

C I M A T

Centro de Investigación en Matemáticas A.C.

---

---

Inferencia Estadística en el  
Análisis de Máximos por Bloques:  
Comparación de Enfoques

T E S I S

que para obtener el grado de

Maestría en Ciencias con Especialidad en  
Probabilidad y Estadística

P R E S E N T A :

Alejandro Cruz Marcelo

DIRECTOR DE TESIS:

Dr. Joaquín Ortega Sánchez

Guamajuato, Gto., México

Agosto del 2006

## Agradecimientos

A mi familia por su paciencia.

Al Consejo Nacional de Ciencia y Tecnología (Conacyt) por financiar mis estudios de maestría. Este trabajo forma parte de los proyectos 05-02-K117-099 de Concyteg y 45974-F de Conacyt.

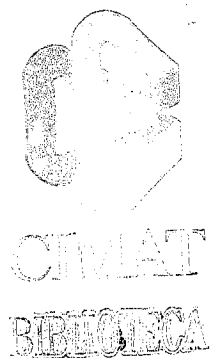
A la Dra. Eloísa Díaz-Francés y Eliud Vilchis por su ayuda para resolver algunos problemas numéricos que aparecieron en los cálculos involucrados en esta tesis. Al Dr. Jorge Domínguez, al Dr. Joaquín Ortega, al Dr. Luis Hernández, a la Dra. Eloísa Díaz-Francés y al FM Fabio Dávila por su apoyo para conseguir el equipo de cómputo que fue requerido para la realización de este trabajo.

A mi tutor el Dr. Joaquín Ortega Sánchez por su disposición para discutir cualquier aspecto relacionado con la tesis. A los integrantes del jurado la Dra. Graciela González Farías y la Dra. Eloísa Díaz-Francés por sus comentarios que mejoraron el contenido de este documento.

## Abstract

En muchas situaciones es de interés modelar los valores más grandes (o pequeños) que presenta un proceso estocástico. Un caso particular es cuando el proceso en cuestión está indexado por el tiempo, lo conforman variables aleatorias iid, y se tienen como "valores más grandes" a los máximos del proceso obtenidos en períodos cuya duración de tiempo es fija. En la teoría de valores extremos, un análisis bajo ese contexto se conoce como de máximos por bloques. Se sabe que, bajo ciertas condiciones, la distribución de esos máximos se puede aproximar por algún tipo de las Distribuciones de Valores Extremos (DVE's) -Fréchet, Gumbel y Weibull-, las cuales es posible representar en una sola familia cuya función de distribución se conoce como Distribución Generalizada de Valores Extremos (DGVE).

En la práctica se puede usar máxima verosimilitud para ajustar una distribución a los máximos y obtener estimaciones de cuantiles altos, sin embargo, se pueden seguir diferentes caminos para ello. Para calcular la función de verosimilitud es posible, por un lado, usar la función de densidad de la DGVE, y por el otro, considerar la función de densidad de algún tipo de las DVE's. Aunque con la segunda alternativa se debe escoger un tipo de DVE's, procediendo de esa forma se tienen modelos cuyos parámetros asociados son más fáciles de interpretar, y además, es posible que se trabaje con menos parámetros si el tipo de distribución Gumbel resulta adecuado. En este trabajo se presentan los resultados de un estudio basado en simulaciones en el que se compararon las estimaciones obtenidas bajo los enfoques mencionados, y donde para la selección de un tipo de las DVE's se utiliza la función de verosimilitud perfil.



# Índice general

Introducción	5
<b>1. Marco Teórico</b>	<b>7</b>
1.1. Teorema de Tipos para Extremos	7
1.2. Dominios de Atracción	11
1.2.1. Dominio de Atracción de las <i>DVE</i> Tipo Fréchet ( $\Phi_\alpha$ )	11
1.2.2. Dominio de Atracción de las <i>DVE</i> Tipo Weibull ( $\Psi_\alpha$ )	13
1.2.3. Dominio de Atracción de las <i>DVE</i> Tipo Gumbel ( $\Lambda$ )	13
1.3. Distribución Generalizada de Valores Extremos	16
1.4. Dominios de Atracción para Algunas Distribuciones Conocidas	17
1.5. Niveles y Períodos de Retorno	17
<b>2. Modelación Estadística para Eventos Extremos</b>	<b>20</b>
2.1. Selección de Modelo	21
2.2. Herramientas Basadas en la Verosimilitud	22
2.2.1. Estimadores de Máxima Verosimilitud	23
2.2.2. Intervalos de Verosimilitud y Aproximación $\chi^2$	25
2.2.3. Verosimilitud Perfil	27
2.2.4. Aplicación a las Distribuciones de Valores Extremos	28
2.3. Enfoques para Hacer Inferencia Usando la Verosimilitud	30
<b>3. Comparación de Enfoques para Obtener Estimaciones</b>	<b>34</b>
3.1. Características del Estudio	34
3.2. Análisis de Resultados	38
3.2.1. Cobertura	38
3.2.2. Cociente Entre las Longitudes de los Intervalos	51

3.2.3. Análisis del Extremo Derecho ..... 62

Conclusiones ..... 113

Bibliografía ..... 116

## Introducción

En muchas situaciones es de interés modelar los valores más grandes (o pequeños) que presenta un proceso estocástico. Por ejemplo, en la construcción de una presa interesa conocer el máximo nivel de agua que deberá soportar, o bien, la intensidad máxima que podría tener un terremoto en la zona donde estará ubicada la presa. En general, el conocimiento de la distribución de los máximos o mínimos de algún fenómeno relevante es importante dentro de la ingeniería: "El diseño en ingeniería debe estar basado en extremos porque los valores más grandes, tales como cargas, terremotos, viento, inundaciones, olas, etc., y los valores más pequeños como fuerza, fatiga, etc. son los parámetros claves que pueden llevar a una falla en los trabajos de ingeniería." (Castillo [2]: 4) Más ejemplos de situaciones en las que es de interés considerar eventos extremos pueden consultarse en Coles [1], pág. 1-17, y también Castillo [2], pág. 3-17.

Una forma particular de modelar eventos extremos ocurre cuando el proceso estocástico en cuestión está indexado por el tiempo, lo conforman variables aleatorias iid, y tenemos como "valores más grandes" a los máximos del proceso obtenidos en períodos cuya duración de tiempo es fija. En la Teoría de Valores Extremos, un análisis bajo ese contexto se conoce como de máximos por bloques. Sin embargo, existen diferentes enfoques que podemos aplicar para hacer inferencia en ese tipo de análisis. El objetivo principal de este documento será comparar dos de esos enfoques.

En el Capítulo 1 presentamos algunos resultados clásicos de la Teoría de Valores Extremos, con los cuales, encontraremos que la distribución de máximos se puede aproximar por algún tipo de las distribuciones de valores extremos (DVE) -Fréchet, Gumbel y Weibull-, las cuales podemos representar en una sola familia cuya función de distribución se conoce como distribución generalizada de valores extremos (DGVE). Por otro lado, en el Capítulo 2 discutimos sobre la selección de un modelo para ajustar la distribución de los máximos, y posteriormente, presentamos diferentes herramientas estadísticas basadas en la función de verosimilitud, las cuales podemos aplicar para obtener estimaciones. Las diferentes combina-

ciones entre tipos de modelo y herramientas estadísticas, dan lugar a que aparezcan diferentes enfoques inferenciales que podemos aplicar en la práctica. Dos de esos enfoques fueron comparados con un estudio basado en simulaciones. Las características de dicho estudio y los resultados obtenidos los presentaremos en el Capítulo 3.

## Capítulo 1

### Marco Teórico

En este capítulo presentamos algunos resultados clásicos de la Teoría de Valores Extremos. Un resultado fundamental es el Teorema de Tipos para Extremos, el cual, como veremos en la siguiente sección, puede usarse para sustentar métodos estadísticos que se aplican para estimar distribuciones asociadas a máximos. Además, también abordaremos el tema de los Dominios de Atracción y su caracterización, y aunque es difícil la aplicación de esos resultados en la práctica dado que parten del supuesto de que es conocida la distribución de interés, para nuestro trabajo es importante porque, al ser una exploración basada en simulaciones, sí contamos con esa información.

Los teoremas que aparecen en este capítulo son sólo enunciados pero las demostraciones están dadas por Resnick [7]. Al final de cada teorema agregamos el número de proposición o corolario que le corresponde en ese documento.

#### 1.1. Teorema de Tipos para Extremos

Sean  $X_1, X_2, \dots$ , una sucesión de variables aleatorias independientes idénticamente distribuidas (*iid*) con función de distribución común  $F$ . La teoría que presentamos a continuación se enfoca al análisis estadístico de máximos asociados a dicha sucesión, de manera específica, al análisis de

$$M_n = \max\{X_1, \dots, X_n\}, \quad n \geq 2.$$

La distribución teórica de  $M_n$  está dada por:

$$\begin{aligned} Pr\{M_n \leq z\} &= Pr\{X_1 \leq z, \dots, X_n \leq z\} \\ &= P(X_1 \leq z) \times \dots \times P(X_n \leq z) \\ &= F^n(z), \end{aligned} \quad (1.1)$$

donde  $z \in \mathbf{R}$  y  $n \in \mathbf{N}$ . Sin embargo, en la práctica no se conoce la distribución  $F$ , y aunque una posibilidad es obtener una estimación para  $F$  a partir de los datos, aparece el problema de que variaciones pequeñas de esa estimación, particularmente en la cola superior de  $F$ , podrían implicar diferencias sustanciales para  $F^n(z)$ .

Otro enfoque consiste en buscar directamente aproximaciones para la distribución de  $M_n$ . Un ejemplo de dicho enfoque se explica a continuación y se refiere a la aplicación del Teorema Central del Límite (TCL) para aproximar la distribución de la suma de variables aleatorias. Sean  $Y_1, Y_2, \dots$  una sucesión de variables aleatorias *iid* con función de distribución  $G$ . Si conocemos a  $G$ , entonces podemos obtener la distribución de  $\sum_{i=1}^n Y_i$  usando la convolución, sin embargo, en la práctica no tenemos esa información. Por otro lado, suponiendo que  $Y_1$  tiene varianza finita, entonces por el TCL se cumple que

$$\frac{\sum_{i=1}^n Y_i - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1), \quad (1.2)$$

donde  $E[Y_1] = \mu$ ,  $Var(Y_1) = \sigma^2$  y  $\xrightarrow{d}$  denota convergencia en distribución. Si se cuenta con estimaciones para  $\mu$  y  $\sigma$ , entonces la convergencia dada en (1.2) puede usarse para aproximar la distribución de la suma pues, con  $n$  "grande", se tiene que

$$\frac{\sum_{i=1}^n Y_i - n\mu}{\sigma\sqrt{n}} \approx N(0, 1).$$

Como analogía al caso anterior, trataremos de encontrar distribuciones límites para aproximar a  $F^n$ . Primero notemos que, denotando al extremo derecho de  $F$  por

$$x_F = \sup\{x \in \mathbf{R} : F(x) < 1\},$$

entonces se cumple que  $M_n \xrightarrow{d} x_F$ , es decir, la distribución de máximos converge a una distribución degenerada. La afirmación anterior se sigue de notar que para  $x < x_F$ , se tiene que  $F(x) < 1$  y

$$P(M_n \leq x) = F^n(x) \rightarrow 0, \quad n \rightarrow \infty,$$

por lo que  $M_n$  converge en probabilidad a  $x_F$ , y dado que la sucesión  $\{M_n\}$  es no decreciente, entonces converge casi seguramente, y en consecuencia, también en distribución.

El argumento anterior nos muestra que es necesario normalizar a  $M_n$ , es decir, estamos buscando garantizar una convergencia débil de la forma

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \rightarrow H(x),$$

con  $\{a_n\}$  y  $\{b_n\}$  sucesiones de números reales y  $H$  una distribución límite no degenerada. El Teorema de tipos para extremos, propuesto por Fisher y Tippett en 1928 y demostrado rigurosamente por Gnedenko en 1943, va en la dirección mencionada. Para enunciar ese teorema requerimos de la siguiente definición.

**Definición 1.1 (Familias de distribuciones).**

Se dice que dos distribuciones  $G$  y  $H$  pertenecen a la misma familia de distribuciones, o son del mismo tipo, si para algunas constantes  $a > 0$  y  $b \in \mathbf{R}$  se cumple

$$H(x) = G(ax + b), \quad x \in \mathbf{R}.$$

En términos de variables aleatorias, tomando  $X \sim G$  y  $Y \sim H$ , la condición anterior es equivalente a que  $Y$  y  $\frac{X-b}{a}$  tengan la misma distribución.

**Teorema 1.1 (Teorema de tipos para extremos)**

Supongamos que existen sucesiones de constantes  $a_n > 0$  y  $b_n \in \mathbf{R}$  tales que

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} H(x),$$

donde  $H$  es propia ( $H(\mathbf{R}) = 1$ ) y no está concentrada en un punto, entonces  $H$  pertenece a alguna de las siguientes familias de distribuciones.

$$\begin{aligned} \text{Fréchet} : \quad \Phi_\alpha(x) &= \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} & \alpha > 0 \\ \text{Weibull} : \quad \Psi_\alpha(x) &= \begin{cases} \exp(-(-x)^\alpha), & x < 0 \\ 1, & x \geq 0 \end{cases} & \alpha > 0 \\ \text{Gumbel} : \quad \Lambda(x) &= \exp\{-\exp\{-x\}\}, & x \in \mathbf{R} \end{aligned}$$

Una demostración puede consultarse en [7], Proposición 0.3.



**Definición 1.2 (Distribuciones de valores extremos)** Las funciones de distribución  $\Phi_\alpha$ ,  $\Psi_\alpha$  y  $\Lambda$ , como se presentaron en el Teorema 1.1, serán llamadas distribuciones de valores extremos estándar, mientras que las funciones de distribución que sean de alguno de esos tipos serán denotadas como distribuciones de valores extremos (DVE).

**Nota 1.1** Las familias de funciones de distribución para las DVE se obtienen al incluir parámetros de localización y escala a sus elementos estándar. Las expresiones correspondientes son:

$$\begin{aligned} \text{Fréchet : } \quad \Phi_{\alpha,a,b}(x) &= \begin{cases} 0, & x \leq b \\ \exp(-(\frac{x-b}{a})^{-\alpha}), & x > b \end{cases} \quad \alpha > 0; \\ \text{Weibull : } \quad \Psi_{\alpha,a,b}(x) &= \begin{cases} \exp(-(-(\frac{x-b}{a}))^\alpha), & x < b \\ 1, & x \geq b \end{cases} \quad \alpha > 0; \\ \text{Gumbel : } \quad \Lambda_{a,b}(x) &= \exp \left\{ -\exp \left\{ -\left( \frac{x-b}{a} \right) \right\} \right\}, \quad x \in \mathbf{R}, \end{aligned}$$

donde  $b \in \mathbf{R}$ ,  $a > 0$  y, para las distribuciones Fréchet y Weibull,  $\alpha > 0$ . El parámetro  $a$  es de escala, mientras que el parámetro  $b$  es de localización y determina, en los casos Fréchet y Gumbel, el soporte de la distribución. Finalmente,  $\alpha$  es un parámetro de forma.

Para demostrar el Teorema 1.1 podemos aplicar el teorema de convergencia a familias de Gnedenko y Khinchin, el cual enunciamos a continuación. Incluir ese resultado es de interés pues con él se garantiza que, dada una distribución de máximos, si existen diferentes sucesiones de constantes normalizadoras con las cuales se puede obtener una distribución límite, entonces esas distribuciones límite son del mismo tipo.

