pesada cuando el valor del parámetro  $\beta$  decrece.

### 4.2.2. Propiedades

En ésta Subsección mostramos algunas propiedades de la distribución HGG.

**Proposición 4.1** Sea  $Z|U = u \sim G(\sigma u^{-1}, \alpha)$  y  $U \sim \text{Beta}(\beta, 1)$  entonces  $Z \sim \text{HGG}(\sigma, \alpha, \beta)$ .

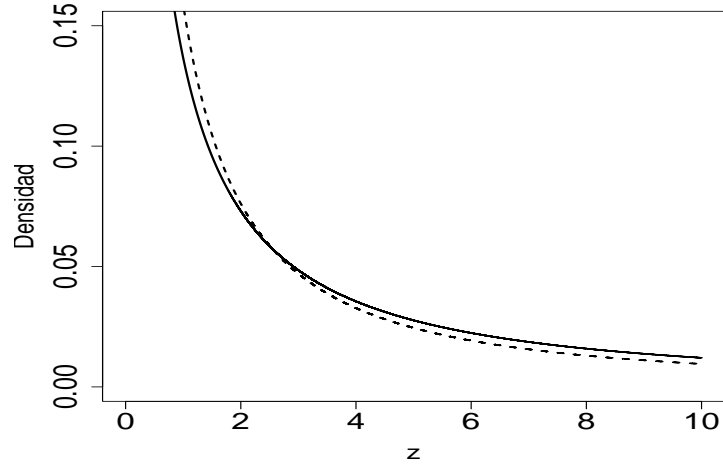


Figura 4.1: Ejemplos de la distribución  $HGG(1,0.5,1)$  (línea continua) y  $HGG(1,0.5,10)$  (línea discontinua)

Cuadro 4.1: Comparación de la cola para la distribución HGG para diferentes valores de  $\beta$

Distribución	$P(Z > 4)$	$P(Z > 5)$	$P(Z > 6)$
$HGG(1,0.5,1)$	0.437	0.406	0.381
$HGG(1,0.5,3)$	0.341	0.311	0.288
$HGG(1,0.5,10)$	0.308	0.280	0.258

**Demostración.** Calculando directamente la integral se tiene

$$\begin{aligned}
f_Z(z; \sigma, \alpha, \beta) &= \int_{\text{Rec}(u)} f_{Z|U}(z) f_U(u) du, \\
&= \int_0^1 \frac{(\sigma u^{-1})^\alpha z^{-\alpha}}{B(\bar{\alpha}, \alpha)(\sigma u^{-1} + z)} \beta u^{\beta-1} du, \\
&= \frac{\beta z^{-\alpha}}{B(\bar{\alpha}, \alpha) \sigma^{\bar{\alpha}}} \int_0^1 \frac{u^{\beta-\alpha}}{1 + \frac{z}{\sigma} u} du,
\end{aligned}$$

e donde el resultado se obtiene utilizando (2.4).  $\square$

**Proposición 4.2** *Sea  $X \sim G(1, \alpha)$  y  $Y \sim \text{Beta}(\beta, 1)$  son variables aleatorias independientes. Entonces  $Z = \sigma \frac{X}{Y} \sim \text{HGG}(\sigma, \alpha, \beta)$ .*

**Demostración.** Utilizando la representación estocástica  $Z = \sigma \frac{X}{Y}$  y la transformación jacobiana, tenemos:

$$\left. \begin{array}{l} Z = \sigma \frac{X}{Y} \\ W = Y \end{array} \right\} \Rightarrow \begin{array}{l} X = \frac{ZW}{\sigma} \\ Y = W \end{array} \Rightarrow J = \begin{vmatrix} \frac{\partial X}{\partial Z} & \frac{\partial X}{\partial W} \\ \frac{\partial Y}{\partial Z} & \frac{\partial Y}{\partial W} \end{vmatrix} = \begin{vmatrix} \frac{w}{\sigma} & \frac{z}{\sigma} \\ 0 & 1 \end{vmatrix} = \frac{w}{\sigma}$$

$$f_{Z,W}(z, w) = |J| f_{X,Y} \left( \frac{zw}{\sigma}, w \right)$$

$$f_{Z,W}(z, w) = \frac{w}{\sigma} f_X \left( \frac{zw}{\sigma} \right) f_Y(w), \quad z > 0, \quad 0 < w < 1$$

luego marginalizando con respecto a la variable  $W$  se obtiene la función de densidad asociada a  $Z$ .

$$f_Z(z; \sigma, \alpha, \beta) = \frac{\beta \sigma^{\alpha-1} z^{-\alpha}}{B(\bar{\alpha}, \alpha)} \int_0^1 \frac{w^{\beta-\alpha}}{1 + \frac{zw}{\sigma}} dw$$

el resultado se obtiene utilizando (2.4).  $\square$

**Proposición 4.3** Sea  $Z \sim HGG(\sigma, \alpha, \beta)$ . Si  $\beta \rightarrow \infty$  entonces  $Z$  converge en distribución a una variable aleatoria  $Z \sim G(\sigma, \alpha)$ .

**Demostración.** Sea  $Z \sim HGG(\sigma, \alpha, \beta)$ . Entonces, por la representación dada en la Proposición 4.2 se tiene que  $Z = \frac{\sigma X}{Y}$ . Primero estudiamos la convergencia en probabilidad de  $Y$ .

Como  $Y \sim \text{Beta}(\beta, 1)$ , se tiene que  $\mathbb{E}[(Y - 1)^2] = \frac{2}{(\beta+1)(\beta+2)}$ , si  $\beta \rightarrow \infty$ , entonces  $\mathbb{E}[(Y - 1)^2] = \frac{2}{(\beta+1)(\beta+2)} \rightarrow 0$ , implica que  $Y \xrightarrow{\mathbb{P}} 1$ . Aplicando el lema de Slutsky se obtiene el resultado.  $\square$

**Comentario 4.2** El resultado de la Proposición 4.1 nos muestra que esta distribución HGG es producto de una mezcla de escala entre la distribución  $G$  y la distribución Beta, por otro lado la Proposición 4.2 nos indica que cuando  $\beta \rightarrow \infty$  en la distribución HGG se obtiene la distribución  $G$ . La Figura 4.2 muestra la relación que existe entre las tres distribuciones, se puede observar que la distribución HGG contiene como casos particulares a las distribuciones  $G$  y MEG.

