



# COLEGIO DE POSTGRADUADOS

---

---

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN  
EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E  
INFORMÁTICA-ESTADÍSTICA

## Inferencia estadística en la distribución Pareto IV

Marisol López Cerino

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA  
OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO  
2016


---

---


La presente tesis titulada: **Inferencia estadística en la distribución Pareto IV**, realizada por la alumna: **Marisol López Cerino**, bajo la dirección del Consejo Particular indicado ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**DOCTOR EN CIENCIAS  
SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA  
ESTADÍSTICA  
CONSEJO PARTICULAR**

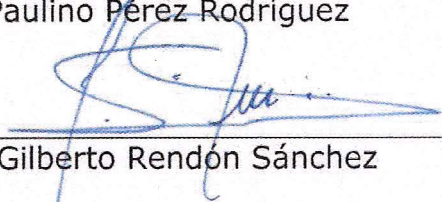
CONSEJERO

  
\_\_\_\_\_  
Dr. Javier Suárez Espinosa

ASESOR:

  
\_\_\_\_\_  
Dr. Paulino Pérez Rodríguez


ASESOR:

  
\_\_\_\_\_  
Dr. Gilberto Rendon Sánchez

ASESOR:

  
\_\_\_\_\_  
Dr. Filemon Ramirez Pérez

ASESOR:

  
\_\_\_\_\_  
Dra. Hortensia Josefina Reyes Cervantes

Montecillo, Texcoco, México, Diciembre 2016.

# Inferencia Estadística en la Distribución Pareto IV

Marisol López Cerino

Colegio de Postgraduados, 2016.

## Resumen

La distribución Pareto Tipo IV es un modelo paramétrico flexible con posibles aplicaciones en confiabilidad, actuaria y ciencias económicas entre otras, por lo que esta distribución será potencialmente útil en muchos casos reales. La distribución tiene cuatro parámetros  $\mu$ ,  $\sigma$ ,  $\gamma$  y  $\alpha$ , y aunque se han utilizado métodos clásicos de estimación, en nuestro conocimiento no existen investigaciones sobre el desempeño de estos métodos para el caso de la distribución Pareto tipo IV con tres parámetros desconocidos. Aunque el caso de los cuatro parámetros desconocidos es un problema abierto, el parámetro de localidad es usualmente conocido o puede ser estimado por las estadísticas mínimas de orden de muestreo. Por lo anterior, se realizó un estudio de simulación para evaluar desempeño del método de momentos (MOM), máxima verosimilitud (MMV) y método de distancia mínima (MDM) para estimar  $\sigma$ ,  $\gamma$  y  $\alpha$ . Además, se propone un algoritmo híbrido alternativo para obtener estimaciones por el MDM. El desempeño de los MOM y MMV fue insatisfactorio en términos de tasas de convergencia y error cuadrado medio (ECM). El estimador obtenido por el MDM tiene mejor desempeño en término de tasas de convergencia pero en término de ECM su desempeño es deficiente. El algoritmo híbrido propuesto para obtener el estimador por el MDM funciona mejor que el MOM, MMV y MDM ordinario. Por lo anterior, la propuesta es una buena alternativa cuando los métodos clásicos no funcionan.

**Palabras claves:** Colas pesadas, inferencia paramétrica, simulación Monte Carlo.

# Statistical inference in the Pareto IV distribution

Marisol López Cerino

Colegio de Postgraduados, 2016.

## Abstract

The type IV Pareto distribution is a flexible parametric model with possible applications in reliability, actuarial and economic sciences among others, so this distribution will be potentially useful in many real cases. This distribution has four parameters  $\mu$ ,  $\sigma$ ,  $\gamma$  and  $\alpha$ , and although the classical estimation methods have been used for estimating some of these parameters, in our best knowledge, there is no research about the performance of these methods for the type IV Pareto distribution with three unknown parameters. Though the four unknown parameters case is an open problem, the location parameter is usually known or can be estimated by the minimum sampling order statistic. So, we did a simulation study to evaluate the performance of the method of moments (MOM), maximum likelihood (MLE) and minimum distance method (MDM) for estimating  $\sigma$ ,  $\gamma$  and  $\alpha$ . In addition, we propose an alternative hybrid algorithm to obtain estimates by the MDM. The performance of the MOM and MLE was unsatisfactory in terms of convergency rate and mean squared error (MSE). The estimator obtained by the ordinary MDM have a better performance in terms of convergency rate but in terms of MSE its performance is poor. The proposed hybrid algorithm to obtain the estimator by the MDM works better than the MLE, MOM and ordinary MDM. The proposal has a good performance in term of convergence and MSE; therefore, the proposal is a good alternative when the ordinary methods do not work.

**Key words:** heavy tails, parametric inference, Monte Carlo simulation.

## AGRADECIMIENTOS

A Dios, por el soplo de vida que hay en mí.

A mi Madre Elvia, pilar fundamental en mi vida.

A mis hermanos Javier, Rafael e Ignacio, ellos son los hermanos que cualquier persona en el mundo, estaría feliz de tener a su lado.

A mi amado esposo José del Carmen, quien ha caminado a mi lado durante todo este proceso, con sus consejos y ejemplo me ha mostrado que los sueños pueden hacerse realidad.

Quiero expresar mi infinito agradecimiento a mi princesa Julieta, por que ella tuvo que soportar las largas horas sin la compañía de su mamá, sin entender, a su corta edad, el porqué mamá Marisol prefería estar al frente de la computadora y no jugar con ella, llevarla de paseo. Y en los momentos difíciles su sonrisa me llenaba de ánimo y fuerza para continuar.

A mi suegra Doña Concha, por su gran apoyo y cuidados.

A mi amiga (hermana) Patricia Broca, por sus oraciones y quien me ha enseñado a no perder la esperanza en los seres humanos.

A mi consejero Dr. Javier Suarez Espinosa, por su paciencia y tiempo. Así como su platicas, anécdotas que compartió conmigo.

Al mi consejo particular integrado por: Dr. Javier Suárez Espinosa, Dr. Paulino Pérez Rodríguez, Dr. Gilberto Rendón Sánchez, Dr. Filemón Ramírez Pérez, Dra. Hortensia Josefina Reyes Cervantes. Por sus observaciones y sugerencia que permitieron, una mejor desarrollo de este trabajo.

Agradezco al CONACyT por el apoyo brindado en el tiempo que duraron mis estudios de doctorado.

## DEDICATORIA

*Este trabajo esta dedicado,  
con mucho amor a Julieta y  
mi amado esposo José del Carmen.*

# Contenido

<b>1. Introducción</b>	<b>1</b>
<b>2. Marco Teórico</b>	<b>4</b>
2.1. Distribución Pareto IV . . . . .	4
2.1.1. Función de densidad y propiedades . . . . .	4
2.1.2. La distribución Pareto IV y su relación con otras distribuciones	6
2.1.3. Momentos de la distribución Pareto IV . . . . .	6
2.2. Estimación . . . . .	7
2.2.1. Conceptos y definiciones . . . . .	7
2.2.2. Métodos de estimación . . . . .	9
2.2.3. Propiedades de los estimadores . . . . .	17
<b>3. Métodos</b>	<b>21</b>
3.1. Introducción . . . . .	21
3.2. Método de momentos . . . . .	22
3.3. Método de momentos fraccionarios . . . . .	24
3.4. Método de máxima verosimilitud . . . . .	25
3.5. Método de distancia mínima . . . . .	26
3.6. Método de distancia mínima obtenido por un algoritmo híbrido (Cuasi-Newton-Malla) . . . . .	27
3.7. Región estudiada del espacio paramétrico . . . . .	31
<b>4. Resultados y Discusión</b>	<b>37</b>
4.1. Método de momentos . . . . .	37
4.2. Método de máxima verosimilitud . . . . .	37
4.3. Método de distancia mínima . . . . .	39
4.3.1. Método de distancia mínima obtenido por un algoritmo híbrido (Cuasi-Newton-Malla) . . . . .	40

4.3.2. Ejemplos de Aplicación . . . . .	42
<b>5. Conclusiones</b>	<b>45</b>
Anexo A: Tablas . . . . .	48
Anexo B: Programas para la estimación de los parámetro en paquete R . . .	56



# Lista de Tablas

4.1. Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ obtenidos con el método de máxima verosimilitud, con $N = 1000$ réplicas Monte Carlo; $n=50, 100, 200$ . . . . .	39
4.2. Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos con el método de distancia mínima, con $N=5000$ réplicas Monte Carlo; $n= 50, 100, 200$ . . . . .	40
4.3. Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , método de distancia mínima obtenido con un algoritmo híbrido (Cuasi-Newton-Malla), con $N=5000$ réplicas Monte Carlo; $n= 50, 100, 200$ . . . . .	41
4.4. Datos del golfistas, ganancia en miles de dólares. Fuente: Golf magazine 1981 yearbook . . . . .	42
5.1. Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de máxima verosimilitud (MMV), con $N=1000$ réplicas Monte Carlo; $n=50$ . . . . .	49
5.2. Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de máxima verosimilitud (MMV), con $N=1000$ réplicas Monte Carlo; $n=100$ . . . . .	50
5.3. Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de máxima verosimilitud (MMV), con $N=1000$ réplicas Monte Carlo; $n=200$ . . . . .	51
5.4. Estimaciones puntuales (ECM) de los pámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de distancia mínima, con $N=5000$ réplicas Monte Carlo; $n=50$ . . . . .	52
5.5. Estimaciones puntuales (ECM) de los pámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de distancia mínima, con $N=5000$ réplicas Monte Carlo; $n=100$ . . . . .	52

5.6.	Estimaciones puntuales (ECM) de los pámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de distancia mínima, con $N=5000$ réplicas Monte Carlo, $n=200$ . . . . .	53
5.7.	Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de distancia mínima obtenido con un algoritmo híbrido (Cuasi–Newton–Malla), con $N=5000$ réplicas Monte Carlo; $n=50$ . . . . .	54
5.8.	Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de distancia mínima obtenido con un algoritmo híbrido (Cuasi–Newton–Malla), con $N=5000$ réplicas Monte Carlo; $n=100$ . . . . .	55
5.9.	Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto $IV(\mu_0, \sigma, \gamma, \alpha)$ , obtenidos por el método de distancia mínima obtenido con un algoritmo híbrido (Cuasi–Newton–Malla), con $N=5000$ réplicas Monte Carlo; $n=200$ . . . . .	56

# Lista de Figuras

2.1. a) Función de densidad de la familia Pareto $IV(\mu, \sigma, \gamma, \alpha)$ , b) Función de distribución de la familia Pareto $IV(\mu, \sigma, \gamma, \alpha)$ . . . . .	5
3.1. Descripción gráfica de la búsqueda del estimador de mínima distancia usando malla. . . . .	28
3.2. En a) y c) Función de densidad para valores de $\gamma < 1$ , fijando $\sigma = 1$ y $\alpha = 1$ ; b) y d) Función de densidad para valores de $\gamma \geq 1$ , fijando $\sigma = 0.5$ y $\alpha = 5$ . . . . .	32
3.3. Función de densidad para valores de $\sigma = 2$ y $\gamma = 0.2$ fijos y variaciones en $\alpha \in \{0.5, \dots, 10\}$ . . . . .	33
3.4. Función de densidad para valores de $\sigma = 2$ y $\gamma = 0.8$ fijos y variaciones en $\alpha \in \{0.5, \dots, 10\}$ . . . . .	34
3.5. Función de densidad para valores de $\gamma = 0.8$ y $\alpha = 1$ fijos y variaciones en $\sigma \in \{1, \dots, 40\}$ . . . . .	35
3.6. Función de densidad para valores de $\gamma = 1$ y $\alpha = 2$ fijos y variaciones en $\sigma \in \{0.5, \dots, 50\}$ . . . . .	36

# Índice de algoritmos

1.	Algoritmo para la evaluación del desempeño del estimador de momentos.	23
2.	Algoritmo para la evaluación del desempeño del estimador de momentos fraccionarios. . . . .	24
3.	Algoritmo para la evaluación del desempeño del estimador de máxima verosimilitud. . . . .	26
4.	Algoritmo para la evaluación del desempeño del estimador de distancia mínima. . . . .	27
5.	Algoritmo para la evaluación del desempeño del estimador de distancia mínima obtenido mediante una técnica de aleatoriedad pura en combinación con un método Cuasi-Newton. . . . .	30

# Capítulo 1

## Introducción

El origen de la hoy llamada distribución Pareto tipo IV (Pareto IV) se remota a los años cuarentas del siglo pasado cuando [Burr \(1942\)](#) buscando obtener los percentiles de una variable aleatoria (v.a.) propuso una distribución llamada Burr XII. La propuesta de Burr fue modificada más tarde por [Johnson \*et al.\* \(1994\)](#) agregándole los parámetros de localidad y escala para obtener la distribución Pareto IV. Como dato interesante la función de distribución Pareto IV fue también obtenida como resultado de una mezcla de distribuciones Weibull por [Dubey \(1968\)](#) y [Harris y Singpurwalla \(1969\)](#). Por otra parte [Arnold \(1985\)](#) estableció la jerarquía de los modelos Pareto iniciando con la distribución Pareto clásica (Pareto I) e introduciendo los parámetros de localidad, escala, forma y desigualdad dando como resultado la familia de distribuciones Pareto IV.

**Definición 1.1** *La función de distribución acumulada,  $F(x)$ , para una variable aleatoria  $X$  con distribución Pareto IV( $\mu, \sigma, \gamma, \alpha$ ) es:*

$$F_X(x; \underline{\theta}) = 1 - \left[ 1 + \left( \frac{x - \mu}{\sigma} \right)^{1/\gamma} \right]^{-\alpha}, \quad x > \mu, \quad (1.1)$$

donde  $\underline{\theta} = (\mu, \sigma, \gamma, \alpha)$ ,  $-\infty < \mu < +\infty$ ,  $\sigma > 0$ ,  $\gamma > 0$  y  $\alpha > 0$  son parámetros de localidad, escala, desigualdad y forma, respectivamente.

Si una variable aleatoria  $X$  tiene la función de distribución anterior,  $F(x)$ , entonces la denotaremos como usualmente se hace:  $X \sim \text{Pareto IV}(\mu, \sigma, \gamma, \alpha)$ .

Naturalmente la jerarquización de Arnold se puede ver fácilmente considerando la reparametrización de  $\mu, \sigma, \gamma, \alpha$ , como sigue:

1. Si  $\gamma = 1$  y  $\mu = \sigma$  se obtiene la distribución Pareto I (Pareto I( $\sigma, \alpha$ )).

2. Si  $\gamma = 1$  se obtiene la distribución Pareto II (Pareto II( $\mu, \sigma, \alpha$ )).
3. Si  $\alpha = 1$  se obtiene la distribución Pareto III (Pareto III( $\mu, \sigma, \gamma$ )).
4. Si  $\mu = 0$  entonces se obtiene la distribución Burr XII (Burr XII( $0, \sigma, \alpha, 1/\gamma$ )).

La distribución Pareto IV tiene la característica de ser un modelo paramétrico flexible como se verá en la siguiente sección, con aplicaciones en confiabilidad, ciencias actuariales, ciencias económicas y finanzas entre otras y por lo anterior su importancia en la modelación de este tipo eventos es relevante. En los últimos años esta distribución ha sido usada para modelar diversos eventos. [Singh y Maddala \(1978\)](#) modelaron tasas de fallas decrecientes y [Crosato y Ganugi \(2007\)](#) modelaron el total de activos de las empresas manufactureras Italianas. También ha sido usada para modelar las pérdidas en carteras de seguros. En [Burr \(1942\)](#) y [Robertson \(1977\)](#) modelaron el tamaño de los incendios forestales en California y [Arnold \(2015\)](#) modeló el ingreso personal de Texanos y los ingresos personales de golfistas. Debido a su gran flexibilidad este modelo puede ser usado en otras áreas; por ejemplo, [Owens et al. \(1978\)](#) han señalado la posibilidad de usar este modelo en experimentos cuánticos.

El modelado de eventos con la distribución Pareto IV usualmente requiere de la estimación de sus parámetros; sin embargo, poca investigación ha sido realizada al respecto, especialmente para el caso de tres parámetros desconocidos. Por ejemplo, [Harris y Singpurwalla \(1969\)](#) realizaron estimaciones por el método de máxima verosimilitud (MMV) y mencionaron que hay que tener cuidado en las situaciones donde la log-verosimilitud es plana ya que la distribución Pareto IV puede presentar dicho problema. [Al-Marzoug y Ahmad \(1985\)](#) realizaron la estimación por momentos fraccionarios para la distribución Pareto IV cuando  $\mu = 0$  y  $\sigma = 1$ , es decir, sólo estiman dos parámetros y comparan sus resultados con la estimación encontrada por máxima verosimilitud y la estimación por momentos ordinarios en términos del sesgo. Por otra parte, aunque [Jalali y Watkins \(2009\)](#) proporcionan una discusión cuidadosa de la estimación por máxima verosimilitud, dicha discusión se limita solamente al mismo caso abordado por Marzoug; es decir, cuando  $\mu = 0$  y  $\sigma = 1$ . Debido a las dificultades encontradas para la estimación de los parámetros de la distribución Pareto IV usando los métodos clásicos [Al-Hussaini y Jaheen \(1992\)](#) proponen una estimación alternativa por el método Bayesiano considerando que  $\mu$  y  $\sigma$  son conocidos; es decir, nuevamente se aborda el problema para solo dos parámetros desconocidos. Por otra parte [Crosato y Ganugi \(2007\)](#) obtienen los parámetros por el MMV para realizar pruebas de bondad de ajuste, [Munir et al. \(2013\)](#) comparan diferentes métodos de estimación de parámetros pero sólo para el caso Pareto tipo I. Recientemente, [Ghosh](#)

(2011) estudió a la distribución Pareto IV bivariada truncada con tres parámetros desconocidos y describe que sus distribuciones marginales pertenecen a la familia Pareto IV. Ghosh obtiene estimaciones por varios métodos (MMV, MOM, método de cuantiles y estimación Bayesiana); aunque se presentan estimaciones para el modelado de datos reales, el estudio por simulación del desempeño de estos métodos es muy limitado ya que solo considera una tripleta de todos valores posibles del espacio paramétrico  $(\mu, \sigma, \alpha)$ . Sin duda la estimación de los cuatro parámetros de la distribución Pareto IV es un problema interesante; sin embargo, Arnold (2015) indica que en la práctica el parámetro de localización con frecuencia es conocido o puede ser estimado con el mínimo de las estadísticas de orden por lo que, mientras tanto, un problema suficientemente interesante es la estimación de los parámetros  $\sigma$ ,  $\gamma$  y  $\alpha$ .