**Teorema 1.2 (Teorema de convergencia a familias)**

Sean  $G(x)$  y  $H(x)$  dos distribuciones ninguna de las cuales está concentrada en un punto. Supongamos que para  $n \geq 1$   $F_n$  es una distribución,  $a_n \geq 0$ ,  $b_n \in \mathbf{R}$ ,  $\alpha_n \geq 0$ ,  $\beta_n \in \mathbf{R}$  y

$$F_n(a_n x + b_n) \rightarrow G(x), \quad F_n(\alpha_n x + \beta_n) \rightarrow H(x) \quad (1.3)$$

débilmente. Entonces existen constantes  $A > 0$  y  $B \in \mathbf{R}$  tales que, cuando  $n \rightarrow \infty$ ,

$$\frac{\alpha_n}{a_n} \rightarrow A, \quad \frac{\beta_n - b_n}{a_n} \rightarrow B \quad (1.4)$$

y

$$H(x) = G(Ax + B). \quad (1.5)$$

Por otro lado, si 1.4 es cierta, entonces 1.3 implica 1.5.

Una demostración de este teorema aparece en [7], Proposición 0.2.

## 1.2. Dominios de Atracción

El teorema enunciado en la sección anterior identifica las distribuciones límite para máximos normalizados. A continuación revisaremos las condiciones que debe cumplir una función de distribución  $F$  para que, cuando existe una distribución límite, ésta sea de un tipo específico de las DVE.

**Definición 1.3 (Dominio de atracción)**

Una variable aleatoria  $X$  (o su función de distribución  $F$ ) pertenece al dominio de atracción de la distribución de valores extremos  $H$  si existen constantes  $a_n > 0$ ,  $b_n \in \mathbf{R}$  tales que

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} H(x).$$

En ese caso se usará la notación  $X \in DA(H)$  ( $F \in DA(H)$ ).

Con la definición anterior podemos reformular el objetivo de esta sección: caracterizar el dominio de atracción para cada tipo de DVE.

### 1.2.1. Dominio de Atracción de las DVE Tipo Fréchet ( $\Phi_\alpha$ )

Comenzamos definiendo un tipo de funciones llamadas de variación regular y la función cuantil asociada a una función de distribución.

**Definición 1.4 (Funciones de variación regular)**

Una función medible  $U : \mathbf{R}^+ \rightarrow \mathbf{R}^+$  se dice que es de variación regular en  $\infty$  con índice  $\alpha \in \mathbf{R}$  ( $U \in VR_\alpha$ ), si para  $x > 0$

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\alpha.$$

**Definición 1.5 (Función cuantil)**

Sea  $F$  una función de distribución, su inversa generalizada

$$Q(t) = F^{-}(t) = \inf\{x \in \mathbf{R} : F(x) \geq t\}, \quad 0 < t < 1,$$

es llamada la función cuantil de la función de distribución  $F$ . El cuantil  $t$  de  $F$  se define como  $x_t = F^{-}(t)$ .

Con las definiciones anteriores y usando la notación  $\bar{F} = 1 - F$ , podemos enunciar un teorema que caracteriza el dominio de atracción Fréchet ( $DA(\Phi_\alpha)$ ).

**Teorema 1.3 (Caracterización del  $DA(\Phi_\alpha)$ )**

Sea  $F$  una función de distribución y  $\alpha > 0$ , entonces  $F \in DA(\Phi_\alpha)$  si y sólo si  $\bar{F} \in VR_{-\alpha}$ . Además, si  $F \in DA(\Phi_\alpha)$ , entonces

$$\frac{M_n}{a_n} \xrightarrow{d} \Phi_\alpha, \quad (1.6)$$

con

$$a_n = Q\left(1 - \frac{1}{n}\right) = \left(\frac{1}{\bar{F}}\right)^{-}(n).$$

Una prueba aparece en [7], Proposición 1.11.

**Nota 1.2** Para que  $F \in DA(\Phi_\alpha)$  es condición necesaria que  $x_F = \infty$ . Lo anterior se obtiene notando que (1.6) es equivalente a que

$$F^n(a_n x) \rightarrow \Phi_\alpha(x)$$

débilmente, y por lo tanto,  $a_n$  debe ser no acotada.

El siguiente corolario presenta condiciones suficientes para que  $F \in DA(\Phi_\alpha)$ . El interés de contar con ese tipo de condiciones es que podrían ser más fáciles de verificar.

**Corolario 1.1 (Condición de Von Mises)**

Sea  $F$  una función de distribución absolutamente continua con función de densidad  $f$  que satisface

$$\lim_{x \rightarrow \infty} \frac{x f(x)}{\bar{F}(x)} = \alpha > 0,$$

entonces  $F \in DA(\Phi_\alpha)$

Para indicaciones sobre la demostración ver [7], Proposición 1.15.

**1.2.2. Dominio de Atracción de las DVE Tipo Weibull ( $\Psi_\alpha$ )**

La caracterización del dominio de atracción de  $\Psi_\alpha$ , al igual que en el caso Fréchet, está relacionada con funciones de variación regular. Esa similitud es congruente con la estrecha relación que existe entre esas DVE las cuales satisfacen la relación:

$$\Psi_\alpha(-x^{-1}) = \Phi_\alpha(x), \quad x > 0.$$

**Teorema 1.4 (Caracterización del  $DA(\Psi_\alpha)$ )**

Sea  $F$  una función de distribución y  $\alpha > 0$ , entonces  $F \in DA(\Psi_\alpha)$  si y sólo si  $x_F < \infty$  y  $\bar{F}(x_F - x^{-1}) \in VR_\alpha$ . Además, si  $F \in DA(\Psi_\alpha)$ , entonces

$$\frac{M_n - x_F}{a_n} \xrightarrow{d} \Psi(\alpha), \quad (1.7)$$

con

$$a_n = x_F - Q\left(1 - \frac{1}{n}\right) = x_F - \left(\frac{1}{\bar{F}}\right)^{-}(n).$$

Una demostración de este teorema se puede encontrar en [7], Proposición 1.13.

Como en el caso Fréchet, si tenemos una función de densidad, entonces podemos encontrar condiciones suficientes para que  $F \in DA(\Psi_\alpha)$ .

**Corolario 1.2 (Condición de Von Mises)**

Sea  $F$  una función de distribución absolutamente continua con función de densidad  $f$  que es positiva en algún intervalo finito  $(z, x_F)$ . Si

$$\lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{\bar{F}(x)} = \alpha > 0,$$

entonces  $F \in DA(\Psi_\alpha)$ .

Para una demostración consultar [7], Proposición 1.16.

**1.2.3. Dominio de Atracción de las DVE Tipo Gumbel ( $\Lambda$ )**

Comenzamos presentando un tipo de funciones llamadas de Von Mises que, como veremos, nos permitirán presentar condiciones suficientes para que una función de distribución pertenezca al  $DA(\Lambda)$ . Posteriormente, enunciaremos algunos teoremas para caracterizar totalmente a ese dominio de atracción.

**Definición 1.6 (Funciones de Von Mises)**

Sea  $F$  una función de distribución con extremo derecho  $x_F \leq \infty$ . Supongamos que existe  $z < x_F$  tal que  $F$  tiene representación

$$\bar{F}(x) = c \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\}, \quad z < x < x_F,$$

donde  $c$  es alguna constante positiva,  $a(\cdot)$  es una función positiva y absolutamente continua (respecto a la medida de Lebesgue) con densidad  $a'$  y  $\lim_{x \uparrow x_F} a'(x) = 0$ . Entonces decimos que  $F$  es una función de Von Mises y que  $a(\cdot)$  es la función auxiliar de  $F$ .

El siguiente teorema afirma que una función de Von Mises pertenece al dominio de atracción de la distribución Gumbel. Además, también presenta condiciones suficientes para que una función de distribución sea de Von Mises.

**Teorema 1.5 (Funciones de Von Mises y  $DA(\Lambda)$ )**

a) Sea  $F$  una función de distribución. Si  $F$  es una función de Von Mises, entonces  $F \in DA(\Lambda)$ . En ese caso se tiene que

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} \Lambda, \quad (1.8)$$

con  $b_n = F^{-1}(1 - \frac{1}{n})$  y  $a_n = a(b_n)$ , donde  $a$  es la función auxiliar de  $F$ .

b) Supongamos que  $F$  tiene extremo derecho  $x_F \leq \infty$  y que existe  $z < x_F$  tal que  $F$  es dos veces diferenciable en  $(z, x_F)$  con densidad positiva  $f = F'$  y  $F''(x) < 0$  para  $z < x < x_F$ . Si

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)F''(x)}{f^2(x)} = -1, \quad (1.9)$$

entonces  $F$  es una función de Von Mises y  $F \in DA(\Lambda)$ . Podemos tomar como función auxiliar

$$a = \frac{\bar{F}}{f}.$$

Inversamente, una función de Von Mises dos veces diferenciable satisface (1.9).

Para una demostración consultar [7], Proposición 1.1.

Como en los otros dominios de atracción, en el caso Gumbel también se cuenta con caracterizaciones completas de su dominio de atracción.

**Teorema 1.6 (Caracterización del  $DA(\Lambda)$ )**

Sea  $F$  una función de distribución con extremo derecho  $x_F \leq \infty$ .  $F \in DA(\Lambda)$  si y sólo si existe algún  $z < x_F$  tal que  $F$  tiene representación

$$\bar{F}(x) = c(x) \exp \left\{ - \int_z^x \frac{g(t)}{a(t)} dt \right\}, \quad z < x < x_F, \quad (1.10)$$

donde  $c$  y  $g$  son funciones medibles que satisfacen  $c(x) \rightarrow c > 0$ ,  $g(x) \rightarrow 1$  cuando  $x \uparrow x_F$ , y  $a(x)$  es una función positiva y absolutamente continua respecto a la medida de Lebesgue con densidad  $a'(x)$  que satisface  $\lim_{x \uparrow x_F} a'(x) = 0$ . Para  $F$  con representación (1.10) podemos escoger

$$b_n = F^{-1}(1 - \frac{1}{n}) \quad \text{y} \quad a_n = a(b_n)$$

como constantes normalizadoras. Una elección posible para la función  $a$  es

$$a(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt, \quad x < x_F.$$

Una demostración aparece en [7], Corolario 1.7 y Proposición 1.9.

**Nota 1.3** Un procedimiento útil para determinar el dominio de atracción y las constantes normalizadoras es a través de transformaciones monótonas. Sea  $X_1, X_2, \dots$ , una sucesión de variables aleatorias y  $g$  una función creciente. Definiendo  $\tilde{X} = g(X)$  se tiene que

$$\tilde{M}_n = \max(\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n) = g(M_n).$$

Si  $X \in DA(\Lambda)$  con

$$\lim_{n \rightarrow \infty} P(M_n \leq a_n x + b_n) = \Lambda(x), \quad x \in \mathbf{R},$$

entonces

$$\lim_{n \rightarrow \infty} P(\tilde{M}_n \leq g(a_n x + b_n)) = \Lambda(x), \quad x \in \mathbf{R},$$

En algunos casos, es posible calcular la expansión de Taylor para  $g$  alrededor de  $b_n$ , y los términos lineales son suficientes para obtener la distribución límite de  $\tilde{M}_n$  considerando como constantes normalizadoras a  $\tilde{a}_n = a_n g'(b_n)$  y  $\tilde{b}_n = g(b_n)$ .

### 1.3. Distribución Generalizada de Valores Extremos

Las familias de distribuciones Fréchet, Gumbel y Weibull pueden ser combinadas dentro de una familia de modelos que se presenta en la siguiente definición.

**Definición 1.7** (*Distribución generalizada de valores extremos*)

Consideremos la familia de distribuciones con parámetros  $\mu \in \mathbf{R}$ ,  $\sigma > 0$  y  $\xi \in \mathbf{R}$  dada por

$$G(z) = \begin{cases} \exp -[1 + \xi(\frac{z-\mu}{\sigma})]^{-\frac{1}{\xi}}, & \xi \neq 0; \\ \exp[-\exp -(\frac{z-\mu}{\sigma})], & \xi = 0, \end{cases} \quad (1.11)$$

definida en

$$\begin{cases} \{z : 1 + \xi(\frac{z-\mu}{\sigma}) > 0\}, & \xi \neq 0; \\ -\infty < z < \infty, & \xi = 0. \end{cases}$$

Ésta es la familia de distribuciones de valores extremos generalizada y nos referiremos a cualquiera de sus elementos como una distribución generalizada de valores extremos (*DGVE*).

Una *DGVE* tiene tres parámetros: un parámetro de localización,  $\mu$ ; un parámetro de escala,  $\sigma$ ; y un parámetros de forma,  $\xi$ . Además, para  $\xi \neq 0$  el soporte está determinado por una combinación de dichos parámetros, en particular, para  $\xi > 0$  la distribución tiene un extremo izquierdo finito, mientras que para  $\xi < 0$ , el extremo derecho es finito.

Sea  $G$  una *DGVE* con parámetros  $(\mu, \sigma, \xi)$ . Si  $\xi > 0$ , entonces  $G$  es de tipo Fréchet, para  $\xi < 0$  es de tipo Weibull, y finalmente, si  $\xi = 0$ , entonces es de tipo Gumbel. Además, es posible determinar los parámetros  $(a, b, \alpha)$  que le corresponden a  $G$  cuando se escribe como una *DVE*, ésta última parametrizada como aparece en la nota 1.1.

$$\text{Cuando } \xi > 0: \begin{cases} b = \mu - \frac{\sigma}{\xi}; \\ a = \frac{\sigma}{\xi}; \\ \alpha = \frac{1}{\xi}. \end{cases}$$

$$\text{Cuando } \xi < 0: \begin{cases} b = \mu - \frac{\sigma}{\xi}; \\ a = -\frac{\sigma}{\xi}; \\ \alpha = -\frac{1}{\xi}. \end{cases}$$

Y, finalmente, para  $\xi = 0$  se tiene  $b = \mu$  y  $a = \sigma$ .

En consecuencia, además de identificar con el signo de  $\xi$  el tipo de *DVE* a la que corresponde una *DGVE*, también podemos determinar el elemento estándar con el que está

relacionada.

$$\begin{cases} \xi = \alpha^{-1} > 0, & \text{corresponde a una distribución de tipo Fréchet } \Phi_\alpha, \\ \xi = 0, & \text{corresponde a una distribución de tipo Gumbel } \Lambda, \\ \xi = -\alpha^{-1} < 0, & \text{corresponde a una distribución de tipo Weibull } \Psi_\alpha. \end{cases} \quad (1.12)$$

### 1.4. Dominios de Atracción para Algunas Distribuciones Conocidas

En los Cuadros 1.1, 1.2 y 1.3, aparecen clasificadas algunas familias de distribuciones según el tipo de dominio de atracción al que pertenecen. Para los tipos Fréchet y Weibull es posible determinar el parámetro  $\alpha$  que determina a  $\Phi_\alpha$  y  $\Psi_\alpha$ , sin embargo, se usó la transformación (1.12) y presentamos el valor para  $\xi$  correspondiente. La utilidad de usar ese cambio en la parametrización de las *DVE* se debe a que, en los capítulos siguientes, serán comparadas con las *DGVE*, y en consecuencia, será cómodo tener la misma parametrización para el parámetro de forma. Finalmente hacemos algunos comentarios sobre la notación que usamos. Dentro de las distribuciones consideradas, la que denotamos como "Weibull<sub>m</sub>" hace referencia a la distribución Weibull de dos parámetros para mínimos -la notación que usamos en este caso coincide con la de Castillo [2]-. Y por otro lado, la distribución Weibull que corresponde a una *DVE* la hemos denotado como "Weibull (*DVE*)".

Demostraciones de que las funciones de distribución que aparecen en los Cuadros 1.1, 1.2 y 1.3 pertenecen al dominio de atracción señalado, están dadas por Embrechts, Klüpelberg y Mikosch [3]. Además, información general sobre dichas distribuciones puede encontrarse en Hogg y Klugman [4], o bien, para un análisis más detallado, se puede consultar Johnson y Kotz [5] y [6].