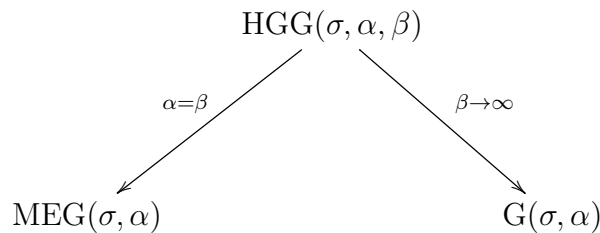


Figura 4.2: Casos particulares de la distribución de HGG.

**Proposición 4.4** Sea  $Z \sim HGG(\sigma, \alpha, \beta)$ . Entonces, la fda de  $Z$  es dada por

$$F_Z(z; \sigma, \alpha, \beta) = I_y(\bar{\alpha}, \alpha) - \frac{z}{\beta} f_Z(z; \sigma, \alpha, \beta), \quad z > 0, \quad (4.5)$$

donde  $\sigma > 0$ ,  $0 < \alpha < 1$ ,  $\beta > 0$ ,  $y = \frac{z}{\sigma+z}$  y  $I_y(\cdot, \cdot)$  es la función beta incompleta regularizada.

**Demostración.** Aplicando en forma directa la definición de fda se tiene

$$\begin{aligned} F_Z(z; \sigma, \alpha, \beta) &= \frac{\beta B(\bar{\alpha} + \beta, 1)}{\sigma^\alpha B(\bar{\alpha}, \alpha)} \int_0^z t^{-\alpha} {}_2F_1\left(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -\frac{t}{\sigma}\right) dt, \\ &= \frac{\beta}{\sigma^\alpha B(\bar{\alpha}, \alpha)} \int_0^z t^{-\alpha} \int_0^1 \frac{v^{\bar{\alpha} + \beta - 1}}{1 + \frac{tv}{\sigma}} dv dt, \\ &= \frac{\beta}{B(\bar{\alpha}, \alpha)} \int_0^1 v^{\beta-1} \left( \int_0^{\frac{vz}{\sigma}} \frac{w^{-\alpha}}{1+w} dw \right) dv \end{aligned}$$

aplicando el método de integración por partes y considerando  $u = \int_0^{\frac{vz}{\sigma}} \frac{w^{-\alpha}}{1+w} dw$ , se obtiene el resultado.  $\square$

**Corolario 4.1** Sea  $Z \sim HGG(\sigma, \alpha, \beta)$ , la función de supervivencia para una variable aleatoria que se distribuye HGG y es dada por

$$\bar{F}_Z(t; \sigma, \alpha, \beta) = 1 - I_y(\bar{\alpha}, \alpha) + \frac{t}{\beta} f_Z(t; \sigma, \alpha, \beta).$$

Por otro lado la función de riesgo,  $h(t)$ , para una variable aleatoria que se distribuye HGG es dada por

$$h(t; \sigma, \alpha, \beta) = \frac{f_Z(t; \sigma, \alpha, \beta)}{1 - I_y(\bar{\alpha}, \alpha) + \frac{t}{\beta} f_Z(t; \sigma, \alpha, \beta)}, \quad t > 0,$$

donde  $y = \frac{t}{\sigma+t}$ .

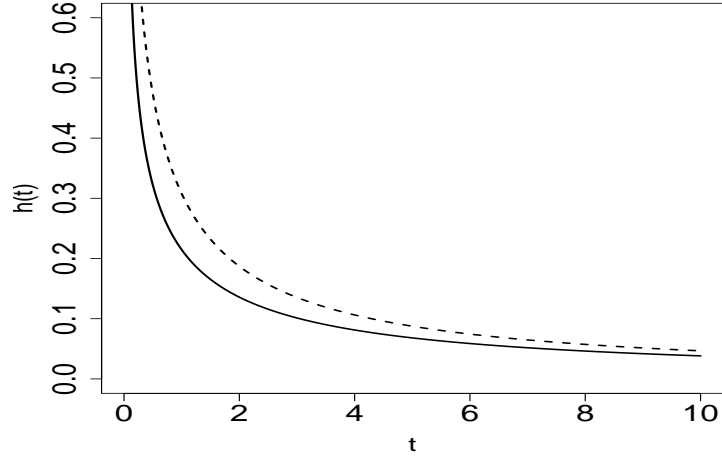


Figura 4.3: Ejemplos de  $h(t)$ , cuando HGG(1,0.5,1) (línea continua) y HGG(1,0.5,10) (línea discontinua)

La Figura 4.3 muestra la forma que tiene la función de riesgo para distintos valores de  $\beta$ , considerando  $\sigma = 1$  y  $\alpha = 0.5$ . Observamos que la función de riesgo es decreciente, esto es un indicador que la distribución HGG es de cola derecha pesada.

En el estudio de simulación como en la estimación de los parámetros por el método de MV consideramos la distribución HGG sólo con dos parámetros, es decir, modificando la Proposición 4.1 como sigue:  $Z|U = u \sim G(u^{-1}, \alpha)$  y  $U \sim Beta(\beta, 1)$ , resulta la siguiente fdp de la variable aleatoria  $Z$ .

$$f_Z(z; \beta, \alpha) = \frac{\beta B(\bar{\alpha} + \beta, 1)}{B(\bar{\alpha}, \alpha)} z^{-\alpha} {}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z), \quad z > 0, \quad (4.6)$$

donde  $\beta > 0$  y  $0 < \alpha < 1$  son parámetros de formas. Esta reducción en el espacio paramétrico se basa en que al considerar un segundo parámetro en la mezcla de escala ( $\beta$ ) esto podría llevar a un problema de sobrestimación en el estudio de simulación.