Por lo anterior el **objetivo general** de este trabajo es la realización de un estudio de simulación para medir el desempeño de los estimadores clásicos usados hasta ahora para el caso Pareto IV con parámetros desconocidos;  $\sigma$ ,  $\gamma$  y  $\alpha$ ; este estudio se encuentra dentro del soporte del espacio paramétrico definido. Otro interés en este trabajo es implementar un método de estimación alternativo en caso de que los métodos clásicos no tengan un buen desempeño.

---

# Capítulo 2

## Marco Teórico

### 2.1. Distribución Pareto IV

En este capítulo se describe la distribución Pareto IV con más profundidad, así como algunas de sus propiedades.

#### 2.1.1. Función de densidad y propiedades

La función de densidad de una v.a. se utiliza en la práctica para modelar la distribución de probabilidades de un suceso o evento, por lo que es importante conocer su forma completa; así para el caso de la distribución Pareto IV se tiene:

**Definición 2.1** *La función de densidad de la Pareto IV( $\mu, \sigma, \gamma, \alpha$ ) está dada por,*

$$f_X(x; \theta) = \frac{\alpha}{\sigma\gamma} \left( \frac{x - \mu}{\sigma} \right)^{1/\gamma - 1} \times \left( 1 + \left( \frac{x - \mu}{\sigma} \right)^{1/\gamma} \right)^{-(\alpha + 1)} \quad (2.1)$$

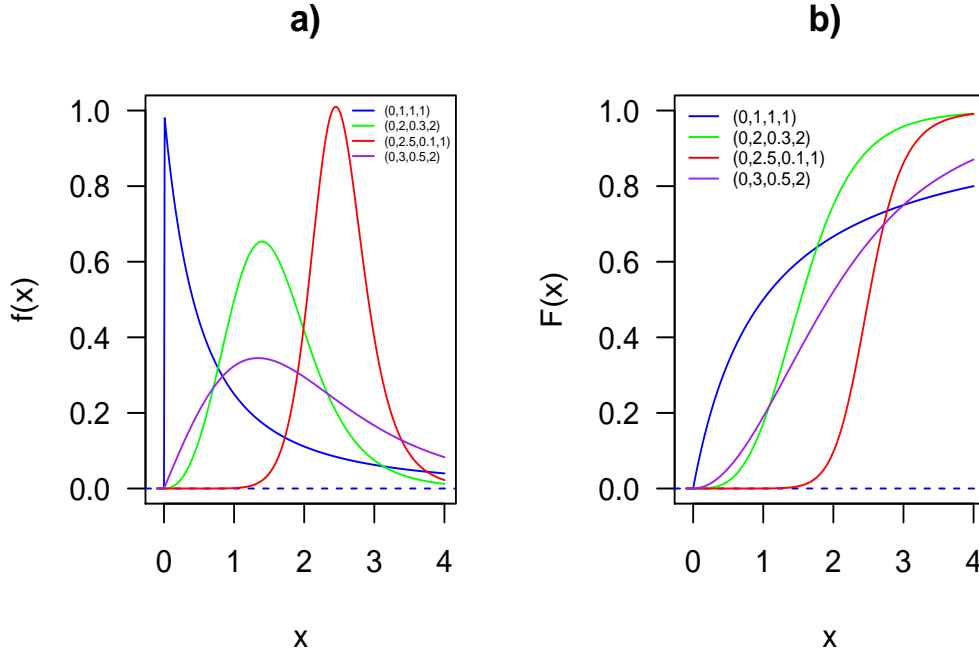
donde  $-\infty < \mu < +\infty$ ,  $\sigma > 0$ ,  $\gamma > 0$  y  $\alpha > 0$  son parámetros de localización, escala, desigualdad y forma, respectivamente.

Como se puede observar la función de densidad del modelo Pareto IV está caracterizada por cuatro parámetros. Si  $\alpha = 1$ ,  $\mu = 0$  y el parámetro de desigualdad  $\gamma < 1$  entonces  $\gamma$  resulta ser el índice de desigualdad de Gini, por lo que este modelo es frecuentemente usado para la modelación de ingresos ver [Arnold \(2015\)](#).

Como el conjunto de valores posibles de los parámetros define a la familia Pareto IV de distribuciones, esta familia resulta ser muy flexible; es decir, es posible usarla para modelar eventos de muy diversa índole y forma. Para ilustrar dicha flexibilidad



en la Figura 2.1 a) se muestra la función de densidad de la distribución Pareto IV para los siguientes parámetros, donde  $\underline{\theta} = (\mu, \sigma, \gamma, \alpha)$ , con  $(\mu = 0, \sigma = 1, 2, 2.5, 3, \gamma = 0.1, 0.3, 0.5, 1, \alpha = 1, 2)$ . La Figura 2.1 b) muestra la función de distribución para los mismos parámetros considerados en el inciso a).



**Figura 2.1:** a) Función de densidad de la familia Pareto IV( $\mu, \sigma, \gamma, \alpha$ ), b) Función de distribución de la familia Pareto IV( $\mu, \sigma, \gamma, \alpha$ ).

Pocas distribuciones siguen una ley de potencia sobre el rango entero. Normalmente, es a partir de un valor donde se comienza a observar la ley de potencia. Se dice que tales variables tienen ley de potencia en la cola, (*power-law tail*). Por lo tanto es común definir una distribución de *cola pesada*, como aquellas cuyas colas decaen siguiendo una ley de potencia.

$$\overline{F(x)} = P(X > x) \propto x^{-a}, \quad 0 < a < 2. \quad (2.2)$$

Al exponente  $a$  se le llama índice de la cola (Newman, 2005). En particular, la función de densidad de la distribución Pareto IV es capaz de modelar la cola pesada ya que cumple con la ley de potencia.

### 2.1.2. La distribución Pareto IV y su relación con otras distribuciones

Si bien se ha mencionado que la jerarquía de las distribuciones Pareto se han establecido a partir de la distribución Pareto clásica (Pareto I), y que la posterior adición de parámetros relacionados con la ubicación, la escala, la forma y la desigualdad dan como resultado un modelo más general llamado Pareto IV, también se ha encontrado que la transformación de algunas variables aleatorias tienen como resultado la distribución Pareto IV. Este tipo de relaciones hace posible obtener más fácilmente sus momentos naturales entre otras propiedades.

Sea  $Y$  una v.a. que sigue una distribución Beta con parámetros  $\gamma_1$  y  $\gamma_2$ , cuya función de densidad es

$$f_Y(y) = y^{\gamma_1-1}(1-y)^{\gamma_2-1}/B(\gamma_1, \gamma_2), \quad 0 < y < 1, \quad (2.3)$$

defina  $W = \mu + \sigma(Y^{-1} - 1)^\gamma$ , donde  $Y \sim \text{Beta}(\gamma_1, \gamma_2)$ ,  $\mu$  es real,  $\sigma > 0$  y  $\gamma > 0$ , entonces esta transformación, sugerida por [Feller \(1971\)](#), sigue una distribución Feller-Pareto y cuando  $\gamma_2 = 1$  y  $\gamma_1 = \alpha$ , se tiene la distribución Pareto IV.

Cabe hacer notar que el resultado anterior indica que la familia de distribuciones Pareto IV, objeto de nuestro estudio, es una subfamilia de la distribución Feller-Pareto y por tanto los resultados obtenidos para esta última distribución pueden ser usados para la Pareto IV; por ejemplo la obtención de sus momentos naturales y la matriz de información de Fisher.

Otro resultado, aunque no tan trascendente como el anterior, es que la distribución Pareto IV tiene una representación como una mezcla de distribuciones Weibull. Así, si

$$P(X > x | Z = z) = e^{z[(x-\mu)/\sigma]^{1/\gamma}}, \quad x > \mu, \quad (2.4)$$

donde  $Z \sim \Gamma(\alpha, 1)$ , entonces  $X \sim \text{Pareto IV}(\mu, \sigma, \gamma, \alpha)$ .

Otra relación interesante con otra distribución es la siguiente: si  $X \sim \text{Pareto IV}(0, 1/\sigma, \gamma, \alpha)$ , entonces  $1/X$  sigue una distribución Dagum(I)( $\sigma, \gamma, \alpha$ ) ([Sarabia, 2008](#)), lo anterior puede ser útil al realizar la estimación de parámetros.

### 2.1.3. Momentos de la distribución Pareto IV

Uno de los métodos de estimación clásicos es el método de momentos (MOM), para lo cual se requiere obtener los momentos naturales de la v.a. en cuestión. Luego entonces para el caso de nuestro estudio se tiene que el  $k$ -ésimo momento de la v.a.

$X \sim \text{Pareto IV}(0, \sigma, \gamma, \alpha)$  se cálcula de la siguiente forma:

$$E[X^k] = \frac{\alpha}{\sigma\gamma} \int_0^\infty x^k \left(\frac{x}{\sigma}\right)^{1/\gamma-1} \left[1 + \left(\frac{x}{\sigma}\right)^{1/\gamma}\right]^{-\alpha-1} dx. \quad (2.5)$$

Naturalmente la solución de esta integral se puede obtener haciendo un cambio de variable; sin embargo, también se puede obtener usando la transformación sugerida por Feller citado por (Arnold, 2015) obteniendo el siguiente resultado :

$$\begin{aligned} E[X^k] &= \int_0^1 (y^{-1} - 1)^k \frac{y^{\alpha-1}}{B(\alpha-1)} dy \\ &= \frac{\sigma^k \Gamma(k\gamma + 1) \Gamma(\alpha - k\gamma)}{\Gamma(\alpha)}. \end{aligned} \quad (2.6)$$

Si consideramos  $k = 1$ , se obtiene su valor esperado:

$$E[X] = \frac{\sigma \Gamma(\gamma + 1) \Gamma(\alpha - \gamma)}{\Gamma(\alpha)}, \quad (2.7)$$

y su varianza respectiva es:

$$Var[X] = \frac{\sigma^2 \Gamma(\alpha - 2\gamma) \Gamma(1 + 2\gamma)}{\Gamma(\alpha)} - \frac{\sigma^2 \Gamma^2(\alpha - \gamma) \Gamma^2(1 + \gamma)}{\Gamma^2(\alpha)}. \quad (2.8)$$

## 2.2. Estimación

En esta sección se revisarán conceptos, definiciones y propiedades deseables de un estimador. Además, se presentarán los métodos de estimación que serán estudiados en este trabajo de investigación para el caso de la distribución Pareto IV.

### 2.2.1. Conceptos y definiciones

La teoría de estimación se ocupa, dentro del marco de la perspectiva clásica, de estudiar las características deseables de los estimadores, permitiendo con este conocimiento la evaluación de diferentes tipos de estimadores. Para lo anterior, se definirán los conceptos más importantes involucrados en este proceso.

La teoría estadística tiene su fundamento en lo que se conoce como un *Espacio de Probabilidad*. Dicho espacio de probabilidad tiene tres componentes  $(\Omega, \mathcal{A}, P)$ , donde  $\Omega$  es un espacio muestral,  $\mathcal{A}$  representa a una  $\sigma$ -álgebra de subconjuntos de  $\Omega$  y  $P$  es una medida de probabilidad en  $(\Omega, \mathcal{A})$ . El espacio muestral denotado por  $\Omega$ ,

es el conjunto de valores posibles resultados de un experimento aleatorio (evento aleatorio). Se dice que cualquier experimento (evento) que involucre incertidumbre puede ser modelado por un *Espacio de Probabilidad*.

**Definición 2.2** (*Variable aleatoria*) Sea  $(\Omega, \mathcal{A}, P)$  un espacio de probabilidad. Una función real cuyo dominio es  $\Omega$  y contradominio los reales  $\mathfrak{R}$  es llamada una variable aleatoria (v.a.) si la imagen inversa bajo  $X$  de todo el conjunto de Borel en  $\mathfrak{R}$  son subconjuntos de  $\Omega$  (eventos), es decir, si

$$X^{-1}(B) = \{w : X(w) \in B\} \in \mathcal{A} \text{ para todo } B \in \mathcal{B}. \quad (2.9)$$

El comportamiento probabilístico de una v.a. está determinado por su función de densidad o su función de distribución. Frecuentemente la función de densidad o distribución depende de un vector de parámetros, comúnmente denotado por  $\underline{\theta}$ . En el enfoque clásico, los elementos de dicho vector de parámetros se consideran cantidades fijas pero desconocidas. Los parámetros de una función de densidad o probabilidad generan una familia de funciones de densidad, lo cual permite modelar diversas situaciones de interés.

El uso de las funciones de densidad o funciones de distribución para modelar eventos de interés requiere del conocimiento de su forma y el valor de sus parámetros. La teoría de estimación que nos ocupa asume que la forma de la función de densidad para modelar el evento de interés es conocida y que solo se requiere investigar cual es el valor de o los parámetros o dicho de otra forma se asume que la familia de densidades para modelar el evento de interés es la correcta y solo falta investigar que función de densidad específica es la adecuada para el conjunto de datos observados.

**Definición 2.3** Sea  $X$  una variable aleatoria con función de distribución  $F_X(x; \underline{\theta})$ , y sean  $X_1, X_2, \dots, X_n$  variables aleatorias independientes idénticamente distribuidas con función de distribución  $F(x; \underline{\theta})$ . Entonces a la colección  $X_1, X_2, \dots, X_n$  se le conoce como una muestra aleatoria (m.a.) de tamaño  $n$  con función de distribución  $F(x; \underline{\theta})$ .

**Definición 2.4** Sean  $X_1, X_2, \dots, X_n$ , una m.a. y sea  $T : \mathcal{R}_n \rightarrow \mathcal{R}_k$  una función Borel medible. Entonces a la variable aleatoria  $T(X_1, X_2, \dots, X_n)$  se le llama una estadística.

**Definición 2.5** Sean  $X_1, X_2, \dots, X_n$ , una m.a. con función de distribución  $F_X(x; \underline{\theta})$  tal que  $\underline{\theta} \in \Theta$ . Sea  $T(X) = T(X_1, X_2, \dots, X_n)$  una estadística, y sea  $\chi$  el conjunto

---

de valores posibles de  $T(X)$ . Luego entonces  $T(X)$  es un estimador de  $\underline{\theta}$  si  $\chi \subseteq \Theta$ , donde  $\Theta$  es el espacio paramétrico de  $\underline{\theta}$ .

**Definición 2.6** Una realización de un estimador  $T(X) = t(X)$ , es conocida como una estimación.

### 2.2.2. Métodos de estimación

En esta sección abordaremos algunos métodos para obtener estimadores. Suponga que se desea fijar las primas de seguros de automóviles de una ciudad determinada y estudios previos han mostrado que la función de densidad Pareto IV es un buen modelo para este tipo de eventos. Entonces para fijar las primas de seguros se requiere conocer las probabilidades de los eventos; el cálculo de probabilidades en este caso requiere del conocimiento de los parámetros de la distribución Pareto IV. Si los parámetros de la distribución no son conocidos, como usualmente lo es, entonces se requiere estimarlos. Para lo cual existen varios métodos de estimación que se revisarán en esta sección.

#### Método de momentos (MOM)

El método de momentos es uno de los procedimientos más antiguos usados para la estimación de parámetros. El MOM fué propuesto por Karl Pearson a finales del siglo XIX y tiene su fundamento en la ley de los Grandes Números. Para la descripción de este método primeramente definamos a los momentos naturales muestrales.

Sea  $X_1, \dots, X_n$  una muestra aleatoria de tamaño  $n$  cuya función de distribución es  $F_X(x; \underline{\theta})$ , donde  $\underline{\theta}$  es un vector de parámetros con elementos  $(\theta_1, \theta_2, \dots, \theta_k)^t$ . El  $i$ -ésimo momento natural muestral se define como:

$$M_i(X) = \frac{1}{n} \sum_{j=1}^n X_j^i. \quad (2.10)$$

Por otra parte, el  $i$ -ésimo momento natural de una v.a.  $X$  se definen como:

$$E(X^i) = \mu_i(\underline{\theta}). \quad (2.11)$$

La ley de los Grandes Números nos dice que el  $i$ -ésimo momento natural muestral converge casi seguro (c. s.) al  $i$ -ésimo momento natural de una v.a. siempre que  $n$

---

sea suficientemente grande; es decir,

$$M_i(X) = \frac{1}{n} \sum_{j=1}^n X_j^i \xrightarrow{c.s.} E(X^i). \quad (2.12)$$

Por lo anterior resulta razonable definir al estimador de momentos como sigue: Sea  $X_1, \dots, X_n$  una muestra aleatoria de tamaño  $n$  cuya función de distribución es  $F_X(x; \underline{\theta})$ , donde  $\underline{\theta}$  es un vector de parámetros con elementos  $(\theta_1, \theta_2, \dots, \theta_k)^t$ . El estimador de momentos obtenido en base a una m.a.  $X_1, \dots, X_n$  es aquel que se obtiene resolviendo el siguiente sistema de  $k$  ecuaciones y  $k$  incógnitas:

$$E(X^i) = M_i(X), \quad \text{para } i = 1, 2, \dots, k; \quad (2.13)$$

es decir,

$$\begin{aligned} M_1(X) &= \mu_1(\theta_1, \theta_2, \dots, \theta_k) \\ M_2(X) &= \mu_2(\theta_1, \theta_2, \dots, \theta_k) \\ &\vdots \\ M_k(X) &= \mu_k(\theta_1, \theta_2, \dots, \theta_k). \end{aligned}$$

Si existe una solución para este sistema de ecuaciones, entonces se dice que  $\tilde{\underline{\theta}} = (\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_k)^t$  es el estimador de momentos de  $\underline{\theta}$ . En muchos casos se tienen métodos numéricos que son computacionalmente tratables y es principalmente por este motivo que se emplea con éxito.

Los estimadores de momentos son muy versátiles, se pueden aplicar en la mayoría de los casos. También suelen ser útiles como valor inicial para métodos numéricos. Los estadísticos obtenidos mediante este método pueden no ser estimadores en el sentido estricto de la definición, es decir, podemos obtener estimaciones que queden fuera del espacio paramétrico. Bajo ciertas condiciones de regularidad, los estimadores de momentos tiene las siguientes propiedades.