### 1.5. Niveles y Períodos de Retorno

Dentro del análisis de eventos extremos aparece la siguiente pregunta: ¿Cuál es el tiempo de espera promedio para que ocurra un evento extremo específico?

Sea  $X_1, X_2, \dots$ , una sucesión de variables aleatorias *iid* con función de distribución continua  $F$  y  $u \in \mathbf{R}$  un umbral dado. Asumamos que  $\{X_i > u\}$  denota a un evento extremo y definamos al tiempo de la primera excedencia como

$$T(u) = \min\{i \geq 1 : X_i > u\}.$$

Name	Distribución $F$ o densidad $f$	Parámetros	Soporte	$\xi$
Burr	$F(x) = 1 - \left(\frac{\kappa}{\kappa+x^\tau}\right)^\alpha$	$\alpha, \kappa, \tau > 0$	$x > 0$	$\frac{1}{\alpha\tau}$
Pareto	$F(x) = 1 - \left(\frac{\kappa}{\kappa+x}\right)^\alpha$	$\alpha, \kappa > 0$	$x > 0$	$\frac{1}{\alpha}$
Loggamma	$f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} (\ln(x))^{\beta-1} x^{-\alpha-1}$	$\alpha, \beta > 0$	$1 < x < \infty$	$\frac{1}{\alpha}$
Fréchet	$F(x) = \exp\{-x^{-\alpha}\}$	$\alpha > 0$	$x > 0$	$\frac{1}{\alpha}$
Cauchy	$f(x) = \left[\pi a \left(1 + \left(\frac{x-b}{a}\right)^2\right)\right]^{-1}$	$b \in R, a > 0$	$-\infty < x < \infty$	1

Cuadro 1.1: Distribuciones que pertenecen al dominio de atracción Fréchet

Name	Distribución $F$ o densidad $f$	Parámetros	Soporte
Weibull <sub>m</sub>	$F(x) = 1 - \exp\{-cx^\tau\}$	$c, \tau > 0$	$x > 0$
Lognormal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}$	$\mu \in R, \sigma > 0$	$x > 0$
Exponencial	$F(x) = 1 - \exp\{-\lambda x\}$	$\lambda > 0$	$x > 0$
Gamma	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\}$	$\alpha, \beta > 0$	$x > 0$
Logística	$F(x) = 1 - \left[1 + \exp\left\{\frac{x-b}{a}\right\}\right]^{-1}$	$b \in R, a > 0$	$-\infty < x < \infty$
Normal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$	$\mu \in R, \sigma > 0$	$-\infty < x < \infty$

Cuadro 1.2: Distribuciones que pertenecen al dominio de atracción Gumbel

Name	Distribución $F$ o densidad $f$	Parámetros	Soporte	$\xi$
Beta	$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1}$	$\alpha, \beta > 0$	$0 < x < 1$	$-\frac{1}{\beta}$
Weibull (DVE)	$F(x) = \exp\{-(-x)^\alpha\}$	$\alpha > 0$	$x > 0$	$-\frac{1}{\alpha}$
Uniforme	$F(x) = \frac{x-a}{b-a}$	$a < b$	$a \leq x \leq b$	-1

Cuadro 1.3: Distribuciones que pertenecen al dominio de atracción Weibull

Con el modelo anterior, la respuesta a la pregunta planteada es  $E[T(u)]$ . Además, tomando  $p = 1 - F(u)$ , obtenemos

$$P(T(u) = k) = (1-p)^{k-1}p, \quad k = 1, 2, \dots,$$

por lo tanto,  $T(u)$  es una variable aleatoria geométrica y  $E[T(u)] = p^{-1}$ . Definimos a  $u$  como el nivel de retorno con un período de retorno  $p^{-1}$  para los eventos  $\{X_i > u\}$ . Nótese que  $u$  es el cuantil de  $F$  con un nivel de probabilidad  $1-p$ .

Aunque en el modelo anterior partimos de un umbral que definía un evento extremo al cual se le calculó su tiempo promedio de ocurrencia, también podemos proceder de manera inversa. Supongamos que buscamos los eventos extremos del tipo  $\{X_i > u\}$  que tardan en ocurrir en promedio un tiempo igual a  $p^{-1}$ , entonces basta tomar como nivel de retorno a  $u = x_p$ , donde  $F(x_p) = 1-p$ .

## Capítulo 2

# Modelación Estadística para Eventos Extremos

Sean  $X_1, X_2, \dots$ , una sucesión de variables aleatorias *iid* cuya distribución común no es conocida. Definimos

$$M_n = \max\{X_1, \dots, X_n\}, \quad n \geq 2.$$

El problema de interés es determinar el comportamiento estadístico de  $M_n$ .

El Teorema 1.1 afirma que las únicas distribuciones límite posibles para  $M_n$  son las *DVE*. Es posible adoptar un enfoque basado en el uso de esas distribuciones límite como aproximaciones a un nivel finito. Específicamente, supongamos que existen sucesiones de constantes  $\{a_n > 0\}$  y  $\{b_n\}$  de modo que

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) \rightarrow G(x)$$

débilmente, con  $G$  una *DVE*. Para  $n$  suficientemente grande

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) \approx G(x),$$

o equivalentemente, con  $z = a_n x + b_n$ ,

$$\begin{aligned} P(M_n \leq z) &\approx G\left(\frac{z - b_n}{a_n}\right) \\ &= G_n^*(z), \end{aligned} \quad (2.1)$$

donde  $G_n^*$  es otra *DVE* del mismo tipo que  $G$ , pero con diferentes parámetros de localización y escala. Por lo tanto, se ha encontrado que la distribución de  $M_n$  se puede aproximar por una

*DVE*. Además, dado que los parámetros de la distribución serán estimados, es irrelevante si los parámetros de la distribución  $G$  son diferentes de los de  $G_n^*$ , y por lo tanto, no es un problema que sean desconocidas las constantes de normalización.

El argumento presentado nos conduce al siguiente enfoque para modelar extremos en la práctica. Sea  $X_1, X_2, \dots$ , una sucesión de variables aleatorias *iid* con distribución común  $F$  y supongamos que tenemos una muestra  $x_1, x_2, \dots, x_m$ . Dividimos la muestra en bloques de  $n$  observaciones formando los vectores

$$\begin{aligned} x^{(1)} &= (x_1^1, \dots, x_n^1), \\ x^{(2)} &= (x_1^2, \dots, x_n^2), \\ &\vdots \\ x^{(k)} &= (x_1^k, \dots, x_n^k). \end{aligned}$$

La muestra *iid* con respecto a la que se realizarán inferencias consiste de

$$M_{n,i} = \max\{x_1^i, \dots, x_n^i\}, \quad i = 1, \dots, k.$$

El enfoque descrito se conoce en la literatura como análisis de máximos por bloques. En ciertas áreas, como en el análisis de datos ambientales, es usual que los bloques se escojan para corresponder con un período de tiempo igual a un año. En ese caso,  $n$  es el número de observaciones consideradas en un año y los máximos por bloques son máximos anuales.

Una vez que contamos con la muestra de máximos, podemos aplicar algún enfoque para obtener inferencias. Para ello, es necesario primero determinar una familia de modelos para aproximar la distribución de los máximos. Esa discusión aparece en la Sección 2.1. Después, debemos considerar el tipo de herramientas estadísticas que usaremos, en este trabajo consideramos las basadas en la función de verosimilitud. La presentación de dichas herramientas aparece en la Sección 2.2. Finalmente, las combinaciones entre herramientas estadísticas y la selección de un modelo, determina diferentes enfoques posibles para hacer inferencia. En la sección 2.3 presentamos tres de ellos.

### 2.1. Selección de Modelo

Consideremos una muestra de máximos por bloques. Ya mostramos que es razonable aproximar la distribución de dicha muestra usando las *DVE*. Sin embargo, existen al menos dos tipos de modelos que podemos usar como base para obtener inferencias, por un lado,

algún tipo de las DVE, y por el otro, la *DGVE* que, como vimos en el Capítulo 1, incluye a las anteriores.

Utilizar las *DVE* como modelo presupone que se tiene algún procedimiento para escoger entre las familias Fréchet, Gumbel y Weibull. Cuando hacemos esa selección, las inferencias siguientes asumen que fue correcta, y por lo tanto, podrían no incluir la incertidumbre de la misma. La selección de una familia es un aspecto relevante si se considera que las colas de las DVE presentan un comportamiento diferenciado. Comparando sus extremos derechos, en las distribuciones Fréchet y Gumbel es infinito, mientras que para las distribuciones de tipo Weibull es finito. Además, la función de densidad decae exponencialmente para la distribución Gumbel y como un polinomio para la distribución Fréchet, lo cual implica diferentes tasas de decrecimiento en las colas de esas distribuciones.

Otro modelo que podemos escoger es el de las *DGVE*. Coles [1] menciona que en este caso se tiene diversas ventajas respecto al caso anterior. "La unificación de las tres familias originales de distribuciones de valores extremos en una familia simplifica enormemente la implementación estadística. A través de la inferencia para  $\xi$ , los datos mismos determinan el tipo más apropiado de comportamiento en la cola, y no hay necesidad de hacer juicios subjetivos a priori sobre cuál familia de valores extremos individual adoptar." (Coles [1]: 48)

Otro factor a considerar en la elección del modelo se refiere a la interpretación de los parámetros involucrados. En las *DVE*, los parámetros tienen una interpretación física clara pues, en particular, el parámetro de localización para los casos Fréchet y Weibull determina el soporte de la distribución. En cambio, en la *DGVE* el soporte es resultado de una combinación de los tres parámetros, situación que hace más complicada la inferencia. Además, cuando la *DVE* de tipo Gumbel es un modelo adecuado, entonces el número de parámetros se reduce a 2.

## 2.2. Herramientas Basadas en la Verosimilitud

En esta sección explicamos diferentes conceptos y resultados que son usados cuando se obtienen inferencias basadas en la función de verosimilitud. Aunque se incluyen algunos comentarios para el caso en que estemos trabajando con las distribuciones de valores extremos, será hasta la siguiente sección donde explicaremos, de manera explícita, algunos enfoques para hacer inferencia cuando se analizan máximos por bloques. Dicha discusión se deja para una sección posterior pues, como veremos, esos enfoques tienen características particulares no sólo en términos de las herramientas que aplican para hacer inferencia -las cuales son

explicadas en esta sección-, sino también en lo que se refiere a los modelos que usan.

### 2.2.1. Estimadores de Máxima Verosimilitud

Sean  $x_1, \dots, x_n$  una muestra de variables aleatorias *iid* cuya distribución común  $F_\theta$  es desconocida y se asume que pertenece a una familia paramétrica

$$\mathcal{F} = \{F_\theta : \theta \in \Theta\},$$

donde  $\Theta \subset \mathbb{R}^k$ , para algún  $k \in \mathbb{N}$ . Supongamos que las distribuciones  $F_\theta$  tienen densidades  $f(x; \theta)$ . Para dar una estimación puntual del valor de  $\theta$  podemos usar el estimador de máxima verosimilitud (*emv*) que definimos a continuación.

Cada valor de  $\theta \in \Theta$  define un modelo potencial, y en consecuencia, determina una función de densidad conjunta posible de la cual  $x_1, \dots, x_n$  es una realización. Asociando a cada vector  $\theta$  con el número real que se obtiene al evaluar la densidad conjunta correspondiente en  $(x_1, \dots, x_n)$ , se define una relación que se denota como **función de verosimilitud**. Por la independencia de las realizaciones, se sigue que esa función tiene la forma

$$L(\theta; \mathbf{x}) = \prod_{i=1}^n f(x_i; \theta),$$

donde  $\mathbf{x} = (x_1, \dots, x_n)$ . Valores de  $\theta$  que tienen una verosimilitud alta corresponden a modelos que le dan alta plausibilidad a los datos observados. El principio de estimación puntual por máxima verosimilitud consiste en adoptar el modelo con la verosimilitud más alta, por lo tanto, el **estimador de máxima verosimilitud** para  $\theta$  es cualquier valor  $\hat{\theta}$  que maximiza  $L$  en  $\Theta$ . Muchas veces la estimación  $\hat{\theta}$  puede obtenerse resolviendo el sistema de ecuaciones de verosimilitud,

$$\left. \frac{\partial \log L}{\partial \theta_i} \right|_{\theta=\hat{\theta}} = 0, \quad (i = 1, \dots, k),$$

y confirmando que la solución  $\hat{\theta}$  maximiza  $L$ . Por otro lado, también es usual tomar logaritmos y trabajar con la **función de log-verosimilitud**

$$\ell(\theta) = \sum_{i=1}^n \log f(x_i; \theta).$$

Dado que la función logaritmo es monótona es equivalente maximizar  $\ell$  o  $L$  para obtener  $\hat{\theta}$ .

Bajo ciertas condiciones de regularidad, los estimadores de máxima verosimilitud tienen algunas propiedades asintóticas como se muestra en el siguiente teorema. En lo que sigue incluiremos en la notación de los estimadores de máxima verosimilitud un subíndice  $n$  el cual hace referencia al tamaño de la muestra usada.

**Teorema 2.1** Sean  $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$  una familia de funciones de distribuciones, donde  $\Theta$  es un subconjunto abierto de  $\mathbf{R}$ . Supongamos que las distribuciones  $F_\theta$  poseen funciones de densidad  $f(x; \theta)$  y que la familia  $\mathcal{F}$  satisface las siguientes condiciones de regularidad.

1. Para cada  $\theta \in \Theta$ , las derivadas

$$\frac{\partial \log [f(x, \theta)]}{\partial \theta}, \frac{\partial^2 \log [f(x, \theta)]}{\partial \theta^2}, \frac{\partial \log [f(x, \theta)]}{\partial \theta}$$

existen para toda  $x$ .

2. Para cada  $\theta_0 \in \Theta$ , existen funciones  $g(x)$ ,  $h(x)$ , y  $m(x)$  (que posiblemente dependen de  $\theta_0$ ) tales que para  $\theta$  en una vecindad  $B(\theta_0)$  las relaciones

$$\left| \frac{\partial \log [f(x, \theta)]}{\partial \theta} \right| \leq g(x), \quad \left| \frac{\partial^2 \log [f(x, \theta)]}{\partial \theta^2} \right| \leq g(x), \quad \left| \frac{\partial^2 \log [f(x, \theta)]}{\partial \theta^2} \right| \leq g(x)$$

son ciertas para toda  $x$ , y

$$\int_{\mathbf{R}} g(x) dx < \infty, \quad \int_{\mathbf{R}} h(x) dx < \infty, \quad \int_{\mathbf{R}} m(x) f(x; \theta) dx < \infty$$

para  $\theta \in B(\theta_0)$ .

3. Para cada  $\theta \in \Theta$ ,

$$0 < E_\theta \left\{ \left( \frac{\partial \log [f(x; \theta)]}{\partial \theta} \right)^2 \right\} < \infty.$$

Bajo las condiciones anteriores, consideremos observaciones iid de  $F_\theta$ , para  $\theta$  un elemento de  $\Theta$ . Entonces, con probabilidad 1, las ecuaciones de verosimilitud admiten una sucesión de soluciones  $\{\hat{\theta}_n\}$  tales que

$$(\hat{\theta}_n - \theta) \sqrt{n \mathcal{I}_\theta} \xrightarrow{d} N(0, 1),$$

cuando  $n \rightarrow \infty$ , donde  $\mathcal{I}_\theta = -E_\theta(\partial^2 \log f(\theta; x_i) / \partial \theta^2)$  es la información esperada de una sola observación, y  $\mathcal{I}_\theta = n \mathcal{I}_\theta$  es la función de información esperada de Fisher.