Se pueden utilizar dos algoritmos para generar números aleatorios de la distribución HGG como se muestra más abajo, el primer algoritmo utiliza la Proposición 4.1 y el método de composición (ver Tanner, 1996). El segundo algoritmo utiliza la representación dada en la Proposición 4.2.

---

Algorithm 1 for simulating from the  $Z \sim HGG(\beta, \alpha)$  can proceed as follows.

1. Step 1: Generate  $U \sim Beta(\beta, 1)$ .
  2. Step 2: Compute  $Z|U = u \sim G(u^{-1}, \alpha)$
- 

---

Algorithm 2 for simulating from the  $Z \sim HGG(\beta, \alpha)$  can proceed as follows.

1. Step 1: Generate  $V \sim Beta(1 - \alpha, \alpha)$ .
  2. Step 2: Compute  $X = \frac{V}{1-V}$
  3. Step 3: Generate  $Y \sim Beta(\beta, 1)$ .
  4. Step 4: Compute  $Z = \frac{X}{Y}$
-

**Proposición 4.5** Sea  $T \sim HGG(\beta, \alpha)$ . Entonces, la función de riesgo de la variable aleatoria  $T$  es decreciente para todo  $t > 0$ .

**Demostración.** Utilizando el Teorema, ítem (b), dado en Glaser(1980), se tiene que

$$\eta(t) = -\frac{f'(t)}{f(t)} = \frac{\alpha}{t} + \left( \frac{\bar{\alpha} + \beta}{\bar{\alpha} + \beta + 1} \right) {}_2F_1(2, \bar{\alpha} + \beta + 1, \bar{\alpha} + \beta + 2; -t),$$

donde  $f(t)$  es la fdp dada en (4.6), luego derivando la función  $\eta(t)$  con respecto a  $t$

$$\eta'(t) = -\left( \frac{\alpha}{t^2} + 2 \left( \frac{\bar{\alpha} + \beta}{\bar{\alpha} + \beta + 2} \right) {}_2F_1(3, \bar{\alpha} + \beta + 2, \bar{\alpha} + \beta + 3; -t) \right) < 0, \quad \forall t > 0$$

se concluye el resultado  $\square$ .

La siguiente Proposición muestra que no existen los momentos de la distribución HGG, esto es algo común en las distribuciones de cola pesada.

**Proposición 4.6** Los  $r$ -ésimos momentos de la variable aleatoria  $Z \sim HGG(\beta, \alpha)$  no existen para  $r \geq \alpha$ .

**Demostración.** Primero haciendo el siguiente cambio de variable  $t = zw$  y luego

utilizando integración por parte, con  $u = \int_0^z \frac{t^{\beta-\alpha}}{1+t} dt$ , se tiene

$$\begin{aligned}\mathbb{E}(Z^r) &= \frac{\beta}{B(\bar{\alpha}, \alpha)} \int_0^\infty z^{r-\alpha} \int_0^1 \frac{w^{\beta-\alpha}}{1+zw} dw dz \\ &= \frac{\beta}{B(\bar{\alpha}, \alpha)} \int_0^\infty z^{r-\beta-1} \int_0^z \frac{t^{\beta-\alpha}}{1+t} dt dz \\ &= \frac{z^{r-\beta}}{r-\beta} \int_0^z \frac{t^{\beta-\alpha}}{1+t} dt \Big|_0^\infty - \frac{1}{r-\beta} \int_0^\infty \frac{z^{r-\alpha}}{1+z} dz,\end{aligned}$$

La expresión del lado izquierdo vale cero si  $r < \beta$  y no existe si  $r \geq \beta$ , la integral que se encuentra en lado derecho fue analizada en la Proposición 1 (e) del trabajo de Olmos et al. (2022), y ésta no converge. Por lo tanto la integral completa diverge y esto significa que no existen los  $r$ -ésimos momentos de la variable aleatoria  $Z \sim HGG(\beta, \alpha)$ .  $\square$

### 4.2.3. Cola de la distribución

Las distribuciones de cola pesada son aproximadamente aquellas cuyas colas decaen más lentamente que a una tasa exponencial. La distribución exponencial se considera como el límite entre colas pesadas y ligeras. En las áreas del conocimiento que son aplicadas, por ejemplo en finanzas y ciencias actuariales, los datos son positivos y con cola derecha pesada. La distribución de Pareto, ha sido muy utilizada en el modelado de conjuntos de datos financieros, pero en muchos casos no proporciona un buen ajuste, de ahí que resulte de interés considerar nuevas distribuciones con cola

derecha pesada.

Un importante tema en la teoría de valores extremos es la variación regular, ver por ejemplo, Feller (1971), Bingham et al. (1987), Beirlant et al. (2001) y Konstantinides (2018), entre otros. La siguiente proposición establece que la función de supervivencia de la distribución HGG es una distribución con variación regular.

**Proposición 4.7** *La función de supervivencia de la variable aleatoria  $Z \sim HGG(\beta, \alpha)$ , tiene colas que varían regularmente.*

**Demostración.** Aplicando la Definición 2.3 y utilizando la regla de L'Hôpital, tenemos que

$$\lim_{z \rightarrow \infty} \frac{\overline{F}_Z(tz)}{\overline{F}_Z(z)} = \lim_{z \rightarrow \infty} \frac{1 - F(tz)}{1 - F(z)} = t^{-\alpha+1} \lim_{z \rightarrow \infty} \frac{\int_0^1 \frac{w^{\beta-\alpha}}{1+tzw} dw}{\int_0^1 \frac{w^{\beta-\alpha}}{1+zw} dw} = t^{-\beta} \lim_{z \rightarrow \infty} \frac{\int_0^{tz} \frac{u^{\beta-\alpha}}{1+u} du}{\int_0^z \frac{u^{\beta-\alpha}}{1+u} du} = t^{-\beta},$$

y teniendo en cuenta que  $\beta > 0$ , se obtiene el resultado.  $\square$

Como la distribución HGG es de colas de variación regular, entonces es de cola pesada (ver Feller, 1971).