1. Consistencia en error cuadrado medio (ECM) y en probabilidad.
2. Asintóticamente normales<sup>1</sup>.

---

<sup>1</sup>La suma de variables aleatorias a veces son MAN.

3. No necesariamente son invariantes.

Para lo anterior, las condiciones de regularidad de los estimadores se pueden ver [Azzalini \(1996\)](#).

**Definición 2.7** *Dada una familia de distribuciones  $P_\theta$ ,  $\theta \in \Theta \subseteq R^k$  con funciones de densidad  $f(x, \theta)$ , decimos que constituyen un modelo regular si se verifican las siguientes condiciones:*

- i) El modelo es identificable, en el sentido de que  $\forall \theta_1 \neq \theta_2 \in \Theta$ , ha de existir al menos un conjunto  $B$  del espacio muestral tal que  $P_{\theta_1}(B) \neq P_{\theta_2}(B)$ .*
- ii) El espacio paramétrico  $\Theta$  es un intervalo abierto de  $R^k$ .*
- iii) Todas las funciones de densidad especificadas por el modelo tienen el mismo soporte.*
- iv) Para la función  $f$ , la derivada con respecto a  $\theta$  y la integración con respecto a  $x$  puede intercambiarse hasta orden dos, concretamente:*

$$1. \int \frac{\partial}{\partial \theta} f(x, \theta) dx = \frac{\partial}{\partial \theta} \int f(x, \theta) dx$$

$$2. \int \frac{\partial^2}{\partial \theta^2} f(x, \theta) dx = \frac{\partial^2}{\partial \theta^2} \int f(x, \theta) dx$$

donde debe entenderse que si  $k > 1$ , entonces  $\frac{\partial^2}{\partial \theta^2}$  representa  $\frac{\partial^2}{\partial \theta_1 \partial \theta_1}$ .

A partir de las hipótesis anteriores pueden obtenerse las dos propiedades siguientes:

$$1. E_\theta \left[ \frac{\partial \log f(x, \theta)}{\partial \theta} \right] = 0$$

$$2. E_\theta \left[ \left( \frac{\partial \log f(x, \theta)}{\partial \theta} \right)^2 \right] = -E_\theta \left[ \frac{\partial^2 \log f(x, \theta)}{\partial \theta^2} \right]$$

## MOM para la distribución Pareto IV

Para el caso de estudio, sea  $X_1, \dots, X_n$  una muestra aleatoria de tamaño  $n$  donde cada  $X_i$  con  $(i = 1, \dots, n)$  sigue una distribución Pareto IV( $0, \sigma, \gamma, \alpha$ ) se desea estimar los parámetros  $\sigma$ ,  $\gamma$  y  $\alpha$  por el método de momentos. En este caso los momentos naturales de  $X$  son:

$$\begin{aligned}\mu_1(\underline{\theta}) = E(X) &= \frac{\sigma\Gamma(\alpha - \gamma)\Gamma(1 + \gamma)}{\Gamma(\alpha)}, \\ \mu_2(\underline{\theta}) = E(X^2) &= \frac{\sigma^2\Gamma(\alpha - 2\gamma)\Gamma(1 + 2\gamma)}{\Gamma(\alpha)}, \\ \mu_3(\underline{\theta}) = E(X^3) &= \frac{\sigma^3\Gamma(\alpha - 3\gamma)\Gamma(1 + 3\gamma)}{\Gamma(\alpha)}.\end{aligned}$$

Igualando los momentos muestrales con los poblacionales se tiene el sistema de ecuaciones a resolver:

$$\frac{1}{n} \sum_i^n x_i = \frac{\sigma\Gamma(\alpha - \gamma)\Gamma(1 + \gamma)}{\Gamma(\alpha)}, \quad (2.14)$$

$$\frac{1}{n} \sum_i^n x_i^2 = \frac{\sigma^2\Gamma(\alpha - 2\gamma)\Gamma(1 + 2\gamma)}{\Gamma(\alpha)}, \quad (2.15)$$

$$\frac{1}{n} \sum_i^n x_i^3 = \frac{\sigma^3\Gamma(\alpha - 3\gamma)\Gamma(1 + 3\gamma)}{\Gamma(\alpha)}. \quad (2.16)$$

Como se puede observar la solución de este sistema de ecuaciones no es una tarea fácil, ya que en el sistema se encuentra involucrada la función gama en términos de los parámetros de interés. Por lo anterior se recurre al uso de métodos numéricos para encontrar una solución. En particular en este trabajo de investigación se probaron varios algoritmos implementados en el paquete estadístico ([R Core Team, 2016](#)). Los cuales se presentan en la sección [3.2](#).

## Método de momentos fraccionarios (MMFR)

Una variante del MOM es el uso de momentos fraccionarios en lugar de los momentos naturales. Aún cuando el método de momentos presenta muchas ventajas, puede ser ineficaz para un problema en particular. Sin embargo, en ocasiones mediante una adecuada elección de los momentos fraccionarios puede aumentar la eficiencia de la estimación; por lo que en esta investigación se exploró este caso.

Dada  $X_1, \dots, X_n$  una m.a de tamaño  $n$  donde cada  $X_i$  con  $(i = 1, \dots, n)$  sigue una



distribución Pareto IV(0,  $\sigma$ ,  $\gamma$ ,  $\alpha$ ), con base en la ecuación 2.13, se plantea el siguiente sistema de ecuaciones:

$$\frac{1}{n} \sum_i^n x_i^{1/2} = \frac{\sigma^{1/2} \Gamma(\alpha - 1/2\gamma) \Gamma(1 + 1/2\gamma)}{\Gamma(\alpha)}, \quad (2.17)$$

$$\frac{1}{n} \sum_i^n x_i^{1/3} = \frac{\sigma^{1/3} \Gamma(\alpha - 1/3\gamma) \Gamma(1 + 1/3\gamma)}{\Gamma(\alpha)}, \quad (2.18)$$

$$\frac{1}{n} \sum_i^n x_i^{1/4} = \frac{\sigma^{1/4} \Gamma(\alpha - 1/4\gamma) \Gamma(1 + 1/4\gamma)}{\Gamma(\alpha)}. \quad (2.19)$$

### Método de máxima verosimilitud (MMV)

El método de máxima verosimilitud es quizá el método de estimación más usado en la actualidad. Lo anterior es posible a las excelentes propiedades que tienen estos estimadores. Fue propuesto por Fisher a principios del siglo pasado y la idea principal del método es que si un evento ocurre es por que su probabilidad de ocurrencia es más alta que la de aquellos eventos que no ocurrieron.

**Definición 2.8** La función de verosimilitud de  $n$  variables aleatorias  $X_1, \dots, X_n$ , se define como la densidad conjunta  $f_{X_1, \dots, X_n}(x_1, \dots, x_n; \underline{\theta})$ , la anterior expresión es vista como una función de  $\underline{\theta}$ . En particular si  $X_1, \dots, X_n$  es una muestra aleatoria de  $f(x; \underline{\theta})$ , entonces la función de verosimilitud es  $f(x_1; \underline{\theta}) \times f(x_2; \underline{\theta}) \times \dots \times f(x_n; \underline{\theta})$ .

La función de verosimilitud se denota como  $L(\underline{\theta}|x)$ ,  $L_{x_1, \dots, x_n}(\underline{\theta})$  o como  $L(\underline{\theta}|x_1, \dots, x_n)$ .

**Definición 2.9** Sea  $X_1, \dots, X_n$  una m.a. cuya función de distribución es  $F_X(x; \underline{\theta})$ , donde  $\underline{\theta}$  es un vector de parámetros con elementos  $(\theta_1, \theta_2, \dots, \theta_k)^t$ , con  $k \geq 1$ . Sea  $x_1, \dots, x_n$  una realización de  $X_1, \dots, X_n$ ; luego entonces la función de verosimilitud de una realización de una m.a. es  $L(\underline{\theta}|x_1, \dots, x_n)$ . Aquí,  $L(\underline{\theta}|x_1, \dots, x_n)$  es el valor de la función de densidad conjunta de una m.a.  $X_1, \dots, X_n$  tal que  $X_1 = x_1, \dots, X_n = x_n$ .

El estimador de máxima verosimilitud basado en una realización de una m.a.  $x_1, \dots, x_n$  es aquel valor de  $\underline{\theta}$  que maximiza la función de verosimilitud  $L(\underline{\theta}|x_1, \dots, x_n)$ , y se denota como  $\hat{\underline{\theta}}$ ; es decir,

$$L(\hat{\underline{\theta}}|x_1, \dots, x_n) = \max_{\underline{\theta} \in \Theta} L(\underline{\theta}|x_1, \dots, x_n). \quad (2.20)$$

La función de verosimilitud suele ser más fácil de maximizar si se considera su logaritmo que la propia función, de hecho cualquier función monótona podría ser de

utilidad. Formalmente, este criterio de trabajo es correcto, ya que, como la verosimilitud es no negativa, para obtener el máximo basta con trabajar donde es estrictamente positiva (el máximo no se encontrará nunca donde la verosimilitud es igual a cero) con lo que el logaritmo de la función de verosimilitud está definida. En la práctica es tan habitual que se utiliza una notación específica para dicho logaritmo, que se denomina log-verosimilitud. Se escribe como  $l(\underline{\theta}|x_1, \dots, x_n)$  o de forma abreviada  $l(\underline{\theta})$ . En concreto se tiene que

$$l(\underline{\theta}|x_1, \dots, x_n) = \ln L(\underline{\theta}|x_1, \dots, x_n).$$

En ocasiones la verosimilitud y su logaritmo suelen ser derivables como función de  $\underline{\theta}$  por lo que en estos casos se puede hacer uso del cálculo para encontrar el máximo; es decir, el estimador de máxima verosimilitud.

A continuación ilustraremos el método de máxima verosimilitud para estimar los parámetros de la distribución Pareto IV. El caso de la distribución Pareto IV con tres parámetros desconocidos  $\sigma$ ,  $\gamma$  y  $\alpha$  es el que hasta ahora algunos investigadores han sugerido abordar; lo anterior debido a que es muy conveniente suponer que el parámetro de localización es conocido, de lo contrario la distribución Pareto IV es una función que no cumple las condiciones de regularidad, lo cual impide que los métodos tradicionales conserven sus propiedades (Serfling, 1980).

Sea  $X_1, \dots, X_n$  una m.a. de tamaño  $n$  de una distribución Pareto IV( $0, \sigma, \gamma, \alpha$ ), la función de log-verosimilitud esta dada por,

$$\begin{aligned} l(0, \sigma, \gamma, \alpha) &= \left(\frac{1}{\gamma} - 1\right) \sum_{i=1}^n \log\left(\frac{x_i}{\sigma}\right) \\ &- (\alpha + 1) \sum_{i=1}^n \log\left(1 + \left(\frac{x_i}{\sigma}\right)^{1/\gamma}\right) \\ &- n \log \gamma - n \log \sigma + n \log \alpha. \end{aligned}$$

De acuerdo al método de máxima verosimilitud se requiere encontrar los valores de  $\sigma$ ,  $\gamma$  y  $\alpha$  que maximicen la verosimilitud. Para lo anterior, se tiene que, para este caso, la verosimilitud es derivable con respecto a los parámetros y por lo tanto podemos usar el cálculo para obtener el máximo. Como es bien sabido en este caso se obtienen las derivadas parciales con respecto a cada parámetro, se igualan a cero y se obtienen un conjunto de ecuaciones cuya solución nos lleva a la obtención del máximo siempre y cuando se cumplan ciertas condiciones. Así, pues las ecuaciones a resolver son las

siguientes:

$$\hat{\sigma} = \frac{1 + \hat{\alpha}}{n} \sum_{i=1}^n \left\{ \frac{\left(\frac{x_i}{\hat{\sigma}}\right)^{1/\hat{\gamma}-1} x_i}{\left[1 + \left(\frac{x_i}{\hat{\sigma}}\right)^{1/\hat{\gamma}}\right]} \right\}, \quad (2.21)$$

$$\hat{\gamma} = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\hat{\alpha} \left(\frac{x_i}{\hat{\sigma}}\right)^{1/\hat{\gamma}} - 1}{\left(\frac{x_i}{\hat{\sigma}}\right)^{1/\hat{\gamma}}} \log \left(\frac{x_i}{\hat{\sigma}}\right) \right\}, \quad (2.22)$$

$$\hat{\alpha} = \left( \frac{1}{n} \sum_{i=1}^n \log \left( 1 + \left(\frac{x_i}{\hat{\sigma}}\right)^{1/\hat{\gamma}} \right) \right)^{-1}. \quad (2.23)$$

Como se puede verificar la solución analítica de estas ecuaciones no es simple y hasta ahora se ha recurrido a métodos Newton y Cuasi-Newton para encontrar una solución aproximada. En esta investigación se exploran algunos algoritmos para encontrar la solución de estas ecuaciones.

Por otra parte, los estimadores de máxima verosimilitud bajo ciertas condiciones de regularidad tienen las siguientes propiedades.

1. Asintóticamente insesgados. Es decir, al incrementar el tamaño muestral, el valor esperado del estimador tiende a ser igual al valor del parámetro, para el caso univariado.
2. Asintóticamente normales. Es decir, la distribución del estimador tiende a una distribución Normal.
3. Asintóticamente de varianza mínima. Es decir, tienden a ser eficientes.
4. Invariantes.

### Método de distancia mínima (MDM)

El método de distancia mínima (MDM) fue propuesto por [Wolfowitz \(1957\)](#) en la década de los 50's y en los últimos años ha recibido una considerable atención debido a que los métodos mencionados anteriormente presentan algunos problemas que abordaremos en el desarrollo de este trabajo de investigación. Aunque el MDM se ha definido en muy diversos contextos y con diferentes de funciones de distancia, a continuación se presenta una definición general del método:

**Definición 2.10** Sean  $X_1, \dots, X_n$  una muestra aleatoria con función de distribución  $F(x, \underline{\theta})$ , donde  $\underline{\theta}$  es un vector de parámetros desconocidos que pertenece a  $\Theta \subseteq \mathbf{R}^k$

( $k \geq 1$ ), y donde  $F_n(x)$  es la función de distribución empírica de  $X_1, \dots, X_n$ . Si existe un  $\tilde{\theta} = \tilde{\theta}(X_1, \dots, X_n)$  en  $\Theta$ , tal que

$$d[F(x, \tilde{\theta}), F_n(x)] = \inf \{d[F(x; \theta), F_n(x)]; \theta \in \Theta\}$$

al estimador resultante se le conoce como estimador de distancia mínima de  $\underline{\theta}$ .

El MDM es útil como método alternativo de estimación cuando otros métodos no son aplicables. Como distancia se suele tomar la de Kolmogorov,

$$d_K(F_n, F_\theta) = \sup_{-\infty < x < \infty} |F_n(x) - F_\theta(x)|,$$

o la de Cramér-von Mises

$$d_c(F_n, F) = \int_{-\infty}^{\infty} [F_n(x) - F_\theta(x)]^2 w_\theta(x) dF_\theta(x).$$

Sin embargo, existen una gran diversidad de distancias que han sido usadas para la implementación de este método. En nuestro caso, la distancia utilizada para obtener el estimador por el MDM es:

$$\begin{aligned} Q = d(F_n(x_i) - F_X(x_i; \theta)) &= \sum_{i=1}^n [F_n(x_i) - F_X(x_i; \theta)]^2 \\ &= \sum_{i=1}^n \left[ F_n(x_i) - 1 + \left[ 1 + \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \right]^{-\alpha} \right]^2. \end{aligned} \quad (2.24)$$

De acuerdo al método distancia mínima se desea encontrar los valores de  $\sigma$ ,  $\gamma$  y  $\alpha$  que minimicen la ecuación (2.24). Para lo anterior, se obtienen las derivadas parciales con respecto a cada parámetro, se igualan a cero y se obtienen un conjunto de ecuaciones cuya solución nos lleva a la obtención del mínimo siempre y cuando se cumplan las siguientes condiciones:

- La distribución tiene que estar totalmente especificada.
- Solo se aplica a distribuciones continuas.

Así, pues las ecuaciones a resolver son las siguientes:

$$\frac{\partial Q}{\partial \sigma} = 2 \sum_{i=0}^n \left[ \left( F_n(x_i) - 1 + \left[ 1 + \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \right]^{-\alpha} \right) \times (-\alpha) \left[ 1 + \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \right]^{-\alpha-1} \times \frac{1}{\gamma} \left( \frac{x_i}{\sigma} \right)^{1/\gamma-1} \times \left( -\frac{1}{\sigma^2} \right) \right] = 0 \quad (2.25)$$

$$\frac{\partial Q}{\partial \gamma} = 2 \sum_{i=0}^n \left[ \left( F_n(x_i) - 1 + \left[ 1 + \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \right]^{-\alpha} \right) \times (-\alpha) \left[ 1 + \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \right]^{-\alpha-1} \times \left( -\frac{1}{\gamma^2} \right) \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \times \ln \left( \frac{x_i}{\sigma} \right) \right] = 0 \quad (2.26)$$

$$\frac{\partial Q}{\partial \alpha} = 2 \sum_{i=0}^n \left[ \left( F_n(x_i) - 1 + \left[ 1 + \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \right]^{-\alpha} \right) \times (-1) \left[ 1 + \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \right]^{-\alpha-1} \times \ln \left( \left[ 1 + \left( \frac{x_i}{\sigma} \right)^{1/\gamma} \right] \right) \right] = 0 \quad (2.27)$$

Del mismo modo que el estimador de máxima verosimilitud, el estimador de distancia mínima tiene muy buenas propiedades, naturalmente bajo ciertas condiciones de regularidad, entre estas aquí algunas:

1. Son asintóticamente normales. Es decir, la distribución del estimador tiende a una distribución Normal.
2. Son consistentes.
3. Son invariantes.

### 2.2.3. Propiedades de los estimadores

En la sección anterior se mencionan algunas de las propiedades de los estimadores, es por eso que en esta sección se definirán con precisión dichas propiedades. A continuación se darán algunas definiciones de las propiedades deseables de un estimador tales como **insesgadez**, **consistencia**, **varianza mínima** y **eficiencia**.