Además, con apropiadas generalizaciones de las condiciones anteriores, el resultado sigue siendo válido para el caso multidimensional. Si suponemos  $\theta = (\theta_1, \dots, \theta_k)$ , entonces, con

probabilidad 1, las ecuaciones de verosimilitud admiten una sucesión de soluciones  $\{\hat{\theta}_n\}$  tales que

$$(\hat{\theta}_n - \theta) \sqrt{n} \xrightarrow{d} N_k(0, \mathcal{I}(\theta)^{-1}),$$

cuando  $n \rightarrow \infty$ , donde  $\mathcal{I}_\theta = -E_\theta \{ \partial^2 \log f(\theta; x_i) / \partial \theta_j \partial \theta_m \}_{k \times k}$  es la matriz de información esperada para una sola observación y  $\mathcal{I}_\theta = n \mathcal{I}_\theta$  es la matriz de información esperada de Fisher.

Una demostración puede consultarse en [9], Sección 4.4.2, pág. 152.

Aunque una familia paramétrica  $F$  pueda ser indexada por un parámetro  $\theta$ , puede ocurrir que el interés es estimar alguna función  $\psi = h(\theta)$ , donde  $\psi$  puede tener una dimensión distinta a  $\theta$ . Los siguientes dos resultados permiten transformar las inferencias para  $\theta$  y obtener las correspondientes para  $\psi$ .

**Teorema 2.2** Si  $\hat{\theta}$  es el estimador de máxima verosimilitud para  $\theta$ , y  $\psi = h(\theta)$  es una función escalar, entonces el estimador de máxima verosimilitud de  $\psi$  está dado por  $\hat{\psi} = h(\hat{\theta})$ .

**Teorema 2.3** Supongamos que las condiciones de regularidad del Teorema 2.1 se satisfacen. En ese caso, para una muestra "grande" obtenemos que la distribución del estimador de máxima verosimilitud  $\hat{\theta}$  es aproximadamente normal, en particular, denotemos a su matriz de covarianza aproximada como  $V_\theta$ . Si  $\psi = h(\theta)$  es una función escalar, entonces el estimador de máxima verosimilitud  $\hat{\psi} = h(\hat{\theta})$  tiene una distribución aproximada  $N(\psi, V_\psi)$ , donde

$$V_\psi = \nabla \psi^t V_\theta \nabla \psi$$

con

$$\nabla \psi = \left[ \frac{\delta \psi}{\delta \theta_1}, \dots, \frac{\delta \psi}{\delta \theta_k} \right]$$

evaluado en  $\hat{\theta}$ .

Indicaciones de la demostración aparecen en Shao [8], pág. 251.

## 2.2.2. Intervalos de Verosimilitud y Aproximación $\chi^2$

Para comparar la plausibilidad asociada a los diferentes valores de  $\theta$  podemos usar la función de verosimilitud relativa que se define como

$$R(\theta) = \frac{L(\theta)}{L(\hat{\theta}_n)}.$$



Al conjunto  $\{\theta : R(\theta) \geq l\}$ , con  $l \in (0, 1)$ , lo llamaremos **region de verosimilitud**, mientras que a  $l$  lo denotaremos como **nivel de verosimilitud**. Para el caso en que la dimensión de  $\Theta$  es igual a 1 esas regiones reciben el nombre de **intervalos de verosimilitud**.

Las afirmaciones que hagamos usando regiones de verosimilitud se refieren sólo a la plausibilidad relativa para los valores de  $\theta$ , y no a la incertidumbre del intervalo. Aunque, en algunos casos, es posible añadir a las regiones de verosimilitud la probabilidad de que contengan el parámetro verdadero. Una forma de hacerlo es usando la función

$$D_n(\theta) = 2(\ell(\hat{\theta}_n) - \ell(\theta)) = -2\log(R(\theta)),$$

porque, como se muestra en el siguiente teorema, bajo ciertas condiciones de regularidad es posible determinar su distribución asintótica.

**Teorema 2.4** Sean  $x_1, \dots, x_n$  una muestra de variables aleatorias iid con función de distribución  $F_\theta$  que pertenece a una familia  $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$ , donde  $\Theta$  es un subconjunto abierto de  $\mathbf{R}^k$ . Supongamos que se satisfacen las condiciones de regularidad del Teorema 2.1. Entonces

$$D_n(\theta) \xrightarrow{d} \chi_k^2, n \rightarrow \infty.$$

Una demostración aparece en Serfling [9], Sección 4.4.3, pág. 155.

Del Teorema 2.4 se sigue que una región de confianza con probabilidad aproximada de  $(1 - \alpha)$  está dada por

$$C_\alpha = \{\theta : D(\theta) \leq c_\alpha\}, \quad (2.2)$$

donde  $c_\alpha$  es el cuantil  $(1 - \alpha)$  de la distribución  $\chi_d^2$ . (ver, por ejemplo, Shao [8] pág. 447) Escribiendo (2.2) en términos de la verosimilitud relativa, obtenemos

$$C_\alpha = \{\theta : R(\theta) \geq \exp(-\frac{1}{2}c_\alpha)\},$$

y por lo tanto, cuando es posible aplicar el Teorema 2.4 se puede aplicar, se cumple que una región de verosimilitud con nivel  $\exp(-\frac{1}{2}c_\alpha)$  tiene una probabilidad de cobertura aproximada de  $(1 - \alpha)$ . En el Cuadro 2.1 se presentan algunos ejemplos de esa relación para el caso  $d = 1$ .

Nivel de Verosimilitud		Nivel de Confianza
0.0362	-	0.99
0.1465	-	0.95
0.2585	-	0.90
0.7965	-	0.50

Cuadro 2.1: Correspondencia entre niveles de verosimilitud y confianza usando la aproximación  $\chi^2$  con  $d = 1$ .

### 2.2.3. Verosimilitud Perfil

Consideremos un espacio parametral  $\Theta$  de dimensión  $k$  con elementos denotados por  $\theta = (\theta_1, \dots, \theta_d)$ . Supongamos que  $\theta$  se particiona en dos componentes  $(\theta^{(1)}, \theta^{(2)})$  y deseamos trabajar sólo con los parámetros incluidos en  $\theta^{(1)}$ . En ese caso, podemos usar la verosimilitud perfil que se define como

$$L_p(\theta^{(1)}) = \max_{\theta^{(2)} | \theta^{(1)}} L(\theta^{(1)}, \theta^{(2)}),$$

donde  $L$  es la función de verosimilitud. En palabras, la verosimilitud perfil se obtiene maximizando la función de verosimilitud evaluada en los elementos de  $\Theta$  con  $\theta^{(1)}$  fijo. Denotando por  $\hat{\theta}_n$  al estimador de máxima verosimilitud, entonces definimos a la verosimilitud relativa perfil para  $\theta^{(1)}$  como

$$R_p(\theta^{(1)}) = \frac{L_p(\theta^{(1)})}{L(\hat{\theta}_n)}.$$

Con la anterior formamos los conjuntos  $\{\theta : R_p(\theta^{(1)}) \geq l\}$  que llamaremos regiones de verosimilitud para  $\theta^{(1)}$  de nivel  $l$ , a los cuales, de modo similar al caso anterior, es posible asociar un nivel de confianza usando el siguiente teorema.

**Teorema 2.5** Sean  $x_1, \dots, x_n$  una muestra de variables aleatorias iid con función de distribución  $F_\theta$  que pertenece a una familia  $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$  con espacio parametral  $\Theta$  de dimensión  $k$ . Supongamos que se satisfacen las condiciones del Teorema 2.1. Sea  $\hat{\theta}_n$  el estimador de máxima verosimilitud para el parámetro del modelo  $\theta = (\theta^{(1)}, \theta^{(2)})$ , donde  $\theta^{(1)}$  es un subconjunto de tamaño  $r$  de las componentes de  $\theta$ . Entonces

$$-2\log(R_p(\theta^{(1)})) \xrightarrow{d} \chi_r^2.$$

Una demostración aparece en Serfling [9], Sección 4.4.4, pág. 158.

### 2.2.4. Aplicación a las Distribuciones de Valores Extremos

Primero presentamos las expresiones para las funciones de log-verosimilitud bajo la *DGVE* y bajo las *DVE*, respectivamente. Usaremos las parametrizaciones dadas en (1.11) y en la Nota 1.1, aunque para los tipos Fréchet y Weibull aplicamos la relación (1.12), y por lo tanto, esas funciones de verosimilitud están en términos de  $\xi$ .

*DGVE*:

$$\ell(\mu, \sigma, \xi; \mathbf{x}) = \begin{cases} \left( -n \log(-\xi\sigma) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^n \log(y_i) - \sum_{i=1}^n (y_i)^{-\frac{1}{\xi}} \right) \mathbf{I}_{\{1+\xi y_i > 0\}}, & \xi \neq 0; \\ -n \log(\sigma) - \sum_{i=1}^n \exp\{-y_i\} - \sum_{i=1}^n y_i, & \xi = 0, \end{cases} \quad (2.3)$$

con  $y_i = \left(\frac{x_i - \mu}{\sigma}\right)$ .

*DVE* tipo Fréchet:

$$\ell(b, a, \xi; \mathbf{x}) = \left( -n \log(\xi a) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^n \log(y_i) - \sum_{i=1}^n (y_i)^{-\frac{1}{\xi}} \right) \prod_{i=1}^n \mathbf{I}_{\{x_i > b\}}, \quad (2.4)$$

con  $y_i = \frac{x_i - b}{a}$ .

*DVE* tipo Gumbel:

$$\ell(b, a; \mathbf{x}) = -n \log(a) - \sum_{i=1}^n (y_i) - \sum_{i=1}^n \exp\{-y_i\}. \quad (2.5)$$

con  $y_i = \frac{x_i - b}{a}$ .

*DVE* tipo Weibull:

$$\ell(b, a, \xi; \mathbf{x}) = \left( -n \log(-\xi a) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^n \log(y_i) - \sum_{i=1}^n (y_i)^{-\frac{1}{\xi}} \right) \prod_{i=1}^n \mathbf{I}_{\{b > x_i\}}, \quad (2.6)$$

con  $y_i = \frac{b - x_i}{a}$ .

Dentro de las expresiones para las funciones de verosimilitud, notemos cómo el soporte de las funciones de densidad que las generan depende de los parámetros. La observación anterior es importante porque entonces no se satisfacen las condiciones usuales de regularidad

que garantizan consistencia, eficiencia y normalidad asintótica de los estimadores de máxima verosimilitud. Sin embargo, Smith [10] demostró que, a pesar de que no se satisfacen las condiciones usuales, los estimadores de máxima verosimilitud tienen las propiedades mencionadas cuando se tiene  $\xi > -0,5$ . Además, para  $-1 < \xi < -0,5$  los estimadores se pueden obtener pero no tienen esas propiedades asintóticas. Por último, cuando  $\xi < -1$  es posible que no se puedan obtener los estimadores de máxima verosimilitud.

Con las funciones de verosimilitud anteriores podemos obtener estimadores de máxima verosimilitud, así como funciones de verosimilitud perfil para los parámetros según el modelo que se haya considerado. Adicionalmente, aplicando el Teorema 2.2 también podemos obtener estimadores de máxima verosimilitud para niveles de retorno. Para ello, notemos que las funciones de distribución para las *DGVE* y las *DVE* son invertibles, y aplicando la inversa del modelo estimado por máxima verosimilitud a  $1-p$ , con  $p \in (0, 1)$ , obtenemos el estimador de máxima verosimilitud para el nivel de retorno asociado a un período de retorno  $p^{-1}$ . Las expresiones específicas para las inversas de las distribuciones, evaluadas en  $(1-p)$ , aparecen a continuación.

$$DGVE : \quad \begin{cases} x_p = \mu + \frac{\sigma}{\xi} \left( [-\log(1-p)]^{-\xi} - 1 \right), & \xi \neq 0; \\ x_p = \mu - \sigma \log(-\log(1-p)), & \xi = 0. \end{cases}$$

$$DVE \text{ tipo Fréchet} : \quad x_p = b + a (-\log(1-p))^{-\xi}.$$

$$DVE \text{ tipo Gumbel} : \quad x_p = b - a \log(-\log(1-p)).$$

$$DVE \text{ tipo Weibull} : \quad x_p = b - a (-\log(1-p))^{-\xi}.$$

Usando las expresiones anteriores también podemos reparametrizar las funciones de verosimilitud y obtener verosimilitudes perfiles para los niveles de retorno. Por ejemplo, para el caso de las *DGVE* con  $\xi \neq 0$ , y después de escoger un período de retorno  $p^{-1}$ , podemos despejar  $\sigma$  de la expresión para la inversa de esa distribución obteniendo

$$\sigma = \xi (x_p - \mu) \left( [-\log(1-p)]^{-\xi} - 1 \right)^{-1}.$$

Al sustituir la anterior en (2.3) obtenemos una parametrización en términos de  $(\mu, x_p, \xi)$ .

### 2.3. Enfoques para Hacer Inferencia Usando la Verosimilitud

Dentro de los enfoques para hacer inferencia que están basados en los conceptos y resultados presentados en la sección anterior, uno que aparece usualmente en la literatura, y que siguiendo a Castillo [2] denominaremos **método de máxima verosimilitud**, tiene las siguientes características. Supongamos que tenemos una muestra de máximos y que hemos seleccionado a la *DGVE* como familia de distribuciones para modelar a los máximos. Entonces, para ajustar un modelo de esa familia a la distribución de máximos, utilizamos como estimador puntual al estimador de máxima verosimilitud  $\hat{\theta} = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$ . Además, evaluando la inversa de la *DGVE* en  $1 - p$ , donde  $p \in [0, 1]$ , obtenemos el estimador de máxima verosimilitud para el nivel de retorno con periodo de retorno igual a  $p^{-1}$ . Por otro lado, para obtener estimadores de tipo intervalo tanto para los parámetros como para los niveles de retorno, utilizamos la normalidad asintótica de los estimadores de máxima verosimilitud. Como ejemplo consideremos un nivel de retorno  $u_p = G^{-1}(1 - p; \theta)$ , donde  $G$  es la *DGVE*. En ese caso, con el método de máxima verosimilitud tomamos como estimador puntual a  $\hat{u}_p = G^{-1}(1 - p; \hat{\theta})$ , y usando el Teorema 2.3 con  $h = F^{-1}$ , obtenemos como intervalo de confianza aproximada del  $(1 - \alpha)100\%$  a

$$\left( \hat{u}_p \pm z_{\alpha/2} \sqrt{\nabla u_p^t V_{\theta} \nabla u_p} \right),$$

donde

$$\nabla x_p = \begin{pmatrix} \frac{\nabla u_p}{\nabla \mu} & \frac{\nabla u_p}{\nabla \sigma} & \frac{\nabla u_p}{\nabla \xi} \end{pmatrix}$$

y  $z_{\alpha/2}$  es el  $(1 - \alpha/2)$  cuantil de la distribución normal estándar.

La presentación del método de máxima verosimilitud era indispensable puesto que es el más difundido, sin embargo, una manera alternativa para encontrar estimadores de intervalo consiste en utilizar las funciones de verosimilitud perfil y su aproximación  $\chi^2$ . Como menciona Coles [1], ésta aproximación es usualmente más exacta que la basada en la normalidad asintótica del estimador de máxima verosimilitud. Lo anterior justifica que en este trabajo nos concentramos en enfoques que, a diferencia del método de máxima verosimilitud, incluyen el uso de la función de verosimilitud perfil. A continuación presentamos dos de esos enfoques los cuales se distinguen por el modelo bajo el cual obtienen inferencias.

En la Sección 2.1 explicamos que es posible escoger entre dos tipos de modelos para hacer inferencia en el análisis de máximos por bloques: por un lado, podemos considerar

la *DGVE*; y por el otro, algún tipo de las *DVE*. El primer enfoque que explicamos toma como modelo a la *DGVE* y consiste en lo siguiente. Usando los estimadores de máxima verosimilitud, ajustamos un modelo de esa familia a la distribución de máximos. Luego, para obtener estimaciones de intervalo para algún parámetro o nivel de retorno, calculamos la función de verosimilitud perfil relativa, y con ésta, obtenemos regiones de verosimilitud cuyo nivel de verosimilitud le asocia un nivel de confianza a través de la aproximación  $\chi^2$  (véase el Cuadro 2.1). En lo que sigue nos referiremos al enfoque que acabamos de describir como **enfoque basado en la *DGVE***.