### 4.3. Inferencia

En esta sección realizamos la estimación de los parámetros del modelo HGG por el método de MV, y un estudio de simulación de los estimadores de MV.

### 4.3.1. Estimación de MV

Para una muestra aleatoria  $z_1, \dots, z_n$  derivada de la distribución HGG( $\beta, \alpha$ ), la función log-verosimilitud se puede escribir como

$$\ell(\beta, \alpha) = c(\beta, \alpha) - \alpha \sum_{i=1}^n \log(z_i) + \sum_{i=1}^n \log({}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z_i)), \quad (4.7)$$

donde  $c(\beta, \alpha) = n \log(\beta) - n \log(\bar{\alpha} + \beta) - n \log(\Gamma(\bar{\alpha})) - n \log(\Gamma(\alpha))$ .

Al igualar a cero las primeras derivadas con respecto a cada parámetro, tenemos

$$(\bar{\alpha} + \beta) \sum_{i=1}^n \frac{\Phi(-z_i, 2, \bar{\alpha} + \beta)}{{}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z_i)} = \frac{n}{\beta}, \quad (4.8)$$

$$(\bar{\alpha} + \beta) \sum_{i=1}^n \frac{\Phi(-z_i, 2, \bar{\alpha} + \beta)}{{}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z_i)} - \sum_{i=1}^n \log(z_i) + n\psi(\bar{\alpha}) = n\psi(\alpha) \quad (4.9)$$

Desde (4.8) y (4.9) obtenemos

$$\hat{\beta}(\hat{\alpha}) = \frac{n}{\sum_{i=1}^n \log(z_i) + n\psi(\hat{\alpha}) - n\psi(1 - \hat{\alpha})}, \quad (4.10)$$

y el estimador de MV para  $\alpha$  ( $\hat{\alpha}$ ) se obtiene resolviendo numéricamente la siguiente ecuación

$$(1 - \hat{\alpha} + \hat{\beta}(\hat{\alpha})) \sum_{i=1}^n \frac{\Phi(-z_i, 2, 1 - \hat{\alpha} + \hat{\beta}(\hat{\alpha}))}{{}_2F_1(1, 1 - \hat{\alpha} + \hat{\beta}(\hat{\alpha}), 2 - \hat{\alpha} + \hat{\beta}(\hat{\alpha}); -z_i)} = \frac{n}{\hat{\beta}(\hat{\alpha})}, \quad (4.11)$$

El estimador  $\hat{\alpha}$  es la solución de la ecuación (4.11), y sustituyéndola en (4.10) obtenemos  $\hat{\beta}$ . Solución para esta ecuación (4.11) se puede obtener mediante procedimientos

numéricos como el algoritmo de Newton-Raphson. Alternativamente, estas estimaciones se pueden encontrar maximizando directamente la función log-verosimilitud dada en (4.7) y utilizando la subrutina `optim` en el software R (ver R Core Team, 2021).

### 4.3.2. Matriz de información de Fisher

Consideremos ahora  $Z \sim HGG(\beta, \alpha)$ , para una sola observación  $z$  de  $Z$ . La función log-verosimilitud para  $\boldsymbol{\theta} = (\beta, \alpha)$  viene dada por

$$\ell(\boldsymbol{\theta}) = \log(\beta) - \log(\bar{\alpha} + \beta) - \log(B(\bar{\alpha}, \alpha)) - \alpha \log(z) + \log({}_2F_1(1, \bar{\alpha} + \beta, \bar{\alpha} + \beta + 1; -z)).$$

La primera y segunda derivadas parciales de la función log-verosimilitud se encuentran en el Apéndice C. Se puede demostrar que la matriz de información de Fisher, para la distribución HGG viene dada por

$$I_F(\boldsymbol{\theta}) = \begin{pmatrix} \frac{1}{\beta^2} - 2\eta_{31} + \eta_{22} & 2\eta_{31} - \eta_{22} \\ 2\eta_{31} - \eta_{22} & \psi'(\alpha) + \psi'(\bar{\alpha}) - \eta_{31} + \eta_{22} \end{pmatrix}$$

donde  $\eta_{ij} = E \left( \frac{\kappa_0 \Phi(-Z, i, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -Z)} \right)^j$  se calculan numéricamente.

Por lo tanto, para muestras grandes, los estimadores de MV,  $\widehat{\boldsymbol{\theta}}$ , de  $\boldsymbol{\theta}$  se distribuyen asintóticamente normal bivariada, es decir,

$$\sqrt{n} \left( \widehat{\boldsymbol{\theta}} - \boldsymbol{\theta} \right) \xrightarrow{\mathcal{L}} N_2(\mathbf{0}, I_F(\boldsymbol{\theta})^{-1}),$$

de ahí que la varianza asintótica de los estimadores de MV de  $\hat{\theta}$  sea la inversa de la matriz de información de Fisher.

### 4.3.3. Estudio de simulación

Para examinar el comportamiento del enfoque de la estimación de MV, se lleva a cabo un estudio de simulación para evaluar el rendimiento del método de estimación, procedimiento para los parámetros  $\beta$  y  $\alpha$  de la distribución HGG. Utilizamos el algoritmo 2 dado en la subsección 4.2 para generar números aleatorios de la distribución HGG. El análisis de simulación se realiza generando 1000 muestras de tamaños  $n = 50, 100$  y  $200$  de la distribución HGG. El Cuadro 4.2 muestra el SE, sesgo medio (SM), EE y ECME. Como muestra el Cuadro 4.2, el desempeño de las estimativas mejora cuando  $n$  aumenta.

## 4.4. Aplicación con datos de ingresos por familia

En esta aplicación utilizamos la misma base de datos del Capítulo 2, pero en este caso la variable de interés son los ingresos anuales de cada familia en dólares. Los estadísticos descriptivos de estos datos se muestran en el Cuadro 4.3.