#### Insesgadez

**Definición 2.11** *Un estimador  $T(\underline{X}) = T(X_1, \dots, X_n)$  de la función  $g(\theta)$  del parámetro  $\theta$ , se le llama estimador insesgado de  $g(\theta)$  si  $E_\theta(T(\underline{X})) = g(\theta)$ , en caso contrario  $E_\theta(T) \neq g(\theta)$  se dice que el estimador es sesgado de  $g(\theta)$ . Donde  $E(T(\underline{X}))$  es la esperanza matemática de  $T(\underline{X})$ .*

**Definición 2.12** *Sea  $T(\underline{X})$  un estimador del parámetro  $g(\theta)$ , se llama sesgo a la función que representa la diferencia entre el valor esperado de un estimador y el parámetro,  $S(\theta) = E(T(\underline{X})) - g(\theta)$ .*

**Observación 2.1** *Sea  $T(\underline{X})$  un estimador del parámetro  $\theta$ , entonces si  $S(\theta) = E(T(\underline{X})) - g(\theta)$  vale cero, esto significa que  $T(\underline{X})$  es un estimador insesgado de  $\theta$ .*

Una de las propiedades deseables en los estimadores puntuales es que sean insesgados, pero el estimador puede no ser insesgado.

## Error cuadrado medio mínimo

Sería ideal escoger un estimador  $T(\underline{X})$  tal que minimice uniformemente en  $\theta$  el error cuadrado medio, el cual se define como:

**Definición 2.13** Sea  $X_1, \dots, X_n$  una muestra aleatoria de densidades  $f(x; \theta)$ ,  $\theta \in \Theta \subset \mathbb{R}$  y  $T(\underline{X}) = T(X_1, \dots, X_n)$  un estimador de  $g(\theta)$ , se llama error cuadrado medio de  $g(\theta)$  con respecto de  $T(\underline{X})$  a

$$ECM_{\theta}(T(\underline{X})) = E_{\theta}[(T(\underline{X}) - g(\theta))^2].$$

**Observación 2.2** Note que el error cuadrado medio (ECM) mide el error esperado del estimador  $T(\underline{X})$  y es una función que depende de  $\theta$  para cada estimador  $T(\underline{X})$ .

Así mismo sirve como un criterio de selección puesto que podemos elegir el estimador  $T(\underline{X})$  que minimice el error cuadrado medio. El criterio de selección es el siguiente: encontrar el  $T^*$  que sea uniformemente mínimo para cualquier valor de  $\theta$ .

**Definición 2.14** Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria de tamaño  $n$  de  $f(x; \theta)$ . Un estimador  $T^*$  de  $g(\theta)$  es denominado un estimador insesgado uniforme de mínima varianza<sup>2</sup> de  $g(\theta)$  si

1.  $T^*$  es insesgado para  $g(\theta)$ .
2. Para cualquier otro estimador insesgado de  $g(\theta)$ ,

$$\text{Var}(T^*) \leq \text{Var}(T^{**})$$

para todo  $\theta \in \Theta$ .

Es evidente que esta es una propiedad deseable para todo estimador insesgado. La razón de ello se fundamenta en el propio concepto de varianza, sobre la cual podemos justificar que al tener menos varianza, la distribución del estimador es menos dispersa, con lo cual el error que cometemos en el proceso de inferencia es más pequeño.

**Proposición 2.1** Sea  $X_1, \dots, X_n$  una muestra aleatoria de una población con parámetros  $\theta$ ,  $\theta \in \Theta \subset \mathbb{R}$  y sea  $T(\underline{X}) = T(X_1, \dots, X_n)$  un estimador de  $g(\theta)$ , entonces

$$ECM_{\theta}(T(\underline{X})) = \text{Var}_{\theta}(T(\underline{X})) + S_{\theta}^2(T(\underline{X}))$$

<sup>2</sup>En inglés se conoce como **UMVUE**: Uniformly minimum variance unbiased estimator.

## Consistencia

Una propiedad que deseamos que tuviera nuestro estimador es que conforme tengamos más información acerca del parámetro, nuestro estimador presentará estimaciones más precisas. A continuación presentamos algunas definiciones que nos ayudarán a ilustrar estas ideas de manera más formal.

**Definición 2.15** Sea  $(\Omega, \mathcal{A}, P)$  un espacio de probabilidad y sea  $\{X_n\}$  una sucesión de variables aleatorias y  $X$  otra variable ambas definidas en  $\Omega$ . Se dice que  $\{X_n\}$  converge en probabilidad a  $X$ , si y sólo si para cada  $\varepsilon > 0$  se cumple que

$$\lim_{n \rightarrow \infty} P[|X_n - X| > \varepsilon] = 0,$$

el cual tiene la siguiente notación  $X_n \xrightarrow{P} X$ .

**Definición 2.16** Se dice que una sucesión de estimadores  $\{T_n\}$  es consistente en probabilidad para estimar a  $\theta$  si y sólo si para cada  $\varepsilon > 0$  se cumple

$$\lim_{n \rightarrow \infty} P[|T_n - \theta| > \varepsilon] = 0.$$

Es decir, un estimador es consistente cuando al incrementar el tamaño de la muestra el estimador se acerca mucho al parámetro. Por ejemplo, la media muestral es un estimador consistente de la media poblacional.

**Definición 2.17** Se dice que una sucesión de estimadores  $\{T_n\}$  es consistente en error cuadrado medio (ECM) si se cumple que

$$\lim_{n \rightarrow \infty} ECM(T_n) = 0 \quad \text{para todo } \theta \in \Theta. \quad (2.28)$$

Probar consistencia en error cuadrado medio no es tarea fácil, sin embargo si un estimador es consistente en ECM, entonces también sería consistente en probabilidad, para ello veremos la afirmación del siguiente teorema.

**Teorema 2.1** Si una sucesión de estimadores  $\{T_n\}$  es consistente en error cuadrado medio (ECM), entonces también es consistente en probabilidad<sup>3</sup>.

**Observación 2.3** Se debe tener cuidado ya que un estimador puede ser consistente en probabilidad pero no en error cuadrado medio.

---

<sup>3</sup>Este resultado se conoce como condición fuerte (ECM).

## Invarianza

**Definición 2.18** Un grupo  $\mathcal{G}$  de transformaciones sobre el espacio de valores de una v.a.  $X$ . Se dice que  $\varphi$  es invariante bajo  $\mathcal{G}$  si

$$\varphi(g(x)) = \varphi(x) \quad \text{para todo } x \text{ y } g \in \mathcal{G}. \quad (2.29)$$

**Definición 2.19** Sea  $\mathcal{G}$  un grupo de transformaciones sobre el espacio de valores de la variable aleatoria  $X$ , se dice que una estadística  $T(X)$  es invariante maximal bajo  $\mathcal{G}$  si

- a)  $T$  es invariante,
- b)  $T$  es maximal, es decir,  $T(x_1) = T(x_2)$  implica  $x_1 = g(x_2)$  para alguna  $g \in \mathcal{G}$ .

Se ha establecido por tanto, la siguiente propiedad, que se denomina propiedad de invarianza de estimadores máximo verosímiles.

**Definición 2.20** Si  $\hat{\theta}$  es el estimador máximo verosímil de  $\theta$ , entonces  $g(\hat{\theta})$  es el máximo verosímil de  $g(\theta)$ .

**Definición 2.21** Los estimadores máximo verosímil son invariantes. Es decir, si  $\hat{\theta}$  es el estimador de máxima verosimilitud del parámetro  $\theta$  y  $g(\cdot)$  es una función con inversa única, entonces  $g(\hat{\theta})$ , es el estimador de máxima verosimilitud de  $g(\theta)$ .



# Capítulo 3

## Métodos

En este capítulo se describe la metodología utilizada para medir el desempeño de los estimadores clásicos que han sido usados para el caso de la distribución Pareto IV con tres parámetros desconocidos. También se presenta un algoritmo alternativo para usar el MDM, ya que los métodos y algoritmos clásicos presentan algunos problemas para obtener estimaciones.

### 3.1. Introducción

La evaluación del desempeño de los estimadores se realizó mediante simulación Monte Carlo. Los métodos Monte Carlo son algoritmos numéricos que utilizan una secuencia de números aleatorios para llevar a cabo una simulación estadística, con el fin de conocer algunas propiedades de las estadísticas bajo estudio. La simulación Monte Carlo sólo requiere conocer una función de densidad de probabilidad, la cual una vez que sea postulada o conocida puede ser usada para generar números aleatorios de esa distribución, y así realizar la simulación haciendo muestreos aleatorios de la misma.

Por otra parte, la obtención del estimador de máxima verosimilitud y la del estimador de mínima distancia involucra como su nombre lo indica la búsqueda del mínimo o del máximo de una función, lo cual es conocido en otros contextos como optimización. Ésta se puede hacerse mediante diversas técnicas, a continuación algunas de las más usadas:

- **Gradiente avanzado.** De este tipo de técnicas las más usadas son: el gradiente conjugado, Newton, Marquardt, quasi-Newton o variable métrica. El método de Newton por ejemplo usa el cálculo del gradiente a partir de la expansión de la

---

serie de Taylor e inversión de la matriz; converge si el punto inicial está cerca del óptimo. El método Marquardt usa pasos ascendentes mientras está lejos del óptimo, y Newton cuando está cerca del óptimo. Estos métodos son quizá los más usados debido posiblemente a su eficiencia para obtener los óptimos y a la solución de problemas multivariados.

- La búsqueda **aleatoria pura o fuerza bruta**. Evalúa en forma repetida la función mediante la selección aleatoria de valores de la variable independiente. Si un número suficiente de muestras es evaluado, el óptimo será eventualmente localizado. Trabaja en discontinuidades y funciones no diferenciables. Siempre encuentra el óptimo global, pero no es eficiente, ya que requiere mucho esfuerzo de implementación dado que no toma en cuenta el comportamiento de la función, ni los resultados de las iteraciones previas para mejorar la velocidad de convergencia. Un ejemplo de esta técnica es la búsqueda por malla <sup>1</sup>, donde las dimensiones de  $x$  y  $y$  se dividen en pequeños incrementos para crear una malla, la función se evalúa en cada nodo: entre más densa es la malla la probabilidad de localizar el punto óptimo es mayor (Lomax *et al.*, 2000).
- **Univariabilidad y búsquedas patrón**. Cambia una variable a la vez para mejorar la aproximación, mientras las otras variables se mantienen constantes, así el problema se reduce a una búsqueda unidimensional que se puede resolver por medio de diversos métodos (Newton, interpolación cuadrática).

En esta investigación se implementaron los métodos del gradiente avanzado y el método de aleatoriedad pura usando el paquete estadístico R.

## 3.2. Método de momentos

La evaluación del desempeño del estimador del método de momentos se realizó mediante simulación Monte Carlo, en la evaluación un paso importante es encontrar la solución de las ecuaciones 2.14, 2.15 y 2.16; el algoritmo usado para la evaluación se

---

<sup>1</sup>En inglés (Grid Search).

presenta a continuación:

<b>Datos:</b>	
	$n$ =tamaño de muestra
	$\sigma$ =parámetro de escala de la distribución Pareto IV
	$\gamma$ =parámetro de desigualdad de la distribución Pareto IV
	$\alpha$ =parámetro de forma de la distribución Pareto IV
	$N$ = número de réplicas Monte Carlo.
	$(\sigma = 1, 1.5, 2, 2.5, 3, 3.5, \gamma = 0.2, 0.4, 0.6, 0.8, 1, \alpha = 1, 1.5, 2, 2.5, 2.5, 3, 3.5)$
	Semilla= $(\sigma_I = 0.1, \gamma_I = 0.1, \alpha_I = 0.1)$ .
<b>Resultado:</b>	
	$\tilde{\sigma}$ = estimación del parámetro $\sigma$ ,
	$\tilde{\gamma}$ = estimación del parámetro $\gamma$ ,
	$\tilde{\alpha}$ = estimación del parámetro $\alpha$ .
<b>1</b>	<b>inicio</b>
<b>2</b>	<b>for</b> $i = 1$ a $N$
<b>3</b>	Generar una muestra aleatoria de tamaño $n$ de la
	<i>ParetoIV</i> ( $0, \sigma_0, \gamma_0, \alpha_0$ ).
<b>4</b>	Calcular los primeros tres momentos muestrales (2.10).
<b>5</b>	Construir el sistema de ecuaciones igualando los momentos muestrales
	con los momentos naturales poblacionales (2.14, 2.15, 2.16).
<b>6</b>	Obtener la solución del sistema de ecuaciones anterior (se obtienen las
	estimaciones de $\sigma$ , $\gamma$ y $\alpha$ ).
<b>7</b>	<b>fin</b>
<b>8</b>	<b>fin</b>
<b>9</b>	<b>devolver</b> $\tilde{\sigma}_i, \tilde{\gamma}_i, \tilde{\alpha}_i$
<b>10</b>	Calcular la varianza, sesgo y ECM con $\{\tilde{\sigma}^1, \dots, \tilde{\sigma}^N\}$ , $\tilde{\gamma} = \{\tilde{\gamma}^1, \dots, \tilde{\gamma}^N\}$ y
	$\tilde{\alpha} = \{\tilde{\alpha}^1, \dots, \tilde{\alpha}^N\}$ .
<b>11</b>	Se obtiene el porcentaje de convergencia numérica.

**Algoritmo 1:** Algoritmo para la evaluación del desempeño del estimador de momentos.

En particular para el paso 6 se usó la función *BBoptim* del paquete *BB* en *R*, el cual es uno de los mejores algoritmos de un método numérico para obtener las raíces de un sistema de ecuaciones no lineales (Varadhan *et al.*, 2009), como es el caso de las ecuaciones del MOM. Debido a que los resultados de la búsqueda de la solución de las ecuaciones 2.14, 2.15 y 2.16 no fué satisfactorio se implementó la opción del mismo método pero con momentos fraccionarios.

### 3.3. Método de momentos fraccionarios

En este caso, el único cambio respecto al algoritmo anterior fué la obtención de la solución del sistema de ecuaciones 2.17, 2.18 y 2.19. Así el algoritmo para este caso queda prácticamente sin modificación:

<b>Datos:</b>	
	$n$ =tamaño de muestra
	$\sigma$ =escala
	$\gamma$ =desigualdad
	$\alpha$ =forma
	$N$ = número de réplicas Monte Carlo.
	$(\sigma = 1, 1.5, 2, 2.5, 3, 3.5, \gamma = 0.2, 0.4, 0.6, 0.8, 1, \alpha = 1, 1.5, 2, 2.5, 2.5, 3, 3.5)$
	Semilla= $(\sigma_I = 0.1, \gamma_I = 0.1, \alpha_I = 0.1)$ .
<b>Resultado:</b>	
	$\tilde{\sigma}_{mmf}$ = estimación del parámetro $\sigma$ ,
	$\tilde{\gamma}_{mmf}$ = estimación del parámetro $\gamma$ ,
	$\tilde{\alpha}_{mmf}$ = estimación del parámetro $\alpha$ .
<b>1</b>	<b>inicio</b>
<b>2</b>	<b>for</b> $i = 1$ a $N$
<b>3</b>	Generar una muestra aleatoria de tamaño $n$ de la ParetoIV( $0, \sigma_0, \gamma_0, \alpha_0$ ).
<b>4</b>	Calcular los primeros tres momentos fraccionarios muestrales (2.10).
<b>5</b>	Construir el sistema de ecuaciones igualando los momentos muestrales con los momentos fraccionarios poblacionales (2.17, 2.18, 2.19).
<b>6</b>	Obtener la solución del sistema de ecuaciones anterior (se obtienen las estimaciones de $\sigma, \gamma$ y $\alpha$ ), usando la función <i>Optim</i> .
<b>7</b>	<b>fin</b>
<b>8</b>	<b>fin</b>
<b>9</b>	<b>devolver</b> $\tilde{\sigma}_i, \tilde{\gamma}_i, \tilde{\alpha}_i$
<b>10</b>	Calcular la varianza, sesgo y ECM con $\{\tilde{\sigma}^1, \dots, \tilde{\sigma}^N\}, \tilde{\gamma} = \{\tilde{\gamma}^1, \dots, \tilde{\gamma}^N\}$ y $\tilde{\alpha} = \{\tilde{\alpha}^1, \dots, \tilde{\alpha}^N\}$ .
<b>11</b>	Se obtiene el porcentaje de convergencia numérica.

**Algoritmo 2:** Algoritmo para la evaluación del desempeño del estimador de momentos fraccionarios.

### 3.4. Método de máxima verosimilitud

En el caso del estimador de máxima verosimilitud la evaluación de su desempeño requiere de la solución de las ecuaciones [2.21](#), [2.22](#) y [2.23](#), tarea que no resulta fácil llevar acabo y que requiere de la implementación de algún método de gradiente avanzado. En esta investigación se implementaron los algoritmos L–BFGS–B, BFGS y Nelder–Nead ver [Lu y Nocedal \(1995\)](#), para obtener la solución de las ecuaciones. De esta manera el algoritmo usado para la evaluación del desempeño del estimador

de máxima verosimilitud se presenta a continuación:

**Datos:**

$n$ =tamaño de muestra  
 $\mu$  =localidad  
 $\sigma$  =escala  
 $\gamma$  =desigualdad  
 $\alpha$  =forma  
 $N$ = número de réplicas Monte Carlo.  
 $(\sigma = 1, 1.5, 2, 2.5, 3, 3.5, \gamma = 0.2, 0.4, 0.6, 0.8, 1, \alpha = 1, 1.5, 2, 2.5, 2.5, 3, 3.5)$   
Semilla=  $(\sigma_I = 0.1, \gamma_I = 0.1, \alpha_I = 0.1)$ .