Por otro lado, al usar como modelo a algún tipo de las *DVE* obtenemos otro tipo de enfoque para hacer inferencia. El que consideramos a continuación está descrito en detalle en Vilchis [11] y es una propuesta de la Dra. Eloísa Díaz Francés, el Dr. Joaquín Ortega Sánchez y el estudiante Eliud Vilchis. Para seleccionar un tipo de las *DVE*, Díaz-Francés, Ortega y Vilchis proponen que consideremos la información que proporciona la función de verosimilitud perfil para el parámetro  $\xi$ , la cual calculamos bajo la *DGVE*. Dicha información se refiere a la plausibilidad que asocia a los valores de  $\xi$  positivos, negativos y al cero. Niveles de verosimilitud altos para valores positivos da evidencia a favor de un modelo tipo Fréchet, mientras que niveles altos para valores negativos es evidencia a favor de un modelo Weibull y un nivel de verosimilitud alto para el cero favorece un modelo Gumbel. Con la información anterior *sugerimos* uno o dos tipos de *DVE* como posibles modelos. Bajo cada uno de esos modelos, ajustamos una distribución a los datos usando los estimadores de máxima verosimilitud. Después, comparamos esas distribuciones considerando el ajuste con los datos. Finalmente, el modelo que *seleccionamos* es aquél con el que se obtuvo la distribución que presentó el mejor ajuste y que mejor explicó el comportamiento del fenómeno de interés. Con ese modelo calculamos como estimadores puntuales a los estimadores de máxima verosimilitud y como estimadores de intervalo a los intervalos de verosimilitud obtenidos a partir de la función de verosimilitud perfil para algún parámetro o nivel de retorno de interés. Una característica relevante del proceso de selección en este enfoque es la postura crítica que se adopta, en la cual, es a través de la validación de las distribuciones ajustadas que finalmente se determina el modelo más adecuado. Ese proceso de validación no sólo se refiere a considerar procedimientos estadísticos estándar como papel de probabilidad o Q-Q plots, sino que también hacen referencia a considerar el contexto del problema que se está considerando y que podría proporcionar información sobre el tipo de modelo más adecuado.

En la Sección 2.1 mencionamos algunos "críticas" a usar un enfoque en el que seleccionamos algún tipo de las *DVE*. Sin embargo, el proceso de selección que proponen Díaz-

Francés, Ortega y Vilchis merecen los siguientes comentarios. Usar la verosimilitud perfil para  $\xi$  como herramienta para sugerir la, o las, familias a considerar, implica que esa información la obtenemos de los datos observados. Además, dejar la posibilidad de que sea más de una familia la sugerida refleja la incertidumbre en el valor de  $\xi$ , y en ese caso, la selección de una de ellas también recurre a la evidencia proporcionada por los datos pues se hace con base en el modelo que presente el mejor ajuste y mejor describa el fenómeno. Por otro lado, notemos que bajo ese enfoque, la incertidumbre respecto al tipo de modelo está incluida en el proceso de selección del modelo pero no dentro de la estimación. Podríamos pensar que al separar, por un lado, la incertidumbre del tipo de modelo adecuado para los datos, y por el otro, la estimación misma del modelo, entonces bajo ese enfoque se obtienen mejores estimadores puntuales así como estimaciones de intervalo para los parámetros.

Para comparar el enfoque de Díaz-Francés, Ortega y Vilchis, por un lado, y el enfoque basado en la *DGVE*, por el otro, una opción es hacer un estudio basado en simulaciones. Dicho estudio consistiría en simular muestras de máximos a partir de una distribución conocida, generar estimaciones por los dos enfoques y compararlas. Sin embargo, lo anterior es difícil de realizar pues cuando la función de verosimilitud perfil para el parámetro de forma sugiera más de una familia, debemos identificar el modelo que mejor ajusta a los datos, para lo cual, es complicado establecer criterios que permitan hacerlo de manera automatizada. En este trabajo, para lograr comparar los enfoques, usaremos una versión más "simple" del propuesto por Díaz-Francés, Ortega y Vilchis y que explicamos a continuación. Primero, tomando como modelo a la *DGVE*, calculamos la función verosimilitud perfil del parámetro  $\xi$ , la cual denotaremos como  $R_p(\cdot)$  y la evaluamos en 0. Luego, seleccionamos -no solo sugerimos- un tipo de *DVE* con el siguiente criterio: si  $R_p(0) \geq 0.8$ , entonces seleccionamos la familia Gumbel; si  $R_p(0) < 0.8$  y  $\hat{\xi} > 0$ , escogemos las distribuciones tipo Fréchet; y finalmente, si  $R_p(0) < 0.8$  y  $\hat{\xi} < 0$ , tomamos como modelo las distribuciones Weibull. Finalmente, bajo el modelo seleccionado obtenemos estimadores de máxima verosimilitud así como regiones de confianza usando la verosimilitud perfil. El enfoque que acabamos de describir será denotado en lo que sigue como **enfoque basado en las *DVE***.

En el siguiente capítulo presentamos los resultados de un análisis basado en simulaciones con el que comparamos las inferencias obtenidas usando el enfoque basado en la *DGVE* y el basado en las *DVE*, respectivamente. Con dicho análisis estudiaremos cómo se comportan las estimaciones al considerar uno de los aspectos presentados en la propuesta de Díaz-Francés, Ortega y Vilchis: la aplicación de la función de verosimilitud perfil de  $\xi$  calculada bajo la *DGVE*, como herramienta que ayude a determinar una familia de *DVE*. Sin embargo,

debe quedar claro que los resultados que se obtengan no son una comparación entre dicha propuesta y el enfoque basado en la *DGVE* pues, al usar una versión "simplificada", no estamos incluyendo la posibilidad de que exista más de un modelo plausible. Debido a esa omisión, es razonable pensar que con la propuesta hecha por Díaz-Francés, Ortega y Vilchis se obtendrán mejores resultados que con la simplificación que será considerada para el estudio basado en simulaciones.

## Capítulo 3

# Comparación de Enfoques para Obtener Estimaciones

En el capítulo anterior explicamos dos tipos de enfoques para hacer inferencia cuando trabajamos con máximos por bloques, a uno lo llamamos enfoque basado en la *DGVE*, mientras que al otro, lo denominamos enfoque basado en las *DVE*. Para comparar esos enfoques realizamos un estudio basado en simulaciones. Las características de dicho estudio las explicamos en la Sección 3.1, mientras que los resultados obtenidos aparecen en la 3.2.

### 3.1. Características del Estudio

A grandes rasgos, el procedimiento que seguimos para hacer la comparación de los enfoques fue el siguiente: simulamos muestras de máximos, obtuvimos estimaciones bajo cada enfoque, y finalmente, comparamos los resultados con los valores reales usados para la simulación. Presentamos los detalles a continuación.

En lo que se refiere a las simulaciones de muestras de máximos por bloques fijamos las siguientes condiciones. Supusimos que las observaciones se obtuvieron diariamente y consideramos períodos de tiempo anual, por lo tanto, el tamaño de bloque fue  $n = 360$ . Además, trabajamos con diferentes escenarios modificando la cantidad de observaciones disponibles, específicamente, consideramos los casos en que se tiene información para 25, 50 y 100 años. Resumiendo, trabajamos con muestras de máximos obtenidas de bloques anuales y tamaños de muestra de 25, 50 y 100.

Para obtener las muestras de máximos dividimos en bloques realizaciones simuladas de

alguna distribución. En este trabajo usamos las familias de distribuciones que aparecen en los Cuadros 1.1, 1.2 y 1.3. Los elementos de esas familias que incluimos en el análisis fueron seleccionados, en general, de modo que su distribución perteneciera al dominio de atracción (DA) de una *DVE* con parámetro de forma  $\xi$  igual a alguna de las componentes del vector

$$[-0.5, -0.3, -0.1, -0.05, 0, 0.05, 0.1, 0.3, 0.5].$$

De manera específica, para cada familia de distribuciones con DA Fréchet escogimos 4 elementos, cada uno con un valor de  $\xi$  asociado distinto y que es igual a alguno de los siguientes: 0.05, 0.1, 0.3 o 0.5. Hicimos algo similar con las distribuciones que pertenecen al DA Weibull pero con valores de  $\xi$  negativos. Para las distribuciones de tipo Gumbel los parámetros que son de localización y escala los fijamos como 0 y 1, respectivamente. Un caso particular, sin embargo, es el de la familia de distribuciones Weibull<sub>m</sub> pues Embrechts, Klüpelberg y Mikosch [3] distinguen dos casos debido a que la elección de parámetros implica un comportamiento diferenciado en las colas. Incluimos esos dos "tipos" de distribuciones y las distinguimos añadiendo al nombre "C.P." o "C.L.", que son las iniciales para colas pesadas y colas livianas, respectivamente. Además, por simplicidad ya no incluiremos el subíndice "m" en el nombre de esa distribución. En los Cuadros 3.1, 3.2 y 3.3 presentamos, divididas por el tipo de dominio de atracción al que pertenezcan, las distribuciones específicas que incluimos en el análisis.

Por otro lado, con relación a las estimaciones que calculamos bajo cada enfoque, éstas se refieren al cálculo de intervalos de verosimilitud para niveles de retorno. Consideraremos un nivel de verosimilitud del 0.1465% y niveles de retorno asociados a períodos de retorno de 20 y 100 años.

Según lo mencionado en esta sección, tenemos que consideramos diferentes escenarios en este estudio que varían según la distribución seleccionada -lo cual incluye familia y parámetros específicos-, el tamaño de muestra de máximos por bloque y el nivel de retorno que se esté estimando. Dado que incluimos 31 distribuciones, 3 tamaños de muestra y estimaciones para 2 niveles de retorno, en total en este estudio trabajamos con 186 escenarios distintos. Para cada uno de ellos hicimos 10,000 repeticiones, es decir, se simularon 10,000 muestras de máximos y con ellas, bajo cada enfoque, obtuvimos intervalos de verosimilitud para el período de retorno correspondiente.

Distribución $F$ o densidad $f$	Parámetros	$\xi$
Burr $F(x) = 1 - \left(\frac{\kappa}{\kappa+x^\tau}\right)^\alpha$	$\alpha, \kappa, \tau > 0$	$\frac{1}{\alpha\tau}$
	$\alpha = \frac{1000}{75}, \kappa = 1, \tau = 1.5$	0.05
	$\alpha = \frac{100}{15}, \kappa = 1, \tau = 1.5$	0.1
	$\alpha = \frac{100}{45}, \kappa = 1, \tau = 1.5$	0.3
	$\alpha = \frac{20}{15}, \kappa = 1, \tau = 1.5$	0.5
Pareto $F(x) = 1 - \left(\frac{\kappa}{\kappa+x}\right)^\alpha$	$\alpha, \kappa > 0$	$\frac{1}{\alpha}$
	$\alpha = 20, \kappa = 1$	0.05
	$\alpha = 10, \kappa = 1$	0.1
	$\alpha = 10/3, \kappa = 1$	0.3
	$\alpha = 2, \kappa = 1$	0.5
Loggamma $f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} (\ln(x))^{\beta-1} x^{-\alpha-1}$	$\alpha, \beta > 0$	$\frac{1}{\alpha}$
	$\alpha = 20, \beta = 0.5$	0.05
	$\alpha = 10, \beta = 0.5$	0.1
	$\alpha = 10/3, \beta = 0.5$	0.3
	$\alpha = 2, \beta = 0.5$	0.5
Fréchet $F(x) = \exp\{-x^{-\alpha}\}$	$\alpha > 0$	$\frac{1}{\alpha}$
	$\alpha = 20$	0.05
	$\alpha = 10$	0.1
	$\alpha = 10/3$	0.3
	$\alpha = 2$	0.5

Cuadro 3.1: Parámetros de distribuciones seleccionadas con DA Fréchet

Distribución $F$ o densidad $f$	Parámetros
Weibull C.L. $F(x) = 1 - \exp\{-cx^\tau\}$	$c, \tau \geq 1$ $c = 1, \tau = 1.5$
Weibull C.P. $F(x) = 1 - \exp\{-cx^\tau\}$	$c, 0 < \tau < 1$ $c = 1, \tau = 0.5$
Lognormal $f(x) = \frac{1}{\sqrt{2\pi\sigma x}} \exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}$	$\mu \in R, \sigma > 0$ $\mu = 0, \sigma = 1$
Exponencial $F(x) = 1 - \exp\{-\lambda x\}$	$\lambda > 0$ $\lambda = 1$
Gamma $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\}$	$\alpha, \beta > 0$ $\alpha = 0.5, \beta = 1$
Logística $F(x) = 1 - [1 + \exp\{\frac{x-b}{a}\}]^{-1}$	$b \in R, a > 0$ $b = 0, a = 1$
Normal $f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$	$\mu \in R, \sigma > 0$ $\mu = 0, \sigma = 1$

Cuadro 3.2: Parámetros de distribuciones seleccionadas con DA Gumbel

Distribución $F$ o densidad $f$	Parámetros	$\xi$
Beta $f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1}$	$\alpha, \beta > 0$	$-\frac{1}{\beta}$
	$\alpha = 1, \beta = 20$	-0.05
	$\alpha = 1, \beta = 10$	-0.1
	$\alpha = 1, \beta = 10/3$	-0.3
	$\alpha = 1, \beta = 2$	-0.5
Weibull (DVE) $F(x) = \exp\{-(-x)^\alpha\}$	$\alpha > 0$	$-\frac{1}{\alpha}$
	$\alpha = 20$	-0.05
	$\alpha = 10$	-0.1
	$\alpha = 10/3$	-0.3
	$\alpha = 2$	-0.5

Cuadro 3.3: Parámetros de distribuciones seleccionadas con DA Weibull

### 3.2. Análisis de Resultados

Para facilitar la presentación de los resultados obtenidos seguiremos las siguientes convenciones. En la sección anterior presentamos las distribuciones que consideramos en este estudio. Como en algunos casos incluimos varios elementos de una misma familia, entonces para hacer referencia a ellos mencionaremos el parámetro de forma  $\xi$  de la *DVE* a cuyo dominio de atracción (DA) pertenezcan. Por ejemplo, las distribuciones de la familia Pareto serán distinguidas mencionando que tienen parámetro de forma  $\xi$  igual a .05, .1, .3 ó .5, según sea el caso. Por otro lado, para diferenciar entre los resultados obtenidos con cada enfoque, usaremos las etiquetas "*DVE*" o "*DGVE*" según se refieran al enfoque basado en las *DVE* o en la *DGVE*, respectivamente. Además, cuando en las gráficas se haga referencia a distribuciones que tienen asociado un parámetro de forma -en los términos explicados al comienzo de éste párrafo-, el valor para  $\xi$  será etiquetado como "ksi". Finalmente, para diferenciar las estimaciones calculadas que se refieren a los niveles de retorno con períodos de retorno de 20 y 100 años, usaremos, en algunos casos, las etiquetas "q95" y "q99" puesto que esos niveles de retorno corresponden a los cuantiles de 95 % y 99 %, respectivamente.

#### 3.2.1. Cobertura

Para comenzar la comparación de los intervalos de verosimilitud obtenidos bajo cada enfoque consideraremos su cobertura, la cual se refiere al porcentaje de veces que los intervalos calculados contienen al verdadero valor según las simulaciones que estemos considerando. Pero antes, presentaremos algunos comentarios relacionados con la aparición de errores que, en algunos casos, no permitieron determinar los intervalos en cuestión.