Cuadro 4.2: SE, SM, EE y ECME de la distribución HGG con diferentes combinaciones en los valores de los parámetros.

valores verdaderos		$n = 50$				$n = 100$				$n = 200$				
$\alpha$	$\beta$	estimativa	SE	SM	EE	ECME	SE	SM	EE	ECME	SE	SM	EE	ECME
0.25	0.5	$\hat{\alpha}$	-0.1374	-0.1299	0.0375	0.1879	-0.1482	-0.1402	0.0239	0.1943	-0.1529	-0.0701	0.0159	0.1981
		$\hat{\beta}$	0.4259	0.3745	0.4085	0.5228	0.4217	0.3923	0.2198	0.4898	0.4144	0.4099	0.1280	0.4727
0.25	1.0	$\hat{\alpha}$	-0.3671	-0.3332	0.0438	0.5257	-0.3737	-0.3655	0.0296	0.5300	-0.3770	-0.3645	0.0203	0.5325
		$\hat{\beta}$	0.5271	0.4453	1.1577	0.9360	0.5898	0.5107	0.9012	0.9740	0.6220	0.5731	0.6249	0.9282
0.5	0.5	$\hat{\alpha}$	-0.0188	-0.0182	0.0634	0.0637	-0.0323	-0.0295	0.0414	0.0593	-0.0407	-0.0373	0.0258	0.0588
		$\hat{\beta}$	0.2778	0.1966	0.3919	0.4199	0.2638	0.1822	0.2371	0.3856	0.2398	0.1798	0.1272	0.3406
0.5	1.0	$\hat{\alpha}$	-0.2531	-0.2053	0.0705	0.3605	-0.2550	-0.1412	0.0510	0.3610	-0.2623	-0.2681	0.0366	0.3647
		$\hat{\beta}$	0.5732	0.3697	1.5618	1.0018	0.5860	0.4185	1.1389	0.9673	0.5905	0.4434	0.8198	0.9221
0.8	0.5	$\hat{\alpha}$	0.1428	0.0656	0.0323	0.2089	0.1476	0.1323	0.0225	0.2119	0.1466	0.1436	0.0158	0.2100
		$\hat{\beta}$	0.0499	0.0279	0.2704	0.2794	-0.0465	-0.0156	0.1363	0.1855	-0.0750	-0.0301	0.0861	0.1844
0.8	1.0	$\hat{\alpha}$	-0.1055	-0.1441	0.0325	0.1486	-0.1045	-0.0729	0.0231	0.1455	-0.1037	-0.0537	0.0160	0.1453
		$\hat{\beta}$	0.3400	0.1629	1.0430	0.7558	0.3325	0.1799	0.7387	0.6662	0.2792	0.1685	0.4242	0.5485

Cuadro 4.3: Estadísticas descriptivas de los datos de ingresos

$n$	Mediana	Media	SD	CS	CK
500	54000	321021.9	3410936	21.127	461.575

La Figura 4.4 muestra el diagrama de caja con algunos datos extremos. Estos datos atípicos hacen que la cola derecha sea muy pesada. Cabe destacar que la mayoría de las observaciones se encuentran alrededor de los 5400 dólares como ingresos de cada

familia, y existe un valor muy extremo de ingreso familiar de 75 millones de dólares.

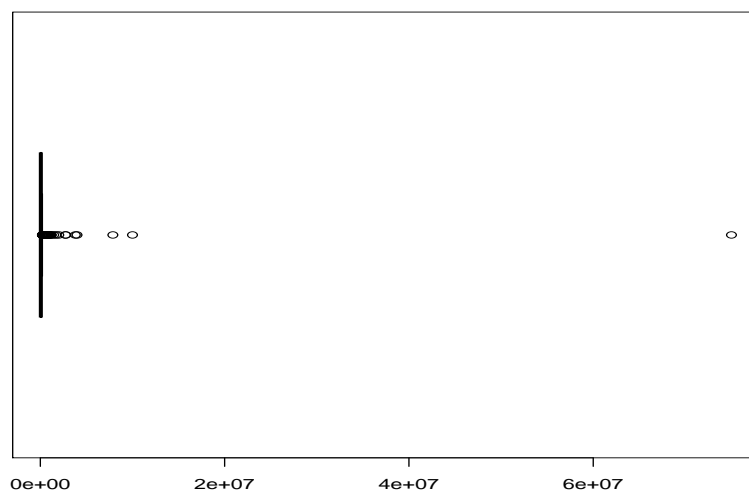


Figura 4.4: Boxplot para los datos de ingresos

Comparamos los ajustes de las distribuciones MEG y Pareto con la distribución HGG de tres parámetros dada en (4.2). El Cuadro 4.4 muestra las estimativas de MV para los parámetros de los modelos de HGG, MEG y Pareto. El Cuadro 4.5 muestra los valores de la función log-verosimilitud y los valores de los criterios AIC y BIC, para cada modelo.

Observamos que los valores más pequeños de los criterios AIC y BIC, corresponden al modelo HGG, lo que significa que el modelo HGG se ajusta mejor a los datos que los otros dos modelos.

Cuadro 4.4: Modelos y estimativas de MV

Modelo	Estimativas de MV
HGG( $\sigma, \alpha, \beta$ )	$\hat{\sigma} = 56417.5$ (2174.647), $\hat{\alpha} = 0.5$ (0.015) , $\hat{\beta} = 756.154$ (62.525)
MEG( $\sigma, \alpha$ )	$\hat{\sigma} = 22584.770$ (4027.346), $\hat{\alpha} = 0.581$ (0.018)
Pareto( $\alpha, \beta$ )	$\hat{\alpha} = 260$ , $\hat{\beta} = 0.186$ (0.008)

Cuadro 4.5: Modelos; valores de la función log-verosimilitud, AIC y BIC.