**Resultado:**

$\hat{\sigma}$  = estimación del parámetro  $\sigma$ ,  
 $\hat{\gamma}$  = estimación del parámetro  $\gamma$ ,  
 $\hat{\alpha}$  = estimación del parámetro  $\alpha$ .

```

1 inicio
2   for  $i = 1$  a  $N$ 
3     Generar una muestra aleatoria de tamaño  $n$  de la
       ParetoIV( $0, \sigma_0, \gamma_0, \alpha_0$ ).
4     Construir el sistema de ecuaciones 2.21, 2.22 y 2.23.
5     Obtener la solución del sistema de ecuaciones anterior (se obtienen las
       estimaciones de  $\sigma$ ,  $\gamma$  y  $\alpha$ ).
6   fin
7 fin
8 devolver  $\hat{\sigma}_i, \hat{\gamma}_i, \hat{\alpha}_i$ 
9 Calcular la varianza, sesgo y ECM con  $\{\hat{\sigma}^1, \dots, \hat{\sigma}^N\}, \hat{\gamma} = \{\hat{\gamma}^1, \dots, \hat{\gamma}^N\}$  y
    $\hat{\alpha} = \{\hat{\alpha}^1, \dots, \hat{\alpha}^N\}$ .
10 Se obtiene el porcentaje de convergencia numérica.
```

**Algoritmo 3:** Algoritmo para la evaluación del desempeño del estimador de máxima verosimilitud.

### 3.5. Método de distancia mínima

La evaluación del desempeño del estimador obtenido por el MDM tiene como uno de sus pasos importantes la solución del sistema de ecuaciones 2.25, 2.26 y 2.27, en este caso la solución se obtuvo usando la función *Optim* del paquete R (R Core Team, 2016), cabe mencionar que la función *Optim* usa un algoritmo Cuasi-Newton

(L–BFHGS–B) para obtener la solución de las ecuaciones. El algoritmo para la evaluación queda como sigue:

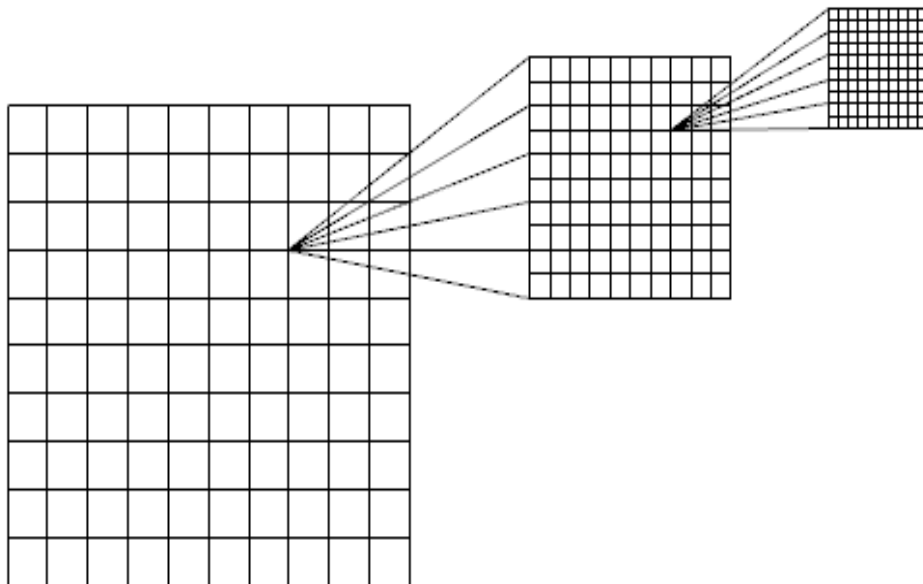
<p><b>Datos:</b>  <math>n</math> = tamaño de muestra  <math>\mu</math> = localidad  <math>\sigma</math> = escala  <math>\gamma</math> = desigualdad  <math>\alpha</math> = forma  <math>N</math> = número de réplicas Monte Carlo.  <math>(\sigma = 1, 1.5, 2, 2.5, 3, 3.5, \gamma = 0.2, 0.4, 0.6, 0.8, 1, \alpha = 1, 1.5, 2, 2.5, 2.5, 3, 3.5)</math>  Semilla = <math>(\sigma_I = 0.1, \gamma_I = 0.1, \alpha_I = 0.1)</math></p> <p>1</p> <p><b>Resultado:</b>  <math>\tilde{\sigma}_{MDM}</math> = estimación del parámetro <math>\sigma</math>,  <math>\tilde{\gamma}_{MDM}</math> = estimación del parámetro <math>\gamma</math>,  <math>\tilde{\alpha}_{MDM}</math> = estimación del parámetro <math>\alpha</math>.</p> <p>2 <b>inicio</b></p> <p>3     <b>for</b> <math>i = 1</math> a <math>N</math></p> <p>4         Generar una muestra aleatoria de tamaño <math>n</math> de la Pareto  IV(0, <math>\sigma_0, \gamma_0, \alpha_0</math>).</p> <p>5         Construir el sistema de ecuaciones (2.25), (2.26), (2.27).</p> <p>6         Obtener la solución del sistema de ecuaciones anterior (se obtienen las  estimaciones de <math>\sigma, \gamma</math> y <math>\alpha</math>).</p> <p>7     <b>fin</b></p> <p>8 <b>fin</b></p> <p>9 <b>devolver</b> <math>\tilde{\sigma}_i, \tilde{\gamma}_i, \tilde{\alpha}_i</math></p> <p>10 Calcular la varianza, sesgo y ECM con <math>\{\tilde{\sigma}^1, \dots, \tilde{\sigma}^N\}, \tilde{\gamma} = \{\tilde{\gamma}^1, \dots, \tilde{\gamma}^N\}</math> y  <math>\tilde{\alpha} = \{\tilde{\alpha}^1, \dots, \tilde{\alpha}^N\}</math>.</p> <p>11 Se obtiene el porcentaje de convergencia numérica.</p>	
--	--

**Algoritmo 4:** Algoritmo para la evaluación del desempeño del estimador de distancia mínima.

### 3.6. Método de distancia mínima obtenido por un algoritmo híbrido (Cuasi-Newton-Malla)

Los resultados de la evaluación de los estimadores de momentos y máxima verosimilitud son muy desalentadores debido a los problemas que presentan para encontrar el estimador de los parámetros de la Pareto IV con tres parámetros desconocidos. El MDM presenta un mejor comportamiento en términos de convergencia; sin embargo,

su desempeño en términos de ECM no es del todo satisfactorio. Es por esta razón que el presente trabajo propone la obtención del estimador por el MDM mediante la técnica de búsqueda usando una malla en conjunto con un método Cuasi-Newton.



**Figura 3.1:** Descripción gráfica de la búsqueda del estimador de mínima distancia usando malla.

La propuesta para la obtención del estimador por el MDM consiste en lo siguiente:

1. Generar una muestra aleatoria de tamaño  $n$  de la Pareto  $IV(0, \sigma_0, \gamma_0, \alpha_0)$ .
2. Obtener la función de distribución empírica  $F_n(\cdot)$  de la muestra generada en el punto 1.
3. Construir el sistema de ecuaciones 2.25, 2.26 y 2.27.
4. Obtener la solución del sistema de ecuaciones anterior mediante el método Cuasi-Newton, L-BFGS-B (se obtienen las estimaciones de  $\sigma$ ,  $\gamma$  y  $\alpha$ ).
5. Para la siguiente fase de estimación considerar que el valor obtenido de  $\gamma$  es conocido, usando su estimación obtenida en 4; es decir  $\hat{\gamma}$ .



- 
6. Crear una malla de valores para los parámetros  $\sigma$  y  $\alpha$  sobre el espacio paramétrico de estudio. Donde  $\sigma \in \{0, \dots, 10\}$  y  $\alpha \in \{0, \dots, 10\}$ .
  7. Evaluar la función de distribución en cada uno de los puntos de la malla.
  8. Obtener la distancia entre la función de distribución empírica y su función de distribución evaluada en los puntos de la malla .
  9. Obtener el mínimo de las distancias obtenidas en 8 y sus correspondientes coordenadas  $(\tilde{\sigma}_{1malla}, \tilde{\alpha}_{1malla})$ .
  - 10 Con el valor de  $\tilde{\sigma}_{1malla} \pm \Delta_i$  y  $\tilde{\alpha}_{1malla} \pm \Delta_i$ ,  $\Delta_i = 0.01 \times i$ ,  $i = 1, \dots, 10$  obtenido en 9, construir una nueva malla alrededor de este punto y obtener  $\tilde{\sigma}_{2malla}$  y  $\tilde{\alpha}_{2malla}$  que minimiza la distancia (punto 8) en esta región.
  - 11 Con el valor de  $\tilde{\sigma}_{2malla} \pm \Delta_j$  y  $\tilde{\alpha}_{2malla} \pm \Delta_j$ ,  $\Delta_j = 0.001 \times j$ ,  $j = 1, \dots, 10$  obtenido en 10, construir una nueva malla alrededor de este último punto y obtener  $\tilde{\sigma}_{3malla}$  y  $\tilde{\alpha}_{3malla}$  que minimiza la distancia (punto 8) en esta región.

La ilustración de los puntos 9, 10 y 11 se pueden ver en la figura 3.1.

De esta forma la estimación obtenida en 4 y las obtenidas en 11 dan como resultado la estimación por el MDM para  $\sigma$ ,  $\gamma$  y  $\alpha$ , mediante el algoritmo híbrido.

A continuación el algoritmo para la evaluación del desempeño del estimador obtenido bajo la propuesta:

---

**Datos:**

$n$  = tamaño de muestra  
 $\mu$  = localidad  
 $\sigma$  = escala  
 $\gamma$  = desigualdad  
 $\alpha$  = forma  
 $N$  = número de réplicas Monte Carlo.  
 $(\sigma = 1, 1.5, 2, 2.5, 3, 3.5, \gamma = 0.2, 0.4, 0.6, 0.8, 1, \alpha = 1, 1.5, 2, 2.5, 2.5, 3, 3.5)$   
Semilla =  $(\sigma_I = 0.1, \gamma_I = 0.1, \alpha_I = 0.1)$

1

**Resultado:**

$\tilde{\sigma}_{malla}$  = estimación del parámetro  $\sigma$ ,  
 $\tilde{\gamma}_{malla}$  = estimación del parámetro  $\gamma$ ,  
 $\tilde{\alpha}_{malla}$  = estimación del parámetro  $\alpha$ .

2 **inicio**

3 **for**  $i = 1$  **a**  $N$

4     Generar una muestra aleatoria de tamaño  $n$  de la ParetoIV( $0, \sigma_0, \gamma_0, \alpha_0$ ).

5     Obtener la función de distribución empírica  $F_n(\cdot)$  de la muestra generada en 4.

6     Construir el sistema de ecuaciones 2.25, 2.26 y 2.27.

7     Obtener la solución del sistema de ecuaciones mediante el método Cuasi-Newton, L-BFGS-B. Construir una malla de valores para los parámetros  $\sigma$  y  $\alpha$  sobre el espacio paramétrico.

8     Considerar que el valor obtenido de  $\gamma$  es conocido, usando su estimación obtenida en el paso anterior; es decir  $\hat{\gamma}$ .

9     Evaluar la función de distribución en cada uno de los valores de la malla.

10     Obtener la distancia entre la función de distribución empírica y su función de distribución evaluada en los puntos de la malla.

11     Obtener la distancia mínima entre toda las distancias y el valor de  $\tilde{\sigma}_{1malla} \pm \Delta_i$  y  $\tilde{\alpha}_{1malla} \pm \Delta_i$ ,  $\Delta_i = 0.01 \times i$ ,  $i = 1, \dots, 10$  que minimizan la distancia.

12     Con el valor  $\tilde{\sigma}_{1malla} \pm \Delta_i$  y  $\tilde{\alpha}_{1malla} \pm \Delta_i$  obtenido en el paso anterior, construir una nueva malla alrededor de este punto y obtener  $\tilde{\sigma}_{2malla} \pm \Delta_j$  y  $\tilde{\alpha}_{2malla} \pm \Delta_j$ ,  $\Delta_j = 0.001 \times j$ ,  $j = 1, \dots, 10$  que minimiza la distancia en esta región.

13     Con el valor de  $\tilde{\sigma}_{2malla} \pm \Delta_i$  y  $\tilde{\alpha}_{2malla} \pm \Delta_i$  obtenido en el paso anterior, construir una nueva malla alrededor de este punto  $\pm 2$  unidades en décimas y obtener  $\tilde{\sigma}_{3malla}$  y  $\tilde{\alpha}_{3malla}$  que minimiza la distancia en esta región .

14 **fin**

15 **fin**

16 **devolver**  $\tilde{\sigma}_{imalla}$  y  $\tilde{\alpha}_{imalla}$

17 Calcular la varianza, sesgo y ECM con  $\{\tilde{\sigma}^1, \dots, \tilde{\sigma}^N\}$ ,  $\tilde{\gamma} = \{\tilde{\gamma}^1, \dots, \tilde{\gamma}^N\}$  y  $\tilde{\alpha} = \{\tilde{\alpha}^1, \dots, \tilde{\alpha}^N\}$ .

18 Se obtiene el porcentaje de convergencia numérica.

**Algoritmo 5:** Algoritmo para la evaluación del desempeño del estimador de distancia mínima obtenido mediante una técnica de aleatoriedad pura en combinación con un método Cuasi-Newton.

---

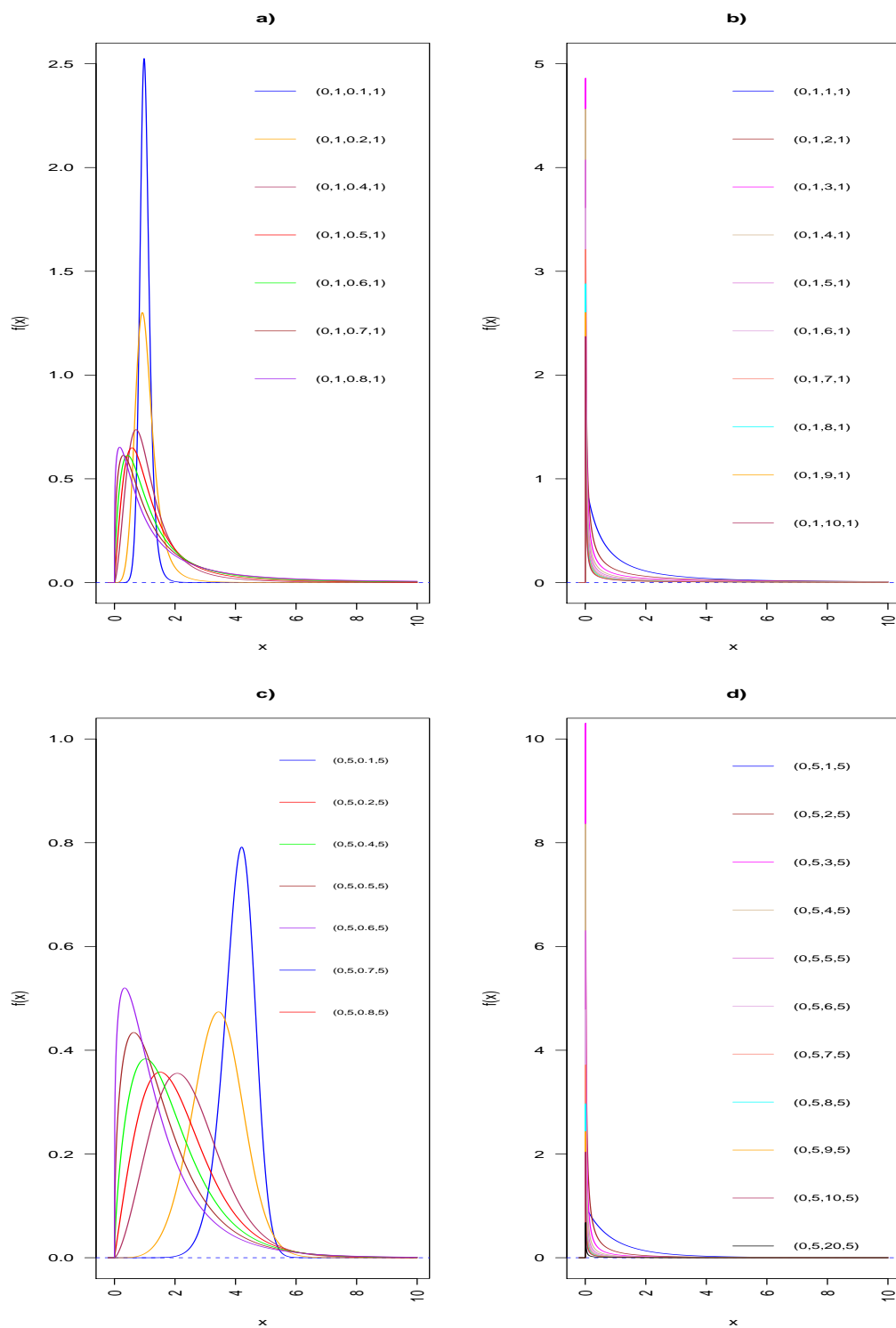
**Nota 3.1** *Se considera que el algoritmo 5 no converge si el estimador obtenido está en algún límite externo de la malla  $\sigma \in \{1, \dots, 12\}$  y  $\alpha \in \{1, \dots, 12\}$ . Naturalmente lo anterior sólo con la finalidad de evaluar el desempeño del estimador.*

### 3.7. Región estudiada del espacio paramétrico

Como se ha justificado anteriormente es necesario realizar una evaluación del desempeño de los estimadores tradicionales mediante simulación en el espacio paramétrico de la distribución Pareto IV. En el caso de la distribución Pareto IV con tres parámetros desconocidos, el espacio paramétrico es  $\mathcal{R}^3$  positivo; es decir, es una región (volumen) infinita. Sin embargo, un estudio de simulación muy pocas veces podrá ser totalmente exhaustivo; más bien, por lo general se realiza una inspección en parte del espacio paramétrico. La primera opción para la determinación de la región de estudio es investigar cuales son los valores de las estimaciones de los parámetros más comúnmente encontrados en la práctica; desafortunadamente para el caso de la Pareto IV con tres parámetros desconocidos existe muy poca información al respecto, de tal manera que no es posible obtener la región por esta vía. Otra opción es investigar la región en donde la distribución tiene potencial para modelación de eventos. Para lograr lo anterior se procedió a investigar las formas que puede adoptar el modelo al fijar dos parámetros y variar uno, encontrando los siguientes resultados:

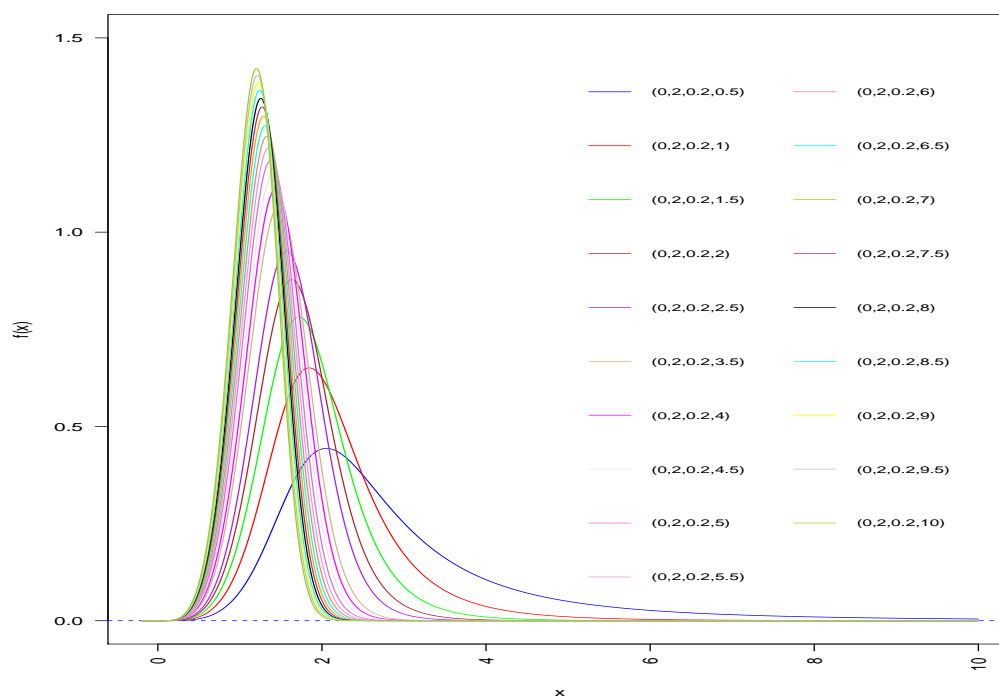
1. Fijando el valor de los parámetros  $\sigma$  y  $\alpha$  en la Figura (3.2) se puede observar que cuando los valores de  $\gamma$  son menores o iguales a uno, la flexibilidad del modelo se manifiesta presentando distribuciones sesgadas con colas de ligeras a pesadas; por lo que podemos decir que en estos casos el modelo es potencialmente útil. Sin embargo, cuando el parámetro  $\gamma$  es mayor a uno la distribución concentra casi toda su densidad en valores cercanos a cero y las colas de la distribuciones caen rápidamente a cero o muy cercana a cero; por lo anterior podemos decir que en estos casos la posibilidad de usar el modelo en situaciones prácticas es baja. También se puede observar que la densidad de la Pareto IV en la ecuación (2.1) tiende a cero cuando el valor de  $\gamma$  tiende a infinito; lo cual explica el comportamiento de la densidad para valores mayores a uno. De lo anterior se puede decir que una exploración del desempeño del estimador en el intervalo de  $0 < \gamma \leq 1$  será suficiente para modelar situaciones reales.