En el estudio que elaboramos es necesario utilizar rutinas de maximización numérica: en el cálculo de los estimadores de máxima verosimilitud y para obtener las funciones de verosimilitud perfil. Calculamos el porcentaje de veces que dichas rutinas fallaron en su búsqueda de un máximo. El comportamiento de esos porcentajes lo describimos a continuación. Para todas las distribuciones con dominio de atracción Fréchet o Gumbel no obtuvimos errores bajo ninguno de los enfoques, cuando usamos muestras de tamaño 50 y 100, mientras que, para muestras de tamaño 25, aparecieron errores que no superaron el 0.55 %. Los porcentajes para esos casos aparecen en los Cuadros 3.4 y 3.5. Por otro lado, para las distribuciones que pertenecen al DA Weibull los porcentajes de errores son mayores para las distribuciones que tienen asociado un menor parámetro de forma, así como en los casos en que se consideran estimaciones para el cuantil de 99 %.(véase el Cuadro 3.6) Para las distribuciones que tienen

Estimaciones para el cuantil de 95%										Estimaciones para el cuantil de 99%									
Distribuciones de la familia Pareto										Distribuciones de la familia Pareto									
tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5			tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5		
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE			DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	
25	0.08	0.02	0.05	0.01	0.06	0.13	0.11	0		25	0.1	0.02	0.07	0	0.03	0.31	0.2	0.01	
50	0	0	0	0	0	0	0.03	0		50	0	0	0	0	0	0	0.03	0	
100	0	0	0	0	0	0	0.05	0		100	0	0	0	0	0	0	0.01	0	
Distribuciones de la familia Loggamma										Distribuciones de la familia Loggamma									
tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5			tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5		
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE			DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	
25	0.08	0.02	0.05	0.01	0.03	0.09	0.05	0		25	0.12	0.02	0.07	0.01	0.04	0.22	0.15	0.01	
50	0	0	0	0	0	0	0.03	0		50	0	0	0	0	0	0	0.01	0	
100	0	0	0	0	0	0	0.02	0		100	0	0	0	0	0	0	0	0	
Distribuciones de la familia Burr										Distribuciones de la familia Burr									
tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5			tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5		
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE			DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	
25	0.15	0.04	0.1	0.02	0.01	0.01	0.17	0.02		25	0.2	0.05	0.13	0.02	0.01	0.01	0.3	0.05	
50	0	0	0	0	0	0	0.02	0		50	0	0	0	0	0	0	0.04	0	
100	0	0	0	0	0	0	0.04	0		100	0	0	0	0	0	0	0.01	0	
Distribuciones de la familia Fréchet										Distribuciones de la familia Fréchet									
tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5			tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5		
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE			DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	
25	0.14	0.06	0.05	0	0.02	0	0.12	0		25	0.17	0.06	0.06	0	0.01	0	0.28	0.02	
50	0	0	0	0	0	0	0.02	0		50	0	0	0	0	0	0	0.02	0	
100	0	0	0	0	0	0	0.04	0		100	0	0	0	0	0	0	0	0	

Cuadro 3.4: Porcentaje de errores en el cálculo de intervalos de verosimilitud para distribuciones que pertenecen al DA Fréchet.

asociado un parámetro de forma igual a -.5 se obtuvo una mayor cantidad de errores cuando se usó el enfoque basado en la *DGVE*, sin embargo, no rebasaron el 5 %.

Para calcular la cobertura de los intervalos de verosimilitud dividimos el número de ocasiones en que dichos intervalos contuvieron al nivel de retorno real entre el número de casos en que sí se obtuvieron los intervalos de verosimilitud. Dado que, en general, la cantidad de errores encontrados fue pequeña, entonces creemos que no afectan la comparación de las coberturas para las diferentes distribuciones consideradas.

Al analizar las coberturas obtenidas, observamos que, en general, son cercanas al 95 % para los intervalos calculados usando el enfoque basado en la *DGVE*. Lo anterior ocurrió para todos los tamaños de muestra que consideramos. Por otro lado, con las estimaciones



Estimaciones para el cuantil de 95%														
tamaño de muestra	Weibull CP		Lognormal		Weibull CL		Exponencial		Gamma		Logística		Normal	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	0.04	0	0.02	0	0.21	0.03	0.16	0.02	0.14	0.04	0.17	0.07	0.34	0.06
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Estimaciones para el cuantil de 99%														
tamaño de muestra	Weibull CP		Lognormal		Weibull CL		Exponencial		Gamma		Logística		Normal	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	0.05	0	0.03	0	0.33	0.03	0.21	0.02	0.23	0.04	0.27	0.08	0.52	0.05
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Cuadro 3.5: Porcentaje de errores en el cálculo de intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel.

Estimaciones para el cuantil de 95%								Estimaciones para el cuantil de 99%								
tamaño de muestra	Distribuciones de la familia Beta								Distribuciones de la familia Beta							
	ksi = -.05		ksi = -.1		ksi = -.3		ksi = -.5		ksi = -.05		ksi = -.1		ksi = -.3		ksi = -.5	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	0.01	0	0.01	0.01	0.11	0.11	1.35	1.3	0.03	0.01	0.02	0.01	0.34	0.15	4.78	1.8
50	0	0	0	0	0	0	0.54	0.41	0	0	0	0	0.02	0	3.71	1.23
100	0	0	0	0	0	0	0.23	0.14	0	0	0	0	0	0	5.38	2.2

tamaño de muestra	Distribuciones de la familia Weibull (DVE)								Distribuciones de la familia Weibull (DVE)							
	ksi = -.05		ksi = -.1		ksi = -.3		ksi = -.5		ksi = -.05		ksi = -.1		ksi = -.3		ksi = -.5	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	0.01	0.01	0	0	0.09	0.09	1.11	1.09	0.05	0.02	0	0	0.2	0.09	4.04	1.64
50	0.06	0	0	0	0.01	0.01	0.4	0.35	0.02	0	0	0	0.01	0	3.58	1
100	0.06	0	0	0	0	0	0.22	0.13	0.05	0	0	0	0	0	5.82	2.12

Cuadro 3.6: Porcentaje de errores obtenidos en el cálculo de intervalos de verosimilitud para distribuciones que pertenecen al DA Weibull.

basadas en las *DVE* aparece una reducción en la cobertura que es mayor con los intervalos de verosimilitud que estiman el cuantil de 99%. A continuación describimos el comportamiento de la cobertura para esas estimaciones con las distribución agrupadas según el tipo de su dominio de atracción.

Para las distribuciones con DA de tipo Fréchet encontramos que, en las cuatro familias consideradas -Pareto, Loggamma, Burr y Fréchet-, con  $\xi = .1$  ocurre la disminución más grande de la cobertura asociada a los intervalos obtenidos con el enfoque basado en las *DVE* (véanse las Figuras 3.1, 3.2, 3.3 y 3.4). Por otro lado, para casi todas los escenarios que involucran distribuciones con  $\xi$  mayor que .1 obtuvimos una reducción en la diferencia entre las coberturas de los intervalos obtenidos con cada enfoque. Las excepciones fueron las distribuciones con  $\xi = .3$  y tamaño de muestra de 25.

En el caso de las distribuciones con dominio de atracción Gumbel, cuyos resultados aparecen en la Figura 3.5, se pueden distinguir dos grupos. El primero incluye a las distribuciones Weibull C.L., Exponencial, Gamma, Logística y Normal, y se caracteriza porque las coberturas obtenidas usando el enfoque basado en las *DVE* tienen un comportamiento estable: aproximadamente 92% para los intervalos que estiman el cuantil de 95% y del 90% para los que estiman el cuantil de 99%. Por otro lado, el segundo grupo está formado por las distribuciones Weibull C.P. y Lognormal para las cuales se presentan caídas de la cobertura cuando se consideran tamaños de muestra de 25 y 50, siendo éstas reducciones mayores con las estimaciones referentes al cuantil de 99%.

Resta considerar los resultados para las familias de distribuciones Beta y Weibull (*DVE*) cuyas gráficas aparecen en las Figuras 3.6 y 3.7. Las coberturas para los intervalos obtenidos con el enfoque basado en las *DVE* y que estiman el cuantil de 95%, son superiores al 92%. Por otro lado, para los intervalos calculados bajo dicho enfoque y que estiman el cuantil de 99%, hay una disminución en la cobertura para las distribuciones con  $\xi = -.1$ . Recordemos que encontramos un comportamiento similar para las distribuciones con DA Fréchet y  $\xi = .1$ , aunque en ese caso las caídas fueron mayores. Para el resto de las distribuciones la diferencia entre las coberturas que obtuvimos con los dos enfoques se redujo, excepto para las distribuciones con  $\xi = -.3$  y tamaño de muestra igual a 25.



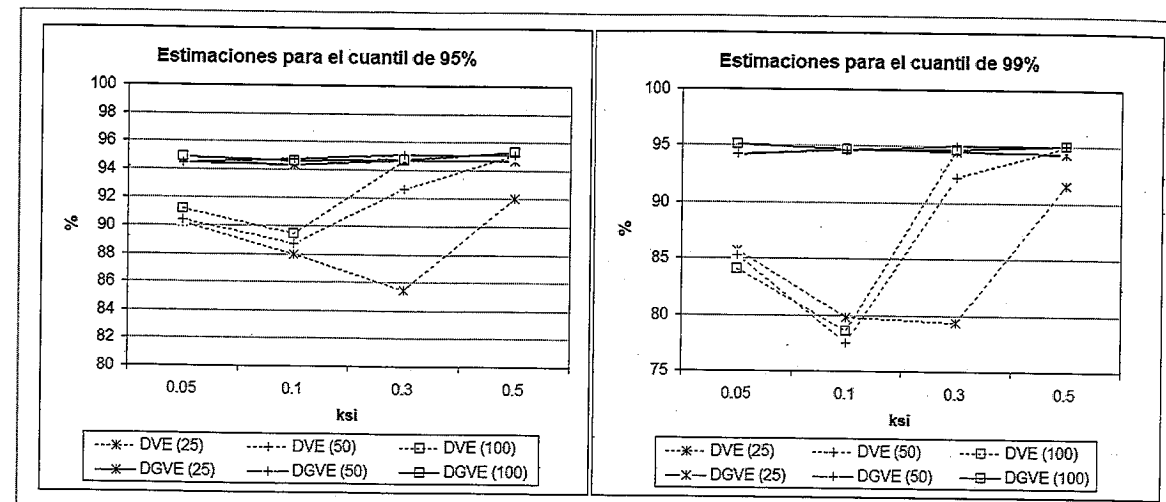


Figura 3.1: Cobertura de los intervalos de verosimilitud para cuantiles obtenidos con distribuciones de la familia Pareto. En las etiquetas los números entre paréntesis se refieren a los tamaños de muestra considerados.

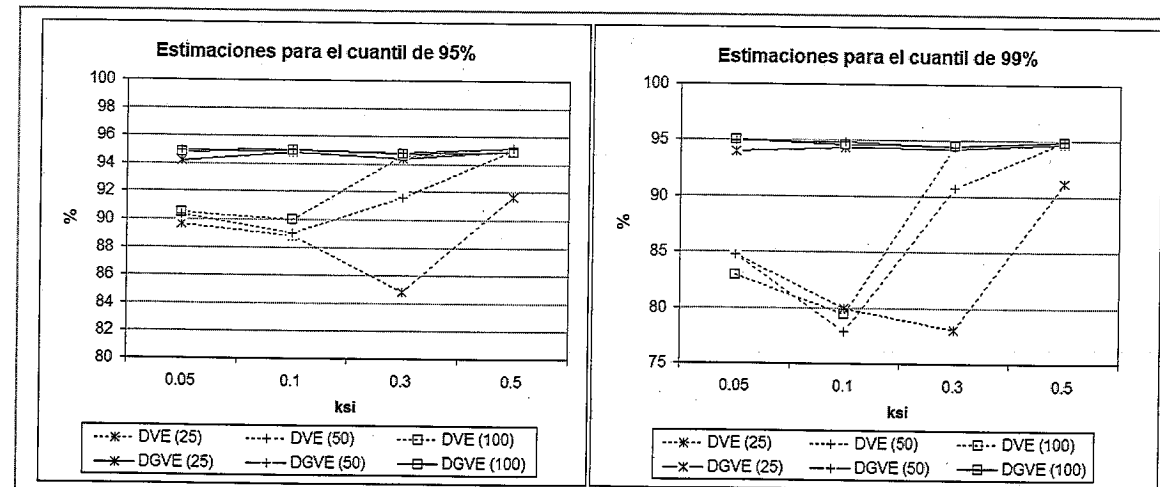


Figura 3.2: Cobertura de los intervalos de verosimilitud para cuantiles obtenidos con distribuciones de la familia Loggamma. En las etiquetas los números entre paréntesis se refieren a los tamaños de muestra considerados.

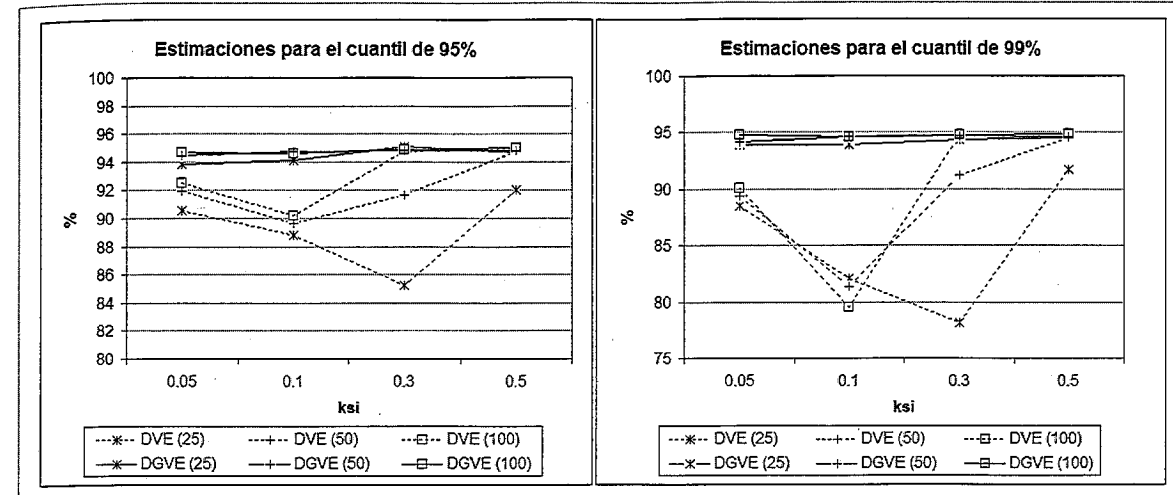


Figura 3.3: Cobertura de los intervalos de verosimilitud para cuantiles obtenidos con distribuciones de la familia Burr. En las etiquetas los números entre paréntesis se refieren a los tamaños de muestra considerados.

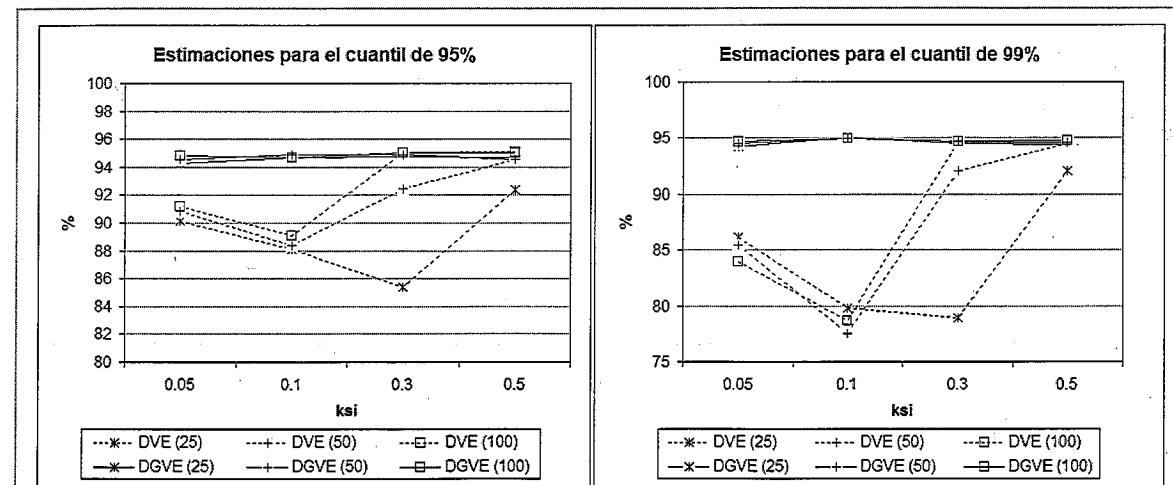


Figura 3.4: Cobertura de los intervalos de verosimilitud para cuantiles obtenidos con distribuciones de la familia Fréchet. En las etiquetas los números entre paréntesis se refieren a los tamaños de muestra considerados.