Modelo	log-verosimilitud	AIC	BIC
GHG( $\sigma, \alpha, \beta$ )	-6466.192	<b>12938.38</b>	<b>12951.03</b>
MEG( $\sigma, \alpha$ )	-6541.776	13087.55	13095.98
Pareto( $\alpha, \beta$ )	-6802.113	13606.23	13610.44

Figura 4.5 muestra la fda empírica con las fda's estimadas de los dos mejores modelos HGG y MEG, lo que también muestra un ajuste mejor del modelo HGG para los datos de ingresos por familia.

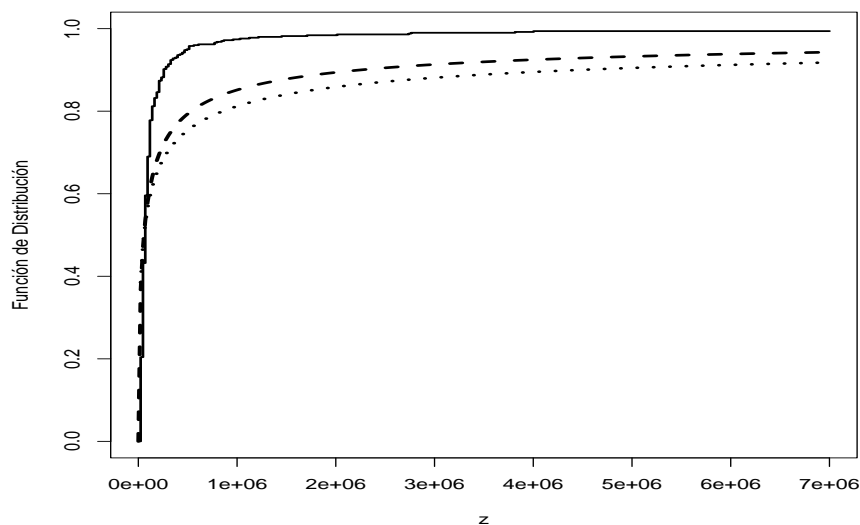


Figura 4.5: Gráficos de la fda empírica (línea continua) con las fda estimadas de los modelos HGG (línea discontinua) y MEG (línea de punteada)

## 4.5. Conclusiones

En este Capítulo se introduce una extensión de la distribución G utilizando la metodología de mezcla de escala, se estudian algunas propiedades, estimación de los parámetros por el método MV y se hace una aplicación a datos reales. La distribución HGG es producto de una mezcla de escala entre las distribuciones G y Beta, obteniendo una nueva distribución con cola derecha pesada y esto la hace interesante, ya que se puede utilizar en estadística actuarial, entre otras áreas. Algunas características

más del modelo HGG son:

- La distribución HGG presenta su cola derecha pesada.
- La distribución HGG tiene dos representaciones dadas en las Proposiciones 4.1 y 4.2.
- Fdp, fda y función de riesgo son explícitas y están representadas por funciones especiales conocidas, por ejemplo la función Gauss hipergeométrica.
- La función de riesgo es decreciente, lo que sugiere información con respecto a la cola de la distribución, es decir, cola derecha pesada (ver Rolski et al., 1999).
- En la aplicación se muestra que la distribución HGG es una buena distribución para el modelado de ingresos por familia, superando a dos distribuciones conocidas como son MEG y Pareto.

## Capítulo 5

### Conclusiones y trabajo futuro

Esta tesis comienza destacando la importancia de los modelos de cola derecha pesada, ya que en diferentes áreas del conocimiento se tienen conjuntos de datos con estas características. Se estudió el modelo G, analizando sus principales propiedades, destacando que este modelo es de cola derecha pesada y que pertenece a las familias subexponenciales y de variación regular. Por otro lado, aunque el modelo G no tiene momentos si tiene función cuantil, lo que permite incorporar el VaR como medida de riesgo. Para ilustrar todos los resultados, en la aplicación se utilizó un conjunto de datos de ingresos per cápita y se comparó su ajuste con el modelo de Pareto obteniéndose buenos resultados. En base al modelo G se obtuvieron los modelos MEG y HGG, ambos con características similares al modelo G. En los tres modelos

se aplicó el método de MV para la estimación de los parámetros, obteniéndose de nuevo buenos resultados. Los estudios de simulación indican que existe una buena recuperación de los valores de los parámetros, aunque en algunos casos se necesita un tamaño de muestra mayor. Se señalan a continuación algunas líneas de trabajo que sin duda pueden constituir un punto de partida para futuras investigaciones.

- Extender los modelos G y MEG, esto se puede lograr incorporado un parámetro que flexibilice la moda de ambos modelos.
- Incorporar una moda más al modelo G, ya que hay muchas situaciones donde hay datos bimodales con cola derecha pesada.
- Realizar regresión en base la mediana del modelo G.
- Se abordará un estudio de estimación más exhaustivo de los tres parámetros de la distribución HGG, donde analizaremos los efectos y las limitaciones de los parámetros  $\alpha$  y  $\beta$  en el método de estimación de MV.
- Se estudiará un extensión bivariada del modelo G.

# Apéndices

## Apéndice A

Las primeras derivadas de  $\ell(\boldsymbol{\theta})$  son dadas por

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\theta})}{\partial \sigma} &= \frac{\alpha}{\sigma} - \frac{1}{\sigma + x} \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \log(\sigma) - \log(x) + \psi(\bar{\alpha}) - \psi(\alpha)\end{aligned}$$

Las segundas derivadas del  $\ell(\boldsymbol{\theta})$  son:

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma^2} = -\frac{\alpha}{\sigma^2} + \frac{1}{(\sigma + x)^2}$$

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma \partial \alpha} = \frac{1}{\sigma}$$

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} = -\psi'(\bar{\alpha}) - \psi'(\alpha)$$

donde  $\psi(\cdot)$  and  $\psi'(\cdot)$  are the digamma and trigamma functions respectively.