---

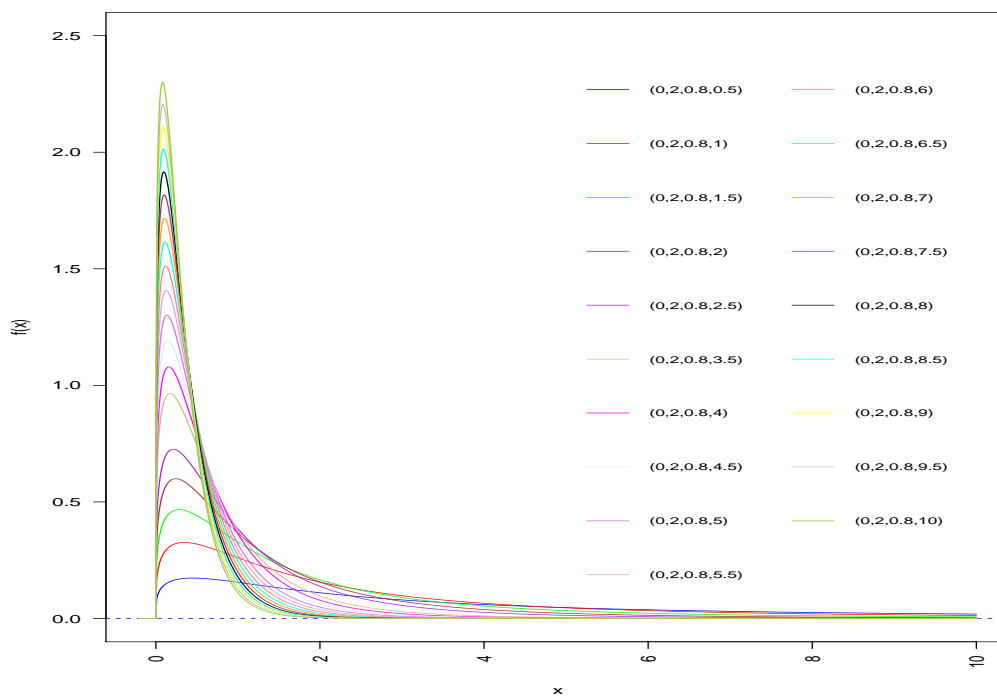


**Figura 3.2:** En a) y c) Función de densidad para valores de  $\gamma < 1$ , fijando  $\sigma = 1$  y  $\alpha = 1$ ; b) y d) Función de densidad para valores de  $\gamma \geq 1$ , fijando  $\sigma = 0.5$  y  $\alpha = 5$ .

2. Cuando fijamos los valores de los parámetros  $\sigma$  y  $\gamma$  se puede observar que si el valor de  $\alpha$  aumenta entonces la densidad nuevamente tiende a concentrarse en el intervalo de 0 a 1 Figura (3.4). Además, se puede notar que cuando el valor de  $\alpha > 8$  las colas de las distribuciones caen rápidamente alrededor de uno; mientras que para valores de  $\alpha < 8$ , se puede apreciar distribuciones sesgadas con colas pesadas. Estas últimas son frecuentemente usadas para modelar eventos extremos. Lo anterior indica que un intervalo razonable para el estudio de simulación en el caso del parámetro  $\alpha$  sería  $0 < \alpha < 8$ .

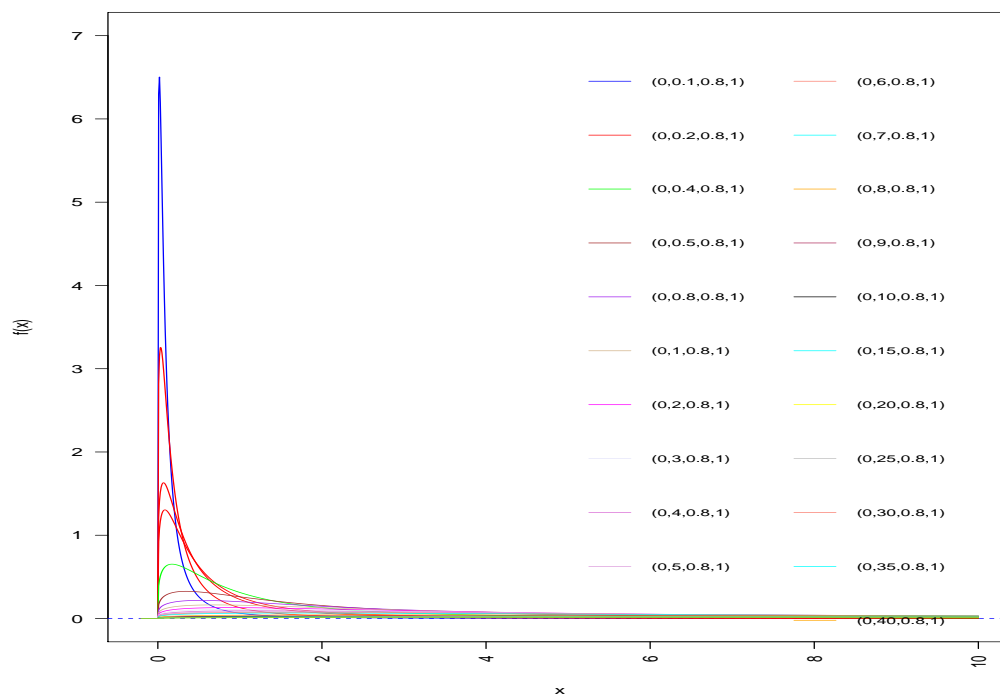


**Figura 3.3:** Función de densidad para valores de  $\sigma = 2$  y  $\gamma = 0.2$  fijos y variaciones en  $\alpha \in \{0.5, \dots, 10\}$ .

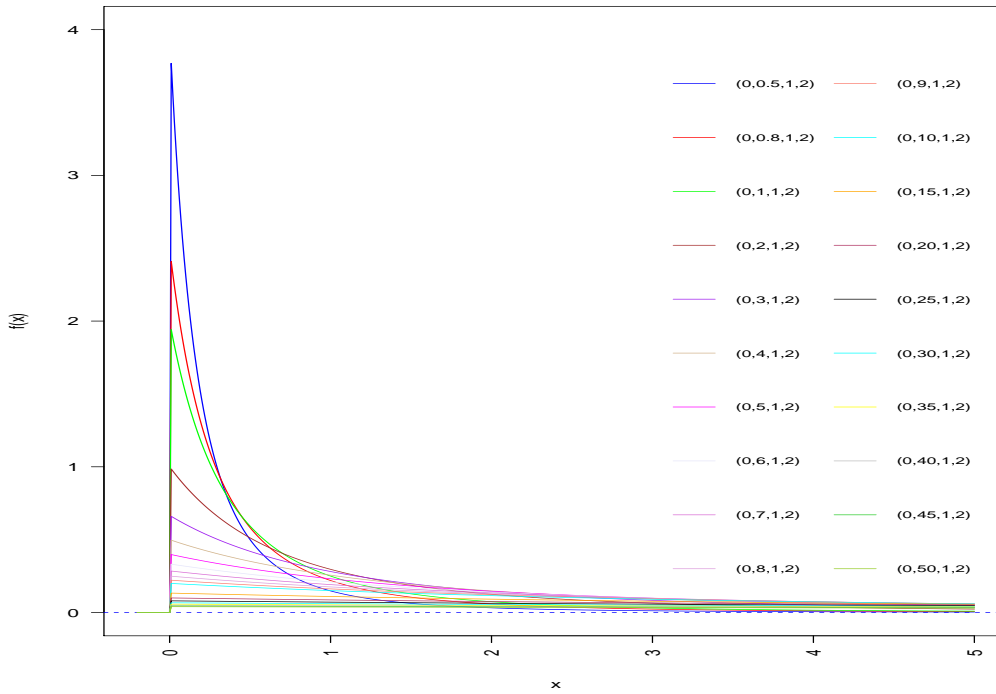


**Figura 3.4:** Función de densidad para valores de  $\sigma = 2$  y  $\gamma = 0.8$  fijos y variaciones en  $\alpha \in \{0.5, \dots, 10\}$ .

3. Finalmente, fijando el valor de los parámetros  $\gamma$  y  $\alpha$ , se puede observar que la forma de la distribución presenta clara tendencia a ser de cola super pesada al aumentar el valor de  $\sigma$  Figura (3.6); además la densidad presenta una forma platicúrtica, con valores de  $\sigma \geq 10$ , tan aplanada que la distribución en estos casos es muy cercana a la distribución uniforme. Por lo anterior, para el estudio de simulación se evaluaron los estimadores en el intervalo  $0 < \sigma \leq 10$ .



**Figura 3.5:** Función de densidad para valores de  $\gamma = 0.8$  y  $\alpha = 1$  fijos y variaciones en  $\sigma \in \{1, \dots, 40\}$ .



**Figura 3.6:** Función de densidad para valores de  $\gamma = 1$  y  $\alpha = 2$  fijos y variaciones en  $\sigma \in \{0.5, \dots, 50\}$ .

En resumen podemos decir que una región razonable para evaluar el desempeño de los estimadores de la distribución Pareto IV con tres parámetros desconocidos en  $\mathcal{R}^3$  mediante simulación es:  $\{(\sigma, \gamma, \alpha) \mid 0 < \sigma < 10, 0 < \gamma \leq 1, 0 < \alpha < 8\}$ . Por lo anterior en este trabajo de investigación, sin ser exhaustivo se investigó del comportamiento de los estimadores en esta región.



# Capítulo 4

## Resultados y Discusión

En este capítulo se presentan los resultados de la investigación del desempeño de los estimadores obtenidos por el MOM, MMV y MDM. Además, se presentan los resultados de un algoritmo propuesto para la obtención del estimador de MDM, ya que los métodos tradicionales para este caso presentan algunas dificultades para la obtención del estimador.

### 4.1. Método de momentos

Los resultados indican que el método de momentos presenta serias dificultades para resolver las ecuaciones 2.14, 2.15 y 2.16. Lo anterior debido a que dicho sistema de ecuaciones tienen incluida a la función gamma, ésta función presenta varias asíntotas en su rango, lo cual dificulta encontrar las raíces. Al estudiar el comportamiento del estimador en un amplio rango del espacio paramétrico, su desempeño fue muy deficiente cuando la semilla está lejos del valor verdadero del parámetro. Prácticamente resulta imposible obtener el error cuadrado medio del estimador debido a la baja tasa de convergencia de los métodos numéricos utilizados. La propuesta de utilizar los momentos fraccionarios también resultó infructuosa en el sentido de que la solución de las ecuaciones presenta el mismo comportamiento que el de momentos naturales.

### 4.2. Método de máxima verosimilitud

El desempeño de este estimador fue deficiente en casi toda la región del espacio paramétrico estudiado. El porcentaje de convergencia varía del 1 % hasta el 100 %. Desafortunadamente en el 80 % de los casos estudiados del espacio paramétrico, la

---

convergencia es menor al 20% (ver Tabla 4.1); lo cual nos indica que obtener el máximo de la función de verosimilitud para el caso de la distribución Pareto IV es muy difícil. No obstante que las tasas de convergencia fueron bajas, el algoritmo se diseñó de tal forma que se obtuvieron 1000 muestras Monte Carlo sin problemas de convergencia de tal forma que los ECM sean comparables. Es decir, 1000 réplicas de estimaciones válidas (cuando el algoritmo converge) para medir el desempeño de este estimador mediante el error cuadrado medio estimado y el valor esperado estimado. Lo primero que resalta en la Tabla 4.1 es el buen desempeño del estimador del parámetro  $\gamma$ ; su ECM oscila entre 0.001 y 0.04 con tamaño de muestra 50; lo cual indica que cuando el estimador de máxima verosimilitud existe el estimador de parámetro  $\gamma$  es muy preciso. Además, para este parámetro, el sesgo es muy pequeño en todas la tripletas estudiadas, con tendencia a cero a medida que el tamaño de muestra aumenta: sesgo=  $-0.0298$  con tamaño de muestra 50; sesgo=  $-0.0064$  con tamaño de muestra 100 y sesgo=  $-0.0023$  con tamaño de muestra 200.

Lo anterior indica que la propiedad de consistencia de este estimador se cumple. Sin embargo, los estimadores de  $\sigma$  y  $\alpha$  presentan ECMs muy altos. Por ejemplo, cuando el valor del parámetro  $\gamma$  se aproxima a 1 los ECMs de  $\sigma$  y  $\alpha$  aumentan en todos los casos estudiados. También cabe resaltar que en algunas tripletas fue imposible evaluar su desempeño debido a la alta frecuencia de no convergencia del algoritmo; este caso se presenta cuando el valor del parámetro es 0.4, aunque con tamaño de muestra de 50 este problema también se presenta para valores del parámetro 0.5, 0.6 y 0.7. (Ver Tablas 5.1, 5.2 y 5.3).

---

**Tabla 4.1:** Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto IV( $\mu_0, \sigma, \gamma, \alpha$ ) obtenidos con el método de máxima verosimilitud, con  $N = 1000$  réplicas Monte Carlo;  $n=50, 100, 200$ .

$(\mu_0, \sigma, \gamma, \alpha)$	n	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\alpha}$	Convergencia (%)
(0, 1, 0.4, 1)	50	1.1251	0.3866	1.2801	5.00
		(0.2667)	(0.0074)	(1.4189)	
	100	1.1100	0.4017	1.2080	59.00
		(0.1896)	(0.0041)	(0.7762)	
	200	1.0210	0.3954	1.0379	50.00
		(0.0372)	(0.0019)	(0.1083)	
(0, 2, 0.6, 2)	50	3.9095	0.5506	4.1067	1.00
		(33.2555)	(0.0253)	(34.1782)	
	100	1.4916	0.5608	1.3484	0.90
		(0.3658)	(0.0086)	(0.5574)	
	200	2.3804	0.6034	2.4636	1.70
		(1.3136)	(0.0042)	(1.9117)	
(0, 3, 0.7, 3)	50	NA	NA	NA	0.00
	100	3.1524	0.6637	3.3002	0.29
		(4.4104)	(0.0084)	(6.6130)	
	200	4.2264	0.6976	4.5515	10.00
		(21.6542)	(0.0039)	(46.7934)	

### 4.3. Método de distancia mínima

En este caso se presentarán los resultados del desempeño del estimador de distancia mínima usando el método Cuasi-Newton L-BFGS-B para obtener el estimador. Cabe mencionar que para lo anterior se utilizó la función *Optim* (R Core Team, 2016), la cual tiene varias opciones para obtener el estimador por el MDM, de las cuales el método que dio los mejores resultados fue el conocido como L-BFGS-B. Los resultados indican un buen desempeño del estimador respecto al porcentaje de convergencia. En el rango del espacio paramétrico estudiado, el porcentaje de convergencia oscila entre el 63% y 99%; lo anterior supera considerablemente el desempeño del estimador de máxima verosimilitud. Se puede observar Tabla 4.2, de forma similar al estimador de máxima verosimilitud, que el mejor desempeño en términos del ECM fue del estimador de  $\gamma$ . Su ECM oscila ente 0.002 y 0.044, con muestra de tamaño 50; 0.0148. Sin embargo, el desempeño de los estimadores de  $\sigma$  y  $\alpha$  es desalentador, presentando ECMs muy altos del orden de  $10^6$ . Se observa que para un par de valores fijos de  $\sigma$  y  $\alpha$ , digamos (1, 1) respectivamente, al aumentar el valor de  $\gamma$  los ECMs de  $\sigma$  y  $\alpha$  aumentan. También, se puede observar que a medida que el tamaño de muestra aumenta

el problema disminuye aunque la tendencia en el ECM es la misma. Los resultados en términos del valor esperado estimado tienen el mismo comportamiento que para el caso del ECM. Lo anterior indica la necesidad de explorar otras alternativas ya que hasta ahora los resultados no son satisfactorios. (Ver Tablas 5.4, 5.5 y 5.6).