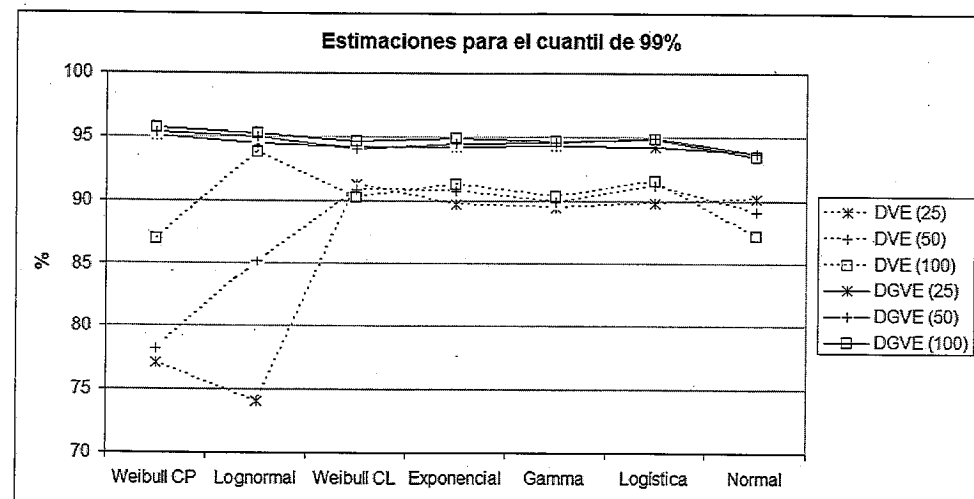
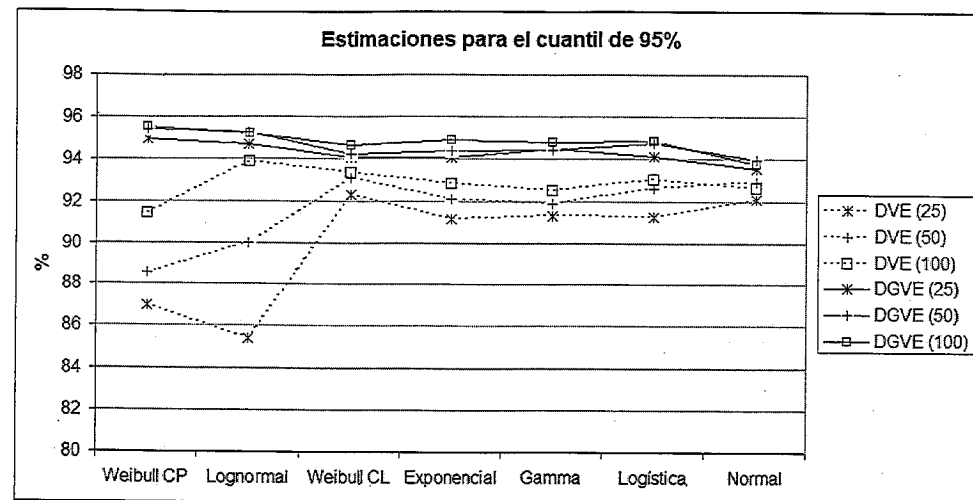


Figura 3.5: Cobertura de los intervalos de verosimilitud para cuantiles con distribuciones que pertenecen al DA Gumbel. En las etiquetas los números entre paréntesis se refieren a los tamaños de muestra considerados.

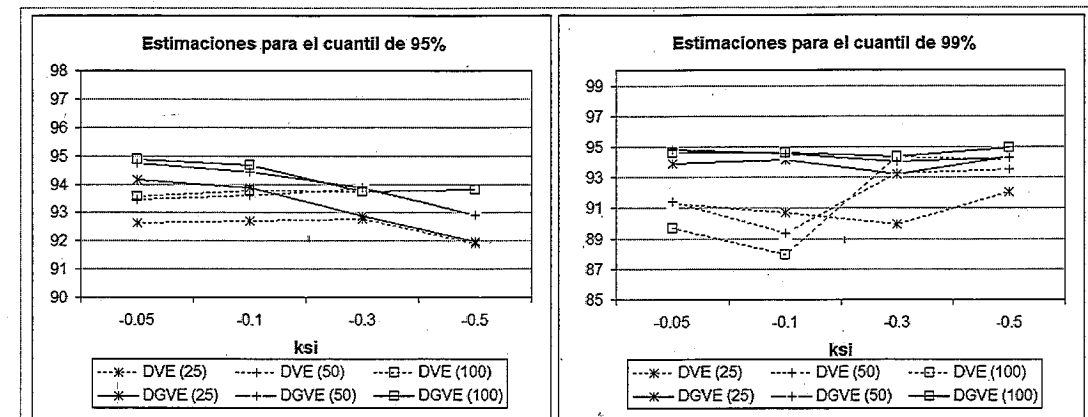


Figura 3.6: Cobertura de los intervalos de verosimilitud para cuantiles con distribuciones de la familia Beta. En las etiquetas los números entre paréntesis se refieren a los tamaños de muestra considerados.

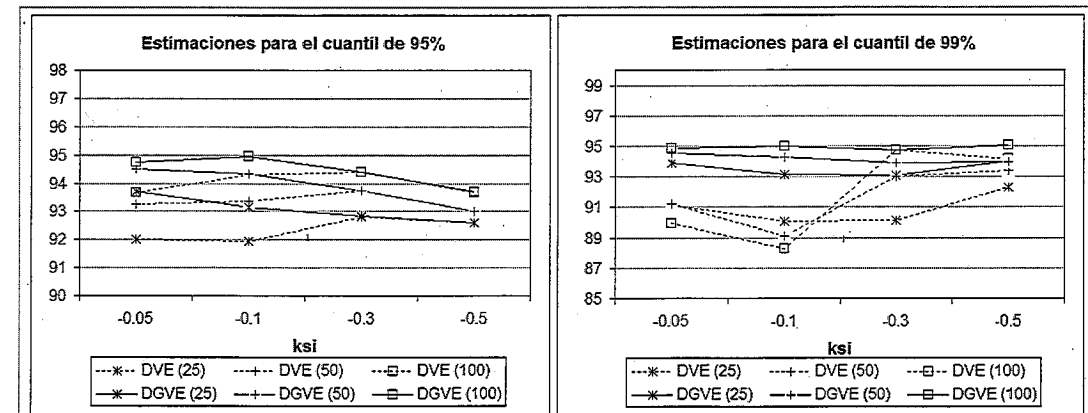


Figura 3.7: Cobertura de los intervalos de verosimilitud para cuantiles con distribuciones de la familia Weibull (DVE). En las etiquetas los números entre paréntesis se refieren a los tamaños de muestra considerados.

Para tener mayor información con relación a la disminución de la cobertura asociada a los intervalos obtenidos con el enfoque basado en las *DVE*, dividimos en dos grupos los casos en que el intervalo estimado no contiene al valor real del cuantil de interés: por un lado, los casos en que el extremo derecho cae por debajo del valor real; y por el otro, aquéllos en que el extremo izquierdo queda por arriba. Dicho de manera informal, dividimos esos intervalos considerando si quedaron por “abajo” o por “arriba” del valor real, respectivamente. Esta división fue sugerida por el Dr. Luis Escobar quien menciona que en las aplicaciones es de interés conocer si los intervalos están subestimando o sobrestimando al valor real. La cantidad de elementos que obtuvimos para cada uno de esos grupos aparecen expresados como porcentajes en los Cuadros 3.7, 3.8 y 3.9. Dichos porcentajes fueron calculados con respecto al número de casos en los que sí fue posible obtener el intervalo de verosimilitud. Para los casos en que consideramos distribuciones que pertenecen al DA Fréchet o Gumbel, notamos la siguiente relación: las combinaciones en las que los intervalos tuvieron las menores coberturas, son también las que presentan porcentajes altos de intervalos que quedaron por “abajo” del valor real. Ejemplos de esa situación son las distribuciones que pertenecen al DA Fréchet y tienen  $\xi$  igual a .05 ó .1, o bien, las distribuciones Weibull C.P. y Lognormal, en ambos casos considerando estimaciones para el cuantil de 99 %.

Para las distribuciones con DA Fréchet, los altos porcentajes de casos en que los intervalos estimados están por “abajo” del valor real podrían explicarse por una elección incorrecta de la *DVE*. El argumento anterior surge al considerar, para esos intervalos, los porcentajes en que fue seleccionado cada tipo de *DVE*. En los Cuadros 3.10 y 3.11 mostramos esos porcentajes. Como podemos apreciar, para las distribuciones con DA Fréchet y que tienen  $\xi$  igual a .05 ó *text.1*, obtuvimos que casi todos los intervalos de verosimilitud que cayeron por “abajo” del valor real fueron calculados usando un tipo de *DVE* distinto al Fréchet.

Estimaciones para el cuantil de 95%										Estimaciones para el cuantil de 99%									
tamaño de muestra	Distribuciones de la familia Pareto								Distribuciones de la familia Pareto										
	ksi = .05		ksi = .1		ksi = .3		ksi = .5		ksi = .05		ksi = .1		ksi = .3		ksi = .5				
	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba			
25	8.31	1.5	10.38	1.66	12.55	1.96	5.86	2.12	12.26	2.04	18.66	1.59	18.23	2.35	5.83	2.66			
50	7.7	1.87	9.49	1.78	5.46	1.85	2.86	2.05	12.57	2.15	20.41	2.05	5.65	2.09	2.69	2.33			
100	6.97	1.85	8.56	1.99	2.95	2.39	2.59	2.09	13.97	1.93	19.1	2.3	2.82	2.5	2.57	2.42			
tamaño de muestra	Distribuciones de la familia Loggamma								Distribuciones de la familia Loggamma										
	ksi = .05		ksi = .1		ksi = .3		ksi = .5		ksi = .05		ksi = .1		ksi = .3		ksi = .5				
	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba			
25	8.49	1.86	9.74	1.42	13.03	2.12	6.49	1.86	13.01	2.24	18	2.01	19.42	2.49	6.51	2.31			
50	7.94	1.71	9.4	1.59	6.39	1.97	2.91	2.04	13.24	1.99	20.26	1.88	6.93	2.33	2.86	2.2			
100	7.3	2.2	8.1	1.89	3.01	2.36	2.63	2.47	14.85	2.21	18.45	2.12	2.95	2.55	2.51	2.64			
tamaño de muestra	Distribuciones de la familia Burr								Distribuciones de la familia Burr										
	ksi = .05		ksi = .1		ksi = .3		ksi = .5		ksi = .05		ksi = .1		ksi = .3		ksi = .5				
	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba			
25	7.73	1.7	9.45	1.77	13.14	1.58	5.68	2.26	9.54	1.99	15.86	2.04	19.69	2.17	5.63	2.62			
50	6.04	2	8.51	1.82	6.44	1.88	2.86	2.37	8.36	2.29	16.72	1.88	6.8	2.01	2.69	2.79			
100	5.14	2.34	7.49	2.27	3.04	2.18	2.57	2.43	7.63	2.3	18.47	2.01	3.08	2.19	2.59	2.52			
tamaño de muestra	Distribuciones de la familia Fréchet								Distribuciones de la familia Fréchet										
	ksi = .05		ksi = .1		ksi = .3		ksi = .5		ksi = .05		ksi = .1		ksi = .3		ksi = .5				
	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba			
25	8.21	1.61	10.39	1.48	12.96	1.69	5.41	2.21	12.09	1.72	18.52	1.65	18.91	2.17	5.13	2.76			
50	7.38	1.79	9.74	1.86	5.45	2.1	3.04	2.41	12.6	1.96	20.56	1.95	5.51	2.38	2.86	2.59			
100	6.8	1.98	8.83	2.06	2.93	2.04	2.55	2.38	13.9	2.09	19.36	2.02	2.78	2.55	2.51	2.75			

Cuadro 3.7: Porcentaje de veces que los intervalos calculados con distribuciones que pertenecen al DA Fréchet, y que fueron calculados con el enfoque basado en las *DVE*, caen por abajo o arriba, respectivamente, del valor real para el cuantil correspondiente.

Estimaciones para el cuantil de 95%														
tamaño de muestra	Weibull CP		Lognormal		Weibull CL		Exponencial		Gamma		Logística		Normal	
	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba
25	11.48	1.54	12.91	1.71	5.89	1.82	7.1	1.76	7.02	1.65	6.94	1.78	5.91	1.99
50	9.88	1.58	8.42	1.59	4.7	2.23	5.87	2.04	6.34	1.8	5.49	1.88	4.64	2.42
100	6.86	1.72	4.13	2.01	3.82	2.83	5.01	2.15	5.16	2.33	4.67	2.31	4.02	3.37

Estimaciones para el cuantil de 99%														
tamaño de muestra	Weibull CP		Lognormal		Weibull CL		Exponencial		Gamma		Logística		Normal	
	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba
25	21.07	1.88	23.68	2.25	5.81	2.95	8.01	2.25	8.16	2.37	7.71	2.46	5.69	4.12
50	19.53	2.23	12.64	2.22	4.59	4.55	6.54	2.69	7.68	2.44	6.18	2.61	4.96	5.92
100	10.81	2.27	3.71	2.51	3.81	5.95	5.86	2.85	6.72	2.92	5.4	3.08	4.56	8.25

Cuadro 3.8: Porcentaje de veces que los intervalos calculados con distribuciones que pertenecen al DA Gumbel, y que fueron calculados con el enfoque basado en las DVE, caen por abajo o arriba, respectivamente, del valor real del cuantil correspondiente.

Estimaciones para el cuantil de 95%							Estimaciones para el cuantil de 99%									
tamaño de muestra	Distribuciones de la familia Beta						tamaño de muestra	Distribuciones de la familia Beta								
	ksi = -.05		ksi = -.1		ksi = -.3			ksi = -.3		ksi = -.5						
	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba		Abajo	Arriba	Abajo	Arriba					
25	5.8	1.56	5.65	1.66	6	1.21	6.6	1.38	5.71	3.02	4.87	4.43	6.2	3.86	6.47	1.32
50	4.53	2.02	4.26	2.15	4.84	1.32	5.69	1.38	4.24	4.34	4.03	6.59	5.05	1.67	5.66	0.71
100	3.78	2.67	3.71	2.5	4.61	1.65	4.79	1.4	3.67	6.62	3.69	8.39	4.49	1.17	4.85	0.85

tamaño de muestra	Distribuciones de la familia Weibull (DVE)						tamaño de muestra	Distribuciones de la familia Weibull (DVE)								
	ksi = -.05		ksi = -.1		ksi = -.3			ksi = -.3		ksi = -.5						
	Abajo	Arriba	Abajo	Arriba	Abajo	Arriba		Abajo	Arriba	Abajo	Arriba					
25	6.05	1.94	5.97	2.09	5.91	1.29	6.22	1.09	5.61	3.26	5.47	4.44	6.03	3.82	6.38	1.19
50	4.6	2.16	4.19	2.45	4.75	1.51	5.79	1.17	4.27	4.5	3.97	6.93	4.87	2.09	5.88	0.64
100	3.78	2.56	3.41	2.3	4.07	1.54	4.88	1.44	3.5	6.54	3.28	8.47	4.02	1.22	4.67	0.85

Cuadro 3.9: Porcentaje de veces que los intervalos calculados con distribuciones que pertenecen al DA Weibull, y que fueron calculados con el enfoque basado en las DVE, caen por abajo o arriba, respectivamente, del valor real para el cuantil correspondiente.