## Apéndice B

Las primeras derivadas de  $\ell(\boldsymbol{\theta})$  son dadas por

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\theta})}{\partial \sigma} &= \frac{\alpha}{\sigma} - \frac{z}{\sigma^2(1+z/\sigma)\log(1+z/\sigma)}, \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \frac{1}{\alpha} + \log(\sigma) + \psi(\bar{\alpha}) - \psi(\alpha) - \log(z).\end{aligned}$$

Las segundas derivadas de  $l(\boldsymbol{\theta})$  son:

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma^2} = -\frac{\alpha}{\sigma^2} + \frac{z(2\sigma+z)}{\sigma^2(\sigma+z)^2 \log(1+z/\sigma)} - \frac{z^2}{\sigma^2(\sigma+z)^2 \log^2(1+z/\sigma)},$$

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \sigma \partial \alpha} = \frac{1}{\sigma},$$

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} = -\frac{1}{\alpha^2} - \psi'(\bar{\alpha}) - \psi'(\alpha).$$

## Apéndice C

Las primeras derivadas de  $\ell(\boldsymbol{\theta})$  son dadas por

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta} &= \frac{1}{\beta} - \frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \psi(\bar{\alpha}) - \psi(\alpha) - \log(z) + \frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)}\end{aligned}$$

Las segundas derivadas de  $l(\boldsymbol{\theta})$  son:

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta^2} = -\frac{1}{\beta^2} + \frac{2\kappa_0 \Phi(-z, 3, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} - \left( \frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} \right)^2$$

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta \partial \alpha} = \left( \frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} \right)^2 - \frac{2\kappa_0 \Phi(-z, 3, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)}$$

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} = -\psi'(\bar{\alpha}) - \psi'(\alpha) + \frac{2\kappa_0 \Phi(-z, 3, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} - \left( \frac{\kappa_0 \Phi(-z, 2, \kappa_0)}{{}_2F_1(1, \kappa_0, \kappa_1; -z)} \right)^2$$

donde  $\kappa_i = \bar{\alpha} + \beta + i$ .

# Bibliografía

- [1] Aban, I.B., Meerschaert, M.M., Panorska, A.K. (2006). Parameter estimation for the truncated Pareto distribution. *J. Am. Statist. Assoc.*, 101(473), 270-277.
- [2] Abramowitz, M., Stegun, I.A. (1968). Handbook of Mathematical Functions number 55. In “Applied Mathematics Series”. National Bureau of Standards.
- [3] Afify, A.Z., Gemeay, A.M., Ibrahim, N.A. (2020). The Heavy-Tailed Exponential Distribution: Risk Measures, Estimation, and Application to Actuarial Data. *Mathematics*, **8**, 1276.
- [4] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19(6)**, 716-723.
- [5] Akinsete, A., Famoye, F., Lee, C. (2008). The Beta-Pareto distribution. *Statistics*, **42**, 547–563.

- [6] Andrews, D.F., Mallows, C.L. (1974). Scale mixtures of normal distributions. *J.Roy. Stat Soc. Ser. B*, **36**, 99-102.
- [7] Arnold, B.C. (2015). Pareto Distributions. Monographs on Statistics & Applied Probability (2nd Ed.).
- [8] Artzner, P. (1999). Application of coherent risk measures to capital requirements in insurance. *N. Am. Actuar. J.*, **3**, 11-25.
- [9] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9(3)**, 203-228.
- [10] Astorga, J.M., H.W. Gómez and H. Bolfarine. (2017). Slashed generalized exponential distribution. *Communications in Statistics - Theory and Methods*, **46(5)**, 2091-2102.
- [11] Barlow, R.E., Proschan, F. (1975). Statistical Theory of Reliability and Life Testing: Probability Models, New York: Holt, Rinehart & Winston.
- [12] Beirlant, J., Teugels, J.L., Vynckier, P. (1996). *Practical analysis of extreme values*. Leuven: Leuven University Press.
- [13] Beirlant, J., Matthys, G., Dierckx, G. (2001). Heavy-tailed distributions and rating. *ASTIN Bull.*, **31**, 37-58.

- [14] Beirlant, J., Joossens, E., Segers, J. (2004). Generalized Pareto fit to the society of actuaries' large claims database. *N. Am. Actuar. J.*, **8**, 108-111.
- [15] Bhati, D., Ravi, S. (2018). On generalized log-Moyal distribution: A new heavy tailed size distribution. *Insur. Math. Econ.*, **79**, 247–259.
- [16] Bingham, N. (1987). Regular Variation. Cambridge University Press, Cambridge.
- [17] Boumaraf, B., Seddik-Ameur, N., Barbu, V.S. (2020). Estimation of Beta-Pareto Distribution Based on Several Optimization Methods. *Mathematics*, **8**, 1055.
- [18] Choulakian, V., Stephens, M.A. (2001). Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics*, **43**, 478–484.
- [19] Cooray K., Ananda, M.M.A. (2008). A generalization of the half-normal distribution with applications to lifetime data. *Commun. Stat.-Theory Methods*, **37**, 1323-1337.
- [20] Dunn, P.K., Smyth, G.K. (1996). Randomized Quantile Residuals. *J. Comput. Graph. Stat.*, **5**, 236-244.
- [21] Embrechts, P., Goldie, C.M. (1980). On closure and factorization properties of subexponential and related distributions. *J. Aust. Math. Soc.*, *29*, 243–256.

- [22] Embrechts, P., Goldie, C.M. (1982). On convolution tails. *Stoch. Process. Their Appl.*, *13*, 263–278.
- [23] Embrechts, P., Sidney I.R., Gennady, S. (1999). Extreme value theory as a risk management tool. *N. Am. Actuar. J.* , **3**, 30-41.
- [24] Feller, W. (1971). An introduction to probability theory and its applications. Vol. II. Second edition. John Wiley & Sons, Inc., New York-London-Sydney.
- [25] Foss, S., Korshunov, D., Zachary, S. (2011). An Introduction to Heavy-Tailed and Subexponential Distributions; Springer Series in Operations Research and Financial Engineering; Springer: New York, NY, USA.
- [26] Galton, F. (1883). Enquiries into Human Faculty and its Development. Macmillan & Company, London.
- [27] Glaser, R.E. (1980). Bathtub and Related Failure Rate Characterizations. *Journal of the American Statistical Association*, **75(371)**, 667-672.
- [28] Gleser, L.J. (1989). The Gamma Distribution as a Mixture of Exponential Distributions. *The American Statistician*, **43(2)**, 115-117.