**Tabla 4.2:** Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto IV( $\mu_0, \sigma, \gamma, \alpha$ ), obtenidos con el método de distancia mínima, con N=5000 réplicas Monte Carlo; n= 50, 100, 200.

$(\mu_0, \sigma, \gamma, \alpha)$	n	$\tilde{\sigma}_{mdm}$	$\tilde{\gamma}_{mdm}$	$\tilde{\alpha}_{mdm}$	Convergencia (%)
(0, 1, 0.4, 1)	50	1.4536	0.4057	2.3759	90.00
		(2.4527)	(0.0113)	(51.6333)	
	100	1.2473	0.4078	1.6052	97.00
		(0.7402)	(0.0061)	(7.4603)	
	200	1.1023	0.4049	1.2136	99.00
		(0.1436)	(0.0032)	(0.8174)	
(0, 2, 0.6, 2)	50	20.9767	0.5759	18.9971	76.00
		(169691.8482)	(0.0182)	(96625.6919)	
	100	3.7756	0.5940	5.0103	87.00
		(83.6697)	(0.0097)	(8325.4485)	
	200	2.9731	0.6005	3.5304	94.00
		(15.7237)	(0.0055)	(46.8097)	
(0, 3, 0.7, 3)	50	244.0255	0.6698	188.3179	69.00
		(2833270.7075)	(0.0245)	(1610321.3019)	
	100	20.5159	0.6781	17.2702	78.00
		(113298.5786)	(0.0128)	(51468.6342)	
	200	7.3729	0.6939	7.8656	87.00
		(12144.4984)	(0.0063)	8370.2804	

### 4.3.1. Método de distancia mínima obtenido por un algoritmo híbrido (Cuasi-Newton-Malla)

Debido a que el MDM presenta buenos resultados en términos de convergencia, se propone hacerle una pequeña modificación y medir su desempeño. La modificación consiste básicamente en obtener el estimador mediante la técnica de búsqueda con malla. Se creó el algoritmo para obtener el estimador mediante la técnica de malla auxiliada con el método Cuasi-Newton, L-BFGS-B, los resultados se muestran en la Tabla 4.3. Como se puede observar las tasas de convergencia son muy parecidas al caso anterior y van del 64% al 99.7%, por lo que su desempeño en este sentido es similar al algoritmo anterior. Además se puede observar que el estimador de  $\gamma$  presenta el mejor desempeño, similarmente a los algoritmos anteriores; en este caso sus ECM oscilan entre 0.002 a 0.06.

El valor esperado estimado de  $\gamma$  es muy cercano al valor del parámetro y se puede apreciar la propiedad de consistencia a medida que aumenta el tamaño de muestra.

Los resultados de los estimadores de  $\sigma$  y  $\alpha$  son bastante aceptables. Los ECMs de estos estimadores, en el espacio paramétrico estudiado para  $\sigma$ , oscilan entre (0.29) y (17.97).

Los resultados muestran que a medida que aumenta el tamaño de muestra los ECMs de estimador de  $\sigma$  disminuyen, lo cual se interpreta con un buen comportamiento del estimador. Los ECMs de estos estimadores, en el espacio paramétrico estudiado para  $\alpha$ , oscilan entre 0.70 y 11.26.

Los valores esperados estimados presentan un sesgo promedio para el caso  $\alpha$ , de 1.2080 con tamaño de muestra 50; de 0.8815 con tamaño de muestra 100 y 0.9861 con tamaño de muestra 200. Los resultados muestran que a medida que aumenta el tamaño de muestra los errores cuadrados medios de  $\tilde{\alpha}$  disminuyen, lo cual se interpreta con un buen comportamiento del estimador (Ver Tablas 5.7, 5.8 y 5.9).

**Tabla 4.3:** Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto IV  $(\mu_0, \sigma, \gamma, \alpha)$ , método de distancia mínima obtenido con un algoritmo híbrido (Cuasi-Newton-Malla), con  $N=5000$  réplicas Monte Carlo;  $n= 50, 100, 200$ .

$(\mu_0, \sigma, \gamma, \alpha)$	n	$\tilde{\sigma}_{mdm}$	$\tilde{\gamma}_{mdm}$	$\tilde{\alpha}_{mdm}$	Convergencia (%)
(0, 1, 0.4, 1)	50	1.5412	0.4717	2.4552	90.00
		(1.2630)	(0.0113)	(8.0305)	
	100	1.3617	0.4110	1.9202	97.30
		(0.6885)	(0.0059)	(4.3304)	
	200	1.1772	0.4067	1.4063	99.00
		(0.2914)	(0.0033)	(1.5491)	
(0, 2, 0.6, 2)	50	2.8512	0.5681	3.1614	76.00
		(6.9994)	(0.0211)	(9.4311)	
	100	2.9841	0.5967	3.4070	88.00
		(5.4804)	(0.0101)	(9.3752)	
	200	2.7373	0.5994	3.0725	94.90
		(3.5538)	(0.0055)	(6.5928)	
(0, 3, 0.7, 3)	50	3.9691	0.6713	4.0277	68.00
		(11.3411)	(0.0226)	(10.1425)	
	100	3.2165	0.6927	3.2268	89.20
		(9.0310)	(0.0134)	(8.1206)	
	200	3.8848	0.6878	4.0387	84.34
		(7.4305)	(0.0064)	(8.499)	

### 4.3.2. Ejemplos de Aplicación

En esta sección se presentan dos ejemplos. Una de las aplicaciones de la metodología propuesta es su uso para realizar pruebas de bondad de ajuste a un conjunto de datos.

**Ejemplo 4.1** *Se presenta a continuación los datos de ingresos de golfistas reportados en 1981 ver [Arnold \(2015\)](#). Representan los ingresos de golfistas durante 1980 de todos aquellos golfistas profesionales cuyos ingresos superiores a los 700,000 dolares durante ese período. El conjunto de datos se muestra en la [Tabla 4.4](#).*

**Tabla 4.4:** Datos del golfistas, ganancia en miles de dólares. Fuente: Golf magazine 1981 yearbook

3581	1433	1066	883	778
2474	1410	1056	878	778
2202	1374	1051	871	771
1858	1338	1031	849	769
1829	1208	1016	844	759
1690	1184	1005	841	753
1684	1171	1001	825	746
1627	1109	965	820	729
1537	1095	944	816	712
1519	1092	912	814	708

A continuación veremos si los datos del ingreso de golfistas sigue una distribución Pareto IV, usando la prueba de bondad de ajuste de Kolmogorov, cuyo juegos de hipótesis es:

$H_0$  : Los datos analizados siguen una distribución Pareto IV.

$H_1$  : Los datos analizados no siguen una distribución Pareto IV.

Para lo anterior primeramente se requiere obtener las estimaciones de los parámetros. Usando la propuesta se obtiene la estimación del vector  $\hat{\theta} = (0, 2.20, 0.03, 0.05)$ , con esta información se calcula la estadística de prueba  $D$ .

$$D = \max_{1 \leq i \leq n} |F(x_i; \underline{\theta}) - F_n(x_i)| = 0.9799. \quad (4.1)$$

La regla de decisión de la prueba de Kolmogorov es la siguiente:

Si  $D \leq D_\alpha$ , entonces aceptar  $H_0$ .

Si  $D > D_\alpha$ , entonces rechazar  $H_0$ .

Eligiendo  $\alpha = 0.05$  el nivel de significancia del contraste.

El valor de tabla para un tamaño de muestra 50 y  $\alpha = 0.05$ , se buscó en la tabla (Test de Kolmogorov-Smirnov sobre Bondad de Ajuste). Es decir  $D_\alpha = 0.18841$ .

Así,  $D = 0.9799 > D_\alpha = 0.18841$ , entonces se rechaza  $H_0$ .

Por lo tanto, con un nivel de significancia  $\alpha = 0.5$ , se rechaza que los datos siguen una distribución Pareto IV.

**Ejemplo 4.2** *Se tiene 157 observaciones que representan el ingreso personal total resultante para las poblaciones de alguna de los 254 condados de Texas en 1969. El conjunto de datos representan todos los condados de Texas en el que el total de ingresos personal es superior a 200,000.00 ver [Arnold \(2015\)](#).*

A continuación veremos si los datos del ingreso personal de ciudadanos de Texas sigue una distribución Pareto IV, usando la prueba de bondad de ajuste de Kolmogorov, cuyo juego de hipótesis es:

$H_0$  : Lo datos analizados siguen una distribución Pareto IV.

$H_1$  : Los datos analizados no siguen una distribución Pareto IV.

Para lo anterior primeramente se requiere obtener las estimaciones de los parámetros. Usando la propuesta se tiene que  $\hat{\underline{\theta}} = (0, 40.80, 0.11, 0.30)$ , con esta información se calcula la estadística de prueba  $D$ .

$$D = \max_{1 \leq i \leq n} |F(x_i; \underline{\theta}) - F_n(x_i)| = 0.2665 \quad (4.2)$$

La regla de decisión de la prueba de Kolmogorov es la siguiente:

Si  $D \leq D_\alpha$ , entonces aceptar  $H_0$ .

Si  $D > D_\alpha$ , entonces rechazar  $H_0$ .

Eligiendo  $\alpha = 0.05$  el nivel de significancia del contraste.

El valor de tabla para un tamaño de muestra 157 y  $\alpha = 0.05$ , se buscó en la tabla (Test de Kolmogorov-Smirnov sobre Bondad de Ajuste). Es decir,

$$D_\alpha = \frac{1.36}{\sqrt{n}} = \frac{1.36}{157} = 0.10853.$$

Así,  $D = 0.2665 > D_\alpha = 0.1085$ , entonces se rechaza  $H_0$ .

Por lo tanto, con un nivel de significancia  $\alpha = 0.05$ , se rechaza que los datos siguen una distribución Pareto IV.



# Capítulo 5

## Conclusiones

El desempeño de los estimadores obtenidos por el método de momentos (MOM), el método de momento fraccionario (MMF) y el método de máxima verosimilitud (MMV) no son satisfactorios para el caso de la distribución Pareto IV( $\mu_0, \sigma, \gamma, \alpha$ ) cuando los parámetros  $\sigma$ ,  $\gamma$  y  $\alpha$  son desconocidos. El método de distancia mínima (MDM), presenta mejores resultados en término de tasa de convergencia; sin embargo, los errores cuadrados medios en algunos casos son insatisfactorios. Se encontraron resultados satisfactorios en términos de los errores cuadrados medios y el porcentaje de convergencia de los estimadores obtenidos por el método de distancia mínima usando el algoritmo híbrido propuesto, lo anterior en la región del espacio paramétrico estudiado. Podemos decir que a propuesta de esta investigación es una buena alternativa para la estimación de los parámetros de la distribución Pareto IV cuando ésta sea utilizada para la modelación de eventos reales.

# Referencias

- Al-Hussaini, E. K. y Jaheen, Z. F. (1992). Bayesian estimation of the parameters, reliability and failure rate functions of the Burr type XII failure model. *Journal of statistical computation and simulation*, 41, 1-2, 31–40.
- Al-Marzoug, A. y Ahmad, M. (1985). Estimation of parameters of Burr probability model using fractional moments. *Pakistan, L Statistics*, 1, 67–77.
- Arnold, B. C. (1985). *Pareto distribution*. Wiley Online Library.
- Arnold, B. C. (2015). *Pareto distribution*. Wiley Online Library.
- Azzalini, A. (1996). *Statistical inference based on the likelihood*, tomo 68. CRC Press.
- Burr, I. W. (1942). Cumulative frequency functions. *The Annals of mathematical statistics*, 13, 2, 215–232.
- Crosato, L. y Ganugi, P. (2007). Statistical regularity of firm size distribution: the Pareto IV and truncated Yule for Italian SCI manufacturing. *Statistical Methods and Applications*, 16, 1, 85–115.
- Dubey, S. D. (1968). A compound Weibull distribution. *Naval Research Logistics Quarterly*, 15, 2, 179–188.
- Feller, W. (1971). An introduction to probability and its applications, Vol. II. *Wiley, New York*.
- Ghosh, I. (2011). Inference for the Bivariate and Multivariate Hidden Truncated Pareto (type II) and Pareto (type IV) Distribution and Some Measures Of Divergence Related To Incompatibility of Probability Distribution.
- Harris, C. M. y Singpurwalla, N. D. (1969). On estimation in Weibull distributions with random scale parameters. *Naval Research Logistics Quarterly*, 16, 3, 405–410.

- Jalali, A. y Watkins, A. J. (2009). On maximum likelihood estimation for the two parameter Burr XII distribution. *Communications in Statistics Theory and Methods*, 38, 11, 1916–1926.
- Johnson, N. L., Kotz, S. y Balakrishnan, N. (1994). Continuous univariate distributions, vol. 1-2.
- Lomax, A., Virieux, J., Volant, P. y Berge-Thierry, C. (2000). Probabilistic earthquake location in 3d and layered models. En *Advances in seismic event location*, 101–134. Springer.
- Lu, R. H. B. P. y Nocedal, J. (1995). A Limited Algorithm for Bound Constrained Optimizatin. *Journal on Scientific and Statistical Computing*, 1190–1208.
- Munir, R., Saleem, M., Aslam, M. y Ali, S. (2013). Comparison of different methods of parameters estimation for Pareto Model. *Caspian Journal of Applied Sciences Research*, 2, 1.
- Newman, M. E. (2005). Power laws, Pareto distributions and Zipf’s law. *Contemporary physics*, 46, 5, 323–351.
- Owens, J. F., Reya, E. y Glück, M. (1978). Detailed quantum-chromodynamic predictions for high- $p_T$  processes. *Phys. Rev. D*, 18, 1501–1514.
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. Bibtex: r\_core\_team\_r: 2016.
- Robertson (1977). Fire managment notes us foreste service. 34, 24.
- Sarabia, J. M. (2008). Parametric lorenz curves: Models and applications. En *Modeling income distributions and Lorenz curves*, 167–190. Springer.
- Singh, S. K. y Maddala, G. S. (1978). A function for the size distribution of income. *Econometrica*, 963–970.
- Varadhan, R., Gilbert, P. *et al.* (2009). BB: An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function. *Journal of Statistical Software*, 32, 4, 1–26.
- Wolfowitz, J. (1957). The minimum distance method. *Ann. Math. Statist.*, 28, 75–88.

# Anexos

## Anexo A: Tablas

En este apartado se presentan las tablas que se obtuvieron como resultado del estudio de simulación en este trabajo de investigación. Las tablas presentadas se describen de la siguiente manera, contiene nueve columnas: en la primera columna se tiene el tamaño de las muestras 50, 100 y 200. Segunda columna contiene cuartetos formada para diferentes parámetros de  $(\mu_0, \sigma, \gamma, \alpha)$ , aquí el parámetro  $\mu_0$  es conocido en nuestro caso es igual a cero;  $\sigma$  toma valores igual a 1, 1.5, 2, 2.5, 3, 3.5,  $\gamma$  toma valores de 0.4, 0.5, 0.6, 0.7, 0.8, 1 y  $\alpha$  toma valores de 1, 1.5, 2, 2.5, 3, 3.5; en la tercer columna se tiene el valor esperado de la estimación encontrada del parámetro  $\sigma$ ; en la cuarta columna se tiene el valor esperado de la estimación encontrada del parámetro  $\gamma$ ; en la quinta columna se tiene el valor esperado de la estimación encontrada del parámetro alfa; en la sexta columna se tiene el valor del ECM encontrado para el parámetro  $\sigma$ ; en la séptima columna se tiene el valor del ECM encontrado para el parámetro  $\gamma$ ; en la octava columna se tiene el valor del ECM encontrado para el parámetro  $\alpha$ ; y finalmente en la novena columna se encuentra la tasa de convergencia expresada en porcentaje.













---

**Método de distancia mínima [L–BFGS–B](MALLA)**

**Tabla 5.7:** Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto IV( $\mu_0, \sigma, \gamma, \alpha$ ), obtenidos por el método de distancia mínima obtenido con un algoritmo híbrido (Cuasi–Newton–Malla), con N=5000 réplicas Monte Carlo; n=50.

n	$(\mu_0, \sigma, \gamma, \alpha)$	$\tilde{\sigma}$	$\tilde{\gamma}$	$\tilde{\alpha}$	$ECM(\tilde{\sigma})$	$ECM(\tilde{\gamma})$	$ECM(\tilde{\alpha})$	Convergencia (%)
50	(0,1,0.4,1)	1.5412	0.4071	2.4552	1.2630	0.0113	8.0305	90.10
	(0,1,0.5,1)	1.6592	0.5046	2.1265	2.7116	0.0166	6.2191	91.40
	(0,1,0.6,1)	2.0139	0.6080	2.1792	6.1630	0.0269	6.9183	93.40
	(0,1,0.7,1)	2.2736	0.7059	2.0294	8.9122	0.0359	5.2104	92.70
	(0,1,0.8,1)	2.4824	0.8267	1.9072	10.2930	0.0444	3.7874	93.20
	(0, 1, 1, 1)	2.5751	1.0063	1.6404	11.8577	0.0649	2.1496	90.28
	(0,1.5,1,1.5)	3.2567	0.9720	2.4159	14.9041	0.0519	4.5093	81.04
	(0,2,0.4,2)	2.4895	0.3801	3.6965	1.5825	0.0075	10.4112	73.00
	(0,2,0.5,2)	2.6013	0.4762	3.4524	2.8070	0.0117	9.7882	72.40
	(0,2,0.6,2)	2.8512	0.5681	3.1614	6.9994	0.0210	9.4311	76.10
	(0,2,0.7,2)	3.6398	0.6898	3.5668	14.0467	0.0268	11.2204	81.10
	(0,2,0.8,2)	4.0032	0.7824	3.6446	17.9775	0.0277	11.2638	82.40
	(0, 2, 1, 2)	3.6379	0.9481	2.9929	15.6573	0.0470	6.3767	75.30
	(0,2.5,1,2.5)	3.9765	0.9311	3.5028	15.9394	0.0423	8.1527	71.66
	(0,3,0.4,3)	3.0676	0.3638	3.8809	1.7568	0.0078	8.6772	64.40
	(0,3,0.5,3)	3.0994	0.4547	3.7054	3.1071	0.0106	8.7308	64.40
	(0,3,0.6,3)	3.4108	0.5553	3.8955	6.0737	0.0170	10.0578	64.90
	(0,3,0.7,3)	3.9691	0.6713	4.0277	11.3411	0.0226	10.1425	68.50
	(0,3,0.8,3)	4.5685	0.7613	4.4005	16.8358	0.0268	12.0710	74.60
	(0,3, 1, 3)	4.1336	0.9194	3.8023	15.5858	0.0411	8.6300	67.34
	(0,3.5,1,3.5)	4.1422	0.9055	4.0017	14.4391	0.0437	8.9164	65.16
	(0, 5, 1, 5)	4.5004	0.8941	4.6427	15.1209	0.0413	10.6277	61.32

---

**Tabla 5.8:** Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto IV( $\mu_0, \sigma, \gamma, \alpha$ ), obtenidos por el método de distancia mínima obtenido con un algoritmo híbrido (Cuasi–Newton–Malla), con N=5000 réplicas Monte Carlo; n=100.

n	$(\mu_0, \sigma, \gamma, \alpha)$	$\tilde{\sigma}$	$\tilde{\gamma}$	$\tilde{\alpha}$	$ECM(\tilde{\sigma})$	$ECM(\tilde{\gamma})$	$ECM(\tilde{\alpha})$	Convergencia (%)
100	(0,1,0.4,1)	1.3617	0.4107	1.9202	0.6885	0.0059	4.3304	97.30
	(0,1,0.5,1)	1.4494	0.5082	1.7382	1.3441	0.0098	3.3121	98.00
	(0,1,0.6,1)	1.6323	0.6113	1.6917	3.0411	0.0145	3.1739	98.20
	(0,1,0.7,1)	1.8978	0.7215	1.7363	5.3233	0.0196	3.3302	98.50
	(0,1,0.8,1)	1.9246	0.8200	1.5942	5.3445	0.0232	2.0483	98.70
	(0,1, 1, 1)	2.0879	1.0203	1.4347	6.9003	0.0364	1.0952	98.18
	(0,1.5,1,1.5)	3.1022	1.0022	2.2939	11.9173	0.0301	3.0192	92.62
	(0,2,0.4,2)	2.5671	0.3965	3.6606	1.4721	0.0040	9.3626	85.30
	(0,2,0.5,2)	2.6824	0.4933	3.4634	2.5448	0.0059	8.8464	85.80
	(0,2,0.6,2)	2.9841	0.5967	3.4070	5.4804	0.0101	9.3752	88.00
	(0,2,0.7,2)	3.2165	0.6927	3.2268	9.0310	0.0134	8.1206	89.20
	(0,2,0.8,2)	3.8390	0.8002	3.5138	14.2926	0.0169	9.2096	91.40
	(0, 2, 1, 2)	2.8274	0.9856	3.0837	1.5111	0.0258	2.3067	86.42
	(0,2.5,1,2.5)	4.200	0.9734	3.5757	15.025	0.0237	6.3229	81.46
	(0,3,0.4,3)	2.5671	0.3965	3.6606	1.4721	0.0040	9.3626	85.30
	(0,3,0.5,3)	2.6824	0.4933	3.4634	2.5448	0.0059	8.8464	85.80
	(0,3,0.6,3)	2.9841	0.5967	3.4070	5.4804	0.0101	9.3752	88.00
	(0,3,0.7,3)	3.2165	0.6927	3.2268	9.0310	0.0134	8.1206	89.20
	(0,3,0.8,3)	3.8390	0.8002	3.5138	14.2926	0.0169	9.2096	91.40
	(0, 3, 1, 3)	4.4346	0.9603	4.0047	14.4398	0.0224	7.3452	77.06
	(0,3.5,1,3.5)	4.6721	0.9528	4.3084	14.5587	0.0222	7.6922	74.36
	(0,5, 1, 5)	5.0476	0.9347	5.0133	13.9696	0.0211	9.2047	67.08

**Tabla 5.9:** Estimaciones puntuales (ECM) de los parámetros de la distribución Pareto IV( $\mu_0, \sigma, \gamma, \alpha$ ), obtenidos por el método de distancia mínima obtenido con un algoritmo híbrido (Cuasi-Newton-Malla), con N=5000 réplicas Monte Carlo; n=200.

n	$(\mu_0, \sigma, \gamma, \alpha)$	$\tilde{\sigma}$	$\tilde{\gamma}$	$\tilde{\alpha}$	$ECM(\tilde{\sigma})$	$ECM(\tilde{\gamma})$	$ECM(\tilde{\alpha})$	Convergencia (%)
200	(0,1,0.4,1)	1.1772	0.4067	1.4063	0.2914	0.0033	1.5491	99.00
	(0,1,0.5,1)	1.2210	0.5081	1.3484	0.4257	0.0050	1.0575	99.10
	(0,1,0.6,1)	1.2205	0.6050	1.2496	0.5717	0.0068	0.7075	99.20
	(0,1,0.7,1)	1.3599	0.7054	1.3067	1.4904	0.0096	1.0451	99.70
	(0,1,0.8,1)	1.3655	0.8101	1.2458	1.1088	0.0119	0.4716	99.30
	(0, 1, 1, 1)	1.5755	1.0162	1.2425	2.6616	0.0198	0.4820	99.50
	(0,1.5,1,1.5)	2.6007	1.0092	2.0458	7.0028	0.01636	1.7250	98.26
	(0,2,0.4,2)	2.5584	0.4031	3.5611	1.2255	0.0023	7.9722	93.70
	(0,2,0.5,2)	2.7155	0.5021	3.4208	2.1138	0.0036	7.1559	92.50
	(0,2,0.6,2)	2.7373	0.5994	3.0725	3.5538	0.0055	6.5928	94.90
	(0,2,0.7,2)	3.0676	0.7034	3.0828	7.1115	0.0077	6.7847	95.10
	(0,2,0.8,2)	3.5957	0.8081	3.3304	11.0952	0.0095	7.3314	96.80
	(0, 2, 1, 1)	2.8700	1.0138	1.1968	4.7582	0.0196	0.2833	99.46
	(0,3,0.4,3)	3.5382	0.3958	4.5963	1.4548	0.0018	8.8171	81.80
	(0,3,0.5,3)	3.7514	0.4951	4.5965	3.0666	0.0033	10.1100	84.60
	(0,3,0.6,3)	3.6717	0.5895	4.0664	4.5732	0.0045	8.3509	82.80
	(0,3,0.7,3)	3.8848	0.6878	4.0387	7.4305	0.0064	8.4999	84.30
	(0,3,0.8,3)	4.6661	0.7972	4.4702	13.1172	0.0085	9.8164	90.90
	(0, 3, 1, 3)	4.7068	0.9878	4.1583	13.6047	0.0124	6.0853	86.62
	(0,3.5,1,3.5)	5.1293	0.9852	4.6386	13.9376	0.0116	6.9361	86.00

## Anexo B: Programas para la estimación de los parámetro en paquete R

En este apartado se presenta el listado de programas sentencias que implementan los algoritmos presentados en el Capítulo 3, resultando el hecho que para cada variación de las combinaciones de los parámetros de escala  $\sigma$ , desigualdad  $\gamma$  y forma  $\alpha$  de la distribución Pareto IV debe considerarse modificaciones en las primeras líneas.

### Programa MVM

---

```

1 ### el programa solo para m\`axima verosimilitud
2 library(VGAM)
3 library(AdequacyModel)

```

---

---

```
4 #rm(list=ls()) # borra
5 #####
6
7 # Pareto IV funci\on de densidad de probabilidad
8 pdf_paretoIV=function(par,x){
9   tet=par[1]
10  c=par[2]
11  a=par[3]
12  a*((x/tet)^((1/c)-1))/((tet*c)*((1+(x/tet)^(1/c)))^(a+1))
13 }
14 # ParetoIV funci\on de dstrubuci\on acumulada
15 cdf_paretoIV<-function(par,x){
16   tet=par[1]
17   c=par[2]
18   a=par[3]
19   1-(1+((x/tet)^(1/c)))^(-a)
20 }
21
22 #vec <- NULL
23 mles<-NULL
24 k<-200#tama\~no de la muestra
25 sigma <- 1.5; gama <- 1; alfa <- 1.5
26 N<-100
27 for (j in 1:N){
28   n<-10
29   matrizp<-matrix(0,n,4)
30   for(i in 1:n){
31     datos<- rparetoIV(k,0,sigma,gama,alfa)
32     res<-goodness.fit(pdf=pdf_paretoIV, cdf=cdf_paretoIV, starts=c
33       (0.1,0.1,0.1),
34       data=datos, method="L-BFGS-B", domain=c(0,Inf),
35       mle=NULL)
36     matrizp[i,]<-c(res$mle[1],res$mle[2],res$mle[3],res$Convergence)
37   }
38   mles<-rbind(mles,matrizp)
39 }
```

---

---

```
38 }
39 mles
40
41 ##
42 #####
43 indice=(mles[,4]==1) # se\~nala los que no convergen
44 queda<-mles[!indice,] # quita los etiquetados con uno
45
46 indexedos=(queda[,4]==52)
47 mequeda<-queda[!indexedos,] # quita uno y 52
48 #####
49
50 mes<-mean(mequeda[,1]); meg<-mean(mequeda[,2]) ;mea<-mean(mequeda[,3])
51 #-----
52 vars<-var(mequeda[,1]); varg<-var(mequeda[,2]); vara<-var(mequeda[,3])
53 #-----
54 sesgos<-(mes-sigma)^2; sesgog<-(meg-gama)^2; sesgoa<-(mea-alfa)^2
55 #-----
56 ecms<- vars+sesgos; ecmg<-varg+sesgog; ecma<-vara+sesgoa
57 conver<-length(mequeda[,3])/1000
58 #length(mequeda[,3])
59
60 vec<-c(mes,meg,mea,ecms,ecmg,ecma,conver)
61 #vec
62 vect<-NULL
63 vect<-rbind(vect,vec)
64 vect
65 #####
66 #dim(queda)
67
68
69 #ms<-mean(queda[,1]); mg<-mean(queda[,2]); ma<-mean(queda[,3])
70 #-----
71 #vars<-var(queda[,1]); varg<-var(queda[,2]); vara<-var(queda[,3])
72 #-----
73 #sesgos<-(ms-sigma)^2; sesgog<-(mg-gama)^2; sesgoa<-(ma-alfa)^2
```

---

---

```

74 #-----
75 #ecm_s<- vars+sesgos; ecm_g<-varg+sesgog; ecm_a<-vara+sesgoa
76 #-----
77 #length(queda[,1])
78 #vec<-NULL
79 #convergencia<-length(queda[,3])/1000
80 #v<- c(ms,mg,ma,ecm_s,ecm_g,ecm_a,convergencia)
81 #vecmmv<-NULL
82 #vecmmv <- rbind(vecmmv,v)
83 #vecmmv

```

---

## Programa MDM

---

```

1 ##### Inicio de programa mdm
2 library(VGAM)
3 library(AdequacyModel)
4 library(fitdistrplus)
5 library(RODBC)
6 #rm(list=ls()) # borra
7 #####
8 iter=5000
9 #vect <- NULL
10
11 n=200
12 sigma=3
13 gamma=0.8
14 alpha=3
15 #####
16
17 mdm<-matrix(0,iter,4)
18 for(i in 1:iter){
19   y<- rparetoIV(n,0,sigma,gamma,alpha)
20
21   par<-c(0.1,0.1,0.1)
22   # Pareto IV funci'on de dstrubuci'on acumulada
23   cdf <-function(par,y){

```

---

---

```

24     x1=par[1]
25     x2=par[2]
26     x3=par[3]
27     1-(1+((y/x1)^(1/x2)))^(-x3)
28 }
29 Fny<-ecdf(y)(y)
30 Fn<-function (par){sum((cdf(par,y)-Fny) ^2) }
31 res1<-optim(par, Fn, gr=NULL, method=c("L-BFGS-B"), lower=0, upper=
      Inf, control=list(), hessian=FALSE)
32 mdm[i,]<-c(res1$par[1],res1$par[2],res1$par[3],res1$convergence)
33 }# termina el for
34 #####
35 indice=(mdm[,4]==1) # se\~nala los que no convergen
36 queda<-mdm[!indice,] # quita los etiquetados con uno
37 indexedos=(queda[,4]==52)
38 mequeda<-queda[!indexedos,] # quita uno y 52
39 #####
40
41 mes<-mean(mequeda[,1]);meg<-mean(mequeda[,2]);mea<-mean(mequeda[,3])
42 #-----
43 vars<-var(mequeda[,1]); varg<-var(mequeda[,2]); vara<-var(mequeda[,3])
44
45 sesgos<-(mes-sigma)^2; sesgog<-(meg-gamma)^2; sesgoa<-(mea-alpha)^2
46 #-----
47 ecms<- vars+sesgos; ecmg<-varg+sesgog; ecma<-vara+sesgoa
48 conv<-length(mequeda[,3])/5000
49 #length(mequeda[,3])
50
51 vec<-c(mes,meg,mea,ecms,ecmg,ecma,conv)
52 #vec
53 vect<-NULL
54 vect<-rbind(vect,vec)
55 vect

```

---

**Programa [L-BFGS-B (MALLA)]**

---



---

```
1 library(VGAM)
2 library(AdequacyModel)
3 library(fitdistrplus)
4 library(RODBC)
5 #rm(list=ls()) # borra
6 #####
7 iter=1000
8
9 n=200      # Tamano de muestra
10 sigma=5
11 gamma=1
12 alpha=5
13 #####
14
15 mdm<-matrix(0,iter,4)
16 for(j in 1:iter){
17   y<- rparetoIV(n,0,sigma,gamma,alpha)
18
19   par<-c(0.1,0.1,0.1)
20   # ParetoIV funci'on de dstrubuci'on acumulada
21   cdf <-function(par,y){
22     x1=par[1]
23     x2=par[2]
24     x3=par[3]
25     1-(1+((y/x1)^(1/x2)))^(-x3)
26   }
27   Fny<-ecdf(y)(y)
28   Fn<-function (par){sum((cdf(par,y)-Fny) ^2) }
29   res1<-optim(par, Fn, gr=NULL, method=c("L-BFGS-B"), lower=0, upper=
30     Inf, control=list(), hessian=FALSE)
31   #mdm[i,]<-res1$par
32 }# termina el for
33 #####
34 #mdm[1,]
```

---

---

```
35 #####
36 #####

37 #### Resultado con mallas o filtro
38
39 # ParetoIV funci\on de dstrubuci\on acumulada
40 #pr<-(0.1,0.1)
41 cdfd <-function(y,xs,xa){
42 1-(1+((y/xs)^(1/res1$par[2])))^(-xa)
43 }
44 #####
45 theta <- expand.grid(0:10,0:10)
46
47 z <- 0 #
48
49 for(i in 1:nrow(theta)){
50     z[i]<-sum((cdfd(y,theta[i,1], theta[i,2])-Fny)^2)
51 }
52 theta.hat <- theta[order(z)[1],] # da la posici\on donde se encuentra
    el m\inimo
53
54 lows <- theta.hat[,1]-2; ups <- theta.hat[,1]+2
55 lowa <- theta.hat[,2]-2; upa <- theta.hat[,2]+2
56 theta <- expand.grid((lows*10):(ups*10)/10, (lowa*10):(upa*10)/10)
57 ###
58 z <- 0
59
60 for(i in 1:nrow(theta)){
61     z[i]<-sum((cdfd(y,theta[i,1], theta[i,2])-Fny)^2)
62 }
63 theta.hat <- theta[order(z)[1],] # da la posici\on donde se encuentra
    el m\inimo
64
65 lows <- theta.hat[,1]-0.2; ups <- theta.hat[,1]+0.2
66 lowa <- theta.hat[,2]-0.2; upa <- theta.hat[,2]+0.2
```

---

---

```
67 theta <- expand.grid((lows*20):(ups*20)/20, (lowa*20):(upa*20)/20)
68
69
70 #####
71 z <- 0
72
73 for(i in 1:nrow(theta)){
74   z[i]<-sum((cdfd(y,theta[i,1], theta[i,2])-Fny)^2)
75 }
76 theta.hatres <- theta[order(z)[1],] # da la posici'on donde se
      encuentra el m'inimo
77
78 #
79 #mdm[j,]<-c(res1$par[1],res1$par[2],res1$par[3],theta.hatres[,1],theta
      .hatres[,2],res1$convergence)
80 mdm[j,]<-c(theta.hatres[,1],res1$par[2],theta.hatres[,2],
      res1$convergence)
81 } # termina el for
82 #####
83 #mdm
84
85 indice=(mdm[,4]==1) # se~nala los que no convergen
86 queda<-mdm[!indice,] # quita los etiquetados con uno
87 indexedos=(queda[,4]==52)
88 quedados<-queda[!indexedos,] # quita uno y 52
89 #####
90 #mdm[1:5,]
91 length(quedados[,1])
92 #colnames(quedados)<-c( "sigma","gamma","alfa", "sigmad","alfad","co")
93 colnames(quedados)<-c( "sigmad", "gamma","alfad","co")
94 datos<-data.frame(quedados)
95 subset(datos,sigmad<12) # quito los de mayores a 12 de sigma malla
96 quedato<-subset(datos,alfad<12) # quito los mayores a 12 de alfa malla
97 #
98 quedato<-data.frame(quedato)
99 #####minima distancia todo
```

---

---

```
100 dmes<-mean(quedato[,1]); dmeg<-mean(quedato[,2]); dmea<-mean(quedato
    [,3])
101
102 dvars<-var(quedato[,1]); dvarg<-var(quedato[,2]); dvara<-var(quedato
    [,3])
103
104 dsesgos<-(dmes-sigma)^2; dsesgog<-(dmeg-gamma)^2; dsesgoa<-(dmea-alpha
    )^2
105 #-----
106 decms<- dvars+dsesgos; decmg<-dvarg+dsesgog; decma<-dvara+dsesgoa
107 conv<-length(quedato[,3])/iter
108 vector<-c(dmes,dmeg,dmea,decms,decmg,decma,conv)
109 #vector
110 #####
111
112 dvect<-NULL
113 dvect<-rbind(dvect,vector)
114 dvect
```

---