- [29] Gómez, Y.M., Bolfarine, H., Gómez, H.W. (2019). Gumbel distribution with heavy tails and applications to environmental data. *Math. Comput. Simul.*, **157**, 115-129.
- [30] Gómez-Déniz, E., Calderín-Ojeda, E. (2014). A suitable alternative to the Pareto distribution. *Hacet. J. Math. Stat.*, **43**, 843-60.
- [31] Gómez-Déniz, E., Calderín-Ojeda, E. (2015). On the use of the Pareto ArcTan distribution for describing city size in Australia and New Zealand. *Physica A*, **436**, 821-832.
- [32] Gómez-Déniz, E., Calderín-Ojeda, E. (2020). Financial and Actuarial Properties of the Beta-Pareto as a Long-Tail Distribution. *Spanish Journal of Statistics*, **2(1)**, 7-21.
- [33] Gómez, H. W., Olivares-Pacheco, J. F., Bolfarine, H. (2009). An extension of the generalized birnbaum-saunders distribution. *Statistics and Probability Letters*, **79(3)**, 331-338.
- [34] Gómez, H. W., Quintana, F. A. and Torres, F. J. (2007). A New Family of Slash-Distributions with Elliptical Contours. *Statistics and Probability Letters*, **77(7)**, 717-725.

- [35] Gómez, H.W., Venegas, O. (2008). Erratum to: A new family of slash-distributions with elliptical contours [Statist. Probab. Lett. 77 (2007) 717-725]. *Statistics and Probability Letters*, **78(14)**, 2273-2274.
- [36] Gupta, R.D., Kundu, D. (1999). Generalized exponential distributions. *Aust N Z J Stat.*, **41(2)**, 173-188.
- [37] Hogg, R.V., Klugman, S.A. (1984). *Loss Distributions*; John Wiley & Sons: Hoboken, NJ, USA.
- [38] Ibragimov, R., Prokhorov, A. (2017). *Heavy Tails and Copulas: Topics in Dependence Modelling in Economics and Finance*; World Scientific: Singapore.
- [39] Jessen, A.H., Mikosch, T. (2006). Regularly varying functions. *Publ. Inst. Math. Nouvelle Série* , **80**, 171–192.
- [40] Jonhson, N.L., Kotz, S., and Balakrishnan, N. (1995). *Continuous univariate distributions*. Vol 1, 2nd edn. New York: Wiley.
- [41] Kafadar, K. (1982). A biweight approach to the one-sample problem. *J. Amer. Statist. Assoc.*, **77**, 416-424.
- [42] Konstantinides, D. (2018). *Risk Theory. A Heavy Tail Approach*. World Scientific Publishing.

- [43] Mahmoudi, E. (2011). The beta generalized Pareto distribution with application to lifetime data. *Math. Comp. Simul.*, **81**, 2414-2430.
- [44] Moors J.J.A. (1988). A quantile alternative for kurtosis. Journal of the Royal Statistical Society. Series D (The Statistician), **37(1)**, 25-32.
- [45] Mosteller, F., Tukey, J.W. (1977). Data analysis and regression. Addison-Wesley, Reading, MA.
- [46] Mudholkar, G.S., Srivastava, D.K., Freimer, M. (1995). The exponentiated Weibull family: A reanalysis of the Bus-Motor-Failure data. *Technometrics*, **37(4)**, 436-445.
- [47] Nassar, M.M., Nada, N.K. (2011). The beta generalized Pareto distribution. *J. Stat. Adv. Theory Appl.*, **6**, 1–17.
- [48] Olmos, N.M., Gómez-Déniz, E.; Venegas, O. (2022). The Heavy-Tailed Gleser Model: Properties, Estimation, and Applications. *Mathematics*, **10(23)**, 4577.
- [49] Olmos, N.M., Gómez-Déniz, E., Venegas, O. (2024). Scale Mixture of Gleser Distribution with an Application to Insurance Data. *Mathematics*, 12(9), 1397. This article belongs to the Special Issue Probabilistic Models in Insurance and Finance.

- [50] Olmos, N.M., Varela, H., Bolfarine, H., Gómez, H.W. (2014). An extension of the generalized half-normal distribution. *Statistical Papers*, **55**, 967-981.
- [51] Olmos, N.M., Varela, H., Gómez, H.W., Bolfarine, H. (2012). An extension of the half-normal distribution. *Statistical Papers*, **53**, 875-886.
- [52] Pareto, V. (1897). *Cours d'économie politique*, Lausanne, Swiss.
- [53] Proschan, F. (1963). Theoretical Explanation of Observed Decreasing Failure Rate, *Technometrics*, **5**, 375-383.
- [54] R Development Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.
- [55] Resnick, S.I. (1997). Discussion of the Danish data on large fire insurance losses. *ASTIN Bulletin*, **27**, 139-151.
- [56] Rolski, T., Schmidli, H., Schmidt, V., Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*. John Wiley & Sons.
- [57] Rogers, W.H., Tukey, J.W. (1972). Understanding some long-tailed symmetrical distributions. *Statist. Neerlandica*, **26**, 211-226.
- [58] Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Statist.*, **6**(2), 461-464.

- [59] Shannon, C.E. (1948). A Mathematical Theory of Communication. *Bell Syst. Tech. J.*, **27**, 379-423.
- [60] Tanner, M.A. (1996). Tools for statistical inference. In *Methods for the Exploration of Posterior Distributions and Likelihood Functions*, 3rd ed.; Springer-Verlag, New York, USA.
- [61] Wang, J., Genton, M.G. (2006). The multivariate skew-slash distribution. *Journal Statistical Planning and Inference*, **136**, 209-220.
- [62] Zhang, J. (2007). Likelihood moment estimation for the generalized Pareto distribution. *Aust. N. Z. J. Stat.*, **49**, 69–77.