Estimaciones para el cuantil de 95%												
tamaño de muestra	Distribuciones de la familia Pareto											
	ksi = .05			ksi = .1			ksi = .3			ksi = .5		
	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet
25	61.97	38.03	0.00	47.78	52.22	0.00	16.57	82.71	0.72	8.87	75.60	15.53
50	52.47	47.53	0.00	33.72	66.07	0.21	6.78	77.11	16.12	0.70	14.34	84.97
100	40.17	59.83	0.00	18.58	79.21	2.22	0.00	13.56	86.44	0.00	0.00	100.00

tamaño de muestra	Distribuciones de la familia Loggamma											
	ksi = .05			ksi = .1			ksi = .3			ksi = .5		
	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet
25	56.54	43.46	0.00	47.33	52.67	0.00	17.58	81.20	1.23	8.01	80.28	11.71
50	45.72	54.03	0.25	34.47	65.11	0.43	6.73	77.47	15.81	0.00	16.84	83.16
100	34.66	64.93	0.41	21.61	77.04	1.36	0.00	16.61	83.39	0.00	0.00	100.00

tamaño de muestra	Distribuciones de la familia Burr											
	ksi = .05			ksi = .1			ksi = .3			ksi = .5		
	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet
25	69.34	30.66	0.00	56.30	43.60	0.11	17.28	81.89	0.84	8.45	76.23	15.32
50	65.07	34.93	0.00	42.07	57.81	0.12	5.75	77.95	16.30	0.70	14.34	84.97
100	62.65	37.35	0.00	32.84	66.89	0.27	0.33	19.08	80.59	0.00	0.00	100.00

tamaño de muestra	Distribuciones de la familia Fréchet											
	ksi = .05			ksi = .1			ksi = .3			ksi = .5		
	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet
25	63.83	36.05	0.12	44.56	55.44	0.00	18.60	80.32	1.08	8.50	74.31	17.19
50	53.66	46.34	0.00	32.55	67.25	0.21	6.42	74.50	19.08	0.33	8.22	91.45
100	44.27	55.59	0.15	20.61	76.44	2.94	0.34	13.65	86.01	0.00	0.00	100.00

Cuadro 3.10: Porcentaje en que fue seleccionado cada tipo de DVE para los casos en que los intervalos de verosimilitud, que se obtienen a partir de distribuciones con DA Fréchet y que fueron calculados con el enfoque basado en las DVE, quedaron por "abajo" del valor real.

Estimaciones para el cuantil de 99%												
Distribuciones de la familia Pareto												
tamaño de muestra	ksi = .05			ksi = .1			ksi = .3			ksi = .5		
	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet
25	57.75	42.25	0.00	41.53	58.47	0.00	12.29	87.66	0.05	8.75	79.76	11.49
50	46.94	53.06	0.00	26.56	73.44	0.00	6.73	84.25	9.03	0.74	15.24	84.02
100	34.15	65.86	0.00	13.46	86.39	0.16	0.00	14.18	85.82	0.00	0.00	100.00

Distribuciones de la familia Loggamma												
tamaño de muestra	ksi = .05			ksi = .1			ksi = .3			ksi = .5		
	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet
25	50.88	49.12	0.00	39.33	60.67	0.00	13.65	86.35	0.00	7.83	85.25	6.91
50	41.62	58.38	0.00	26.01	73.99	0.00	6.35	85.14	8.51	0.00	17.13	82.87
100	28.49	71.52	0.00	15.45	84.50	0.05	0.00	17.29	82.71	0.00	0.00	100.00

Distribuciones de la familia Burr												
tamaño de muestra	ksi = .05			ksi = .1			ksi = .3			ksi = .5		
	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet
25	69.60	30.40	0.00	51.77	48.24	0.00	13.76	86.03	0.20	8.53	81.00	10.48
50	65.43	34.57	0.00	36.84	63.16	0.00	5.59	86.03	8.38	0.74	15.24	84.02
100	61.08	38.93	0.00	24.26	75.74	0.00	0.32	18.83	80.84	0.00	0.00	100.00

Distribuciones de la familia Fréchet												
tamaño de muestra	ksi = .05			ksi = .1			ksi = .3			ksi = .5		
	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet	Weibull	Gumbel	Fréchet
25	59.39	40.61	0.00	38.82	61.18	0.00	15.12	84.72	0.16	8.58	81.29	10.14
50	48.25	51.75	0.00	27.87	72.13	0.00	6.53	83.85	9.62	0.35	8.74	90.91
100	36.48	63.53	0.00	15.70	84.19	0.10	0.36	14.39	85.25	0.00	0.00	100.00

Cuadro 3.11: Porcentaje en que fue seleccionado cada tipo de DVE para los casos en que los intervalos de verosimilitud, que se obtienen a partir de distribuciones con DA Fréchet y que fueron calculados con el enfoque basado en las DVE, quedaron por "abajo" del valor real.

### 3.2.2. Cociente Entre las Longitudes de los Intervalos

Cuando evaluamos una función de verosimilitud perfil es necesario maximizar la función de verosimilitud. Para los cálculos en este estudio que se refieren a las funciones de verosimilitud perfil para cuantiles, dicha maximización se hizo numéricamente, y por lo tanto, fue necesario seleccionar en cada caso un punto de inicio para comenzar la búsqueda del máximo. Dicha selección puede originar un error de tipo numérico que explicamos a continuación. Tomando como punto de inicio para la búsqueda del máximo a los estimadores de máxima verosimilitud, entonces puede ocurrir que, al evaluar la función de verosimilitud perfil en valores cada vez más alejados del estimador de máxima verosimilitud del cuantil de interés  $\hat{x}_p$ , encontremos que los valores obtenidos como evaluaciones dan un "salto" hacia cero. Es decir, la función de verosimilitud perfil puede comportarse como si tuviera una discontinuidad de salto. La situación anterior se puede evitar "actualizando" los puntos de inicio como se explica a continuación. Supongamos que tenemos reales  $x_1 \leq \dots \leq x_n$  mayores que  $\hat{x}_p$  en los cuales se desea evaluar la función de verosimilitud perfil. Para evaluar en  $x_1$  tomamos como punto de inicio a los estimadores de máxima verosimilitud  $\hat{\theta}$  y maximizamos la función de verosimilitud con  $x_1$  fijo, con lo cual, obtenemos el valor de la función de verosimilitud perfil evaluada en  $x_1$  y un vector  $\theta_1$  en el cual la función de verosimilitud alcanza su máximo. Posteriormente, para evaluar la función de verosimilitud perfil en  $x_2$  utilizamos como punto de inicio a  $\theta_1$ , y maximizando, obtenemos el valor de la función de verosimilitud perfil y un vector  $\theta_2$  donde alcanza su máximo la función de verosimilitud. Procediendo de esa forma continuamos hasta obtener todas las evaluaciones requeridas.

La situación descrita en el párrafo anterior es de interés porque afectó la forma en que calculamos los intervalos de verosimilitud para los cuantiles. Supongamos que estamos calculando el extremo derecho de uno de esos intervalos y sigamos denotando por  $\hat{x}_p$  al estimador de máxima verosimilitud del cuantil de interés. Entonces, el primer paso consistió en buscar un intervalo en el cual, al evaluar a la función de verosimilitud perfil, encontráramos valores por arriba y por abajo del nivel de verosimilitud deseado, y en consecuencia, fuera candidato a contener el extremo derecho del intervalo de verosimilitud. Para ello, avanzamos progresivamente con valores a la derecha de  $\hat{x}_p$ , los cuales obtuvimos sumando múltiplos de algún real positivo  $\delta$ , evaluando la función de verosimilitud en dichos puntos y continuando hasta que se encontraba un valor inferior al nivel de verosimilitud deseado. Si  $\delta$  es "muy grande", entonces puede ocurrir un error numérico como el descrito en el párrafo anterior que no permitiría determinar el extremo derecho del intervalo de verosimilitud, mientras que si es muy

“pequeño”, es necesario una gran cantidad de evaluaciones para encontrar una que esté por debajo del nivel de verosimilitud del intervalo que estamos calculando, y en consecuencia, el tiempo para el cálculo de dichos intervalos es mayor.

Para enfrentar la situación anterior, diseñamos el programa de modo que calculara los intervalos de verosimilitud comenzando con un valor para  $\delta$  fijo, y si encontraba un error, entonces repetía el procedimiento usando un valor más pequeño para  $\delta$ . Aunque procediendo de esa manera eliminamos el problema numérico descrito al comienzo de esta sección, implicó que el tiempo de ejecución se elevara considerablemente. Particularmente ese problema apareció cuando trabajamos con distribuciones que pertenecen al dominio de atracción tipo Fréchet o Gumbel. Para paliar esa situación, decidimos establecer un “límite” en la búsqueda del extremo derecho de los intervalos como se explica a continuación. Al evaluar la función de verosimilitud perfil en puntos de la forma  $\hat{x}_p + k\delta$ , donde  $k$  es un entero positivo, si antes de encontrar una evaluación de la función de verosimilitud perfil que sea menor al nivel de verosimilitud de interés, ocurría que  $\hat{x}_p + k\delta$  era mayor que un número real fijado con anticipación, entonces el programa se detenía registrando que el extremo derecho del intervalo había superado ese cantidad. Para fijar ese “límite”, fue necesario utilizar un criterio aplicable a todas las distribuciones, el cual consistió en tomar el nivel de retorno asociado a un período de retorno de 10 millones de años calculado con respecto a la distribución de máximos que corresponda según la distribución que se esté considerando. Notemos que con esa elección el límite se ajusta a las condiciones específicas de cada distribución.

De modo análogo a cómo el análisis de cobertura resultaba “afectado” por la existencia de errores en la determinación de los intervalos de verosimilitud, en este caso la decisión de utilizar un límite en el cálculo de los intervalos de verosimilitud impone una restricción cuando queremos comparar la longitud de los intervalos de verosimilitud obtenidos bajo cada enfoque. Por lo tanto, es de interés calcular, para cada escenario, el porcentaje de casos en que el extremo derecho excedió el límite fijado.

En general, los porcentajes de excesos del límite son más altos cuando consideramos intervalos de verosimilitud para el cuantil de 99% que cuando se refieren al cuantil de 95%. Además, también son mayores cuando disminuye el tamaño de muestra. Presentamos esos porcentajes en los Cuadros 3.12 y 3.13.

Para poder comparar las longitudes de los intervalos de verosimilitud, además de que es necesario que bajo ninguno de los enfoques el extremo derecho de los intervalos de verosimilitud rebasen el límite, también requerimos que ninguno de los casos presente un error numérico que evite el cálculo de los intervalos. Al considerar el efecto combinado de los as-

Estimaciones para el cuantil de 95%								Estimaciones para el cuantil de 99%									
Distribuciones de la familia Pareto								Distribuciones de la familia Pareto									
tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5		tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE		DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	1.29	1.28	0.51	0.49	0.04	0.04	0.01	0.01	25	29.21	24.4	20.77	19.9	4.39	4.39	0.99	1
50	0	0	0	0	0	0	0	0	50	2.48	2.48	0.81	0.81	0	0	0	0
100	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
Distribuciones de la familia Loggamma								Distribuciones de la familia Loggamma									
tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5		tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE		DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	1.3	1.29	0.56	0.56	0.02	0.02	0	0	25	30.66	26	22.03	21.1	4.54	4.53	0.7	0.7
50	0	0	0	0	0	0	0	0	50	2.49	2.49	1.03	1.03	0	0	0	0
100	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
Distribuciones de la familia Burr								Distribuciones de la familia Burr									
tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5		tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE		DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	1.24	1.24	0.49	0.49	0.02	0.02	0.01	0.01	25	29.08	20.8	20.61	19.4	4.25	4.25	1.22	1.23
50	0	0	0	0	0	0	0	0	50	2.2	2.2	0.7	0.7	0.01	0.01	0	0
100	0	0	0	0	0	0	0	0	100	0.02	0.02	0	0	0	0	0	0
Distribuciones de la familia Fréchet								Distribuciones de la familia Fréchet									
tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5		tamaño de muestra	ksi = .05		ksi = .1		ksi = .3		ksi = .5	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE		DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	1.05	1.05	0.52	0.5	0	0	0	0	25	29.54	25	21.08	20.3	4.83	4.82	0.99	0.99
50	0	0	0	0	0	0	0	0	50	2.22	2.22	1.01	1.01	0.02	0.02	0	0
100	0	0	0	0	0	0	0	0	100	0.01	0.01	0	0	0	0	0	0

Cuadro 3.12: Porcentaje de casos en que el extremo derecho de los intervalos de verosimilitud excedió el límite para las distribuciones que pertenecen al DA Fréchet.

Estimaciones para el cuantil de 95%														
tamaño de muestra	Weibull CP		Lognormal		Weibull CL		Exponencial		Gamma		Logística		Normal	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	1.01	1.01	0.31	0.31	3.2	3.06	2.17	2.11	0.27	0.05	2.26	2.19	3.38	2.83
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Estimaciones para el cuantil de 99%														
tamaño de muestra	Weibull CP		Lognormal		Weibull CL		Exponencial		Gamma		Logística		Normal	
	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE	DGVE	DVE
25	32.85	31.59	16.81	16.71	40.92	14.53	38.79	22.28	0.98	0.08	39.74	22.48	42.7	10.17
50	3.26	3.26	0.39	0.39	7.01	5.77	5.49	5.35	0.55	0.09	5.18	5.06	7.63	4.17
100	0.01	0.01	0	0	0.04	0.04	0.06	0.06	0	0	0.02	0.02	0.07	0.07

Cuadro 3.13: Porcentaje de casos en que el extremo derecho de los intervalos de verosimilitud excedió el límite para las distribuciones que pertenecen al DA Gumbel.

pectos anteriores obtenemos, para cada escenario, el porcentaje de casos para los cuales sí se pueden comparar la longitud de los intervalos. Esos porcentajes aparecen en los Cuadros 3.14, 3.15 y 3.16.

En lo que sigue de esta sección analizamos el comportamiento de los cocientes que resultan de dividir la longitud de los intervalos obtenidos bajo cada enfoque tomando como numerador las estimaciones basadas en las *DVE*. Los cocientes así obtenidos los presentamos usando box-plots, o también llamadas, gráficas de “cajas” y “bigotes”. Esas gráficas fueron construidas con los siguientes criterios. Las líneas inferior y superior de la caja son los percentiles del 25% y 75% de la muestra, respectivamente. La distancia entre la parte superior e inferior de la caja lo denominamos rango intercuartílico. La línea en la mitad de la caja es la mediana de la muestra. Si la mediana no está localizada en el centro, es indicativo de que existe sesgo. Los bigotes son las líneas que se extienden por arriba y abajo de la caja. Muestran la extensión del resto de la muestra (a menos que existan outliers). Suponiendo que no existen outliers, el máximo de la muestra es la parte más alta del bigote superior, mientras que el mínimo de la muestra es la parte más baja del bigote inferior. Consideraremos a una observación un outlier cuanto tienen un valor que esta a una distancia de más de 1.5 veces el rango intercuartílico con respecto a la parte superior o inferior de la caja. En las gráficas esos errores son denotados por “+”.

Calculamos los box-plots para los cocientes correspondientes, los cuales, en la mayoría de

Estimaciones para el cuantil de 95%				
tamaño de muestra	Distribuciones de la familia Pareto			
	ksi = .05	ksi = .1	ksi = .3	ksi = .5
25	98.63	99.44	99.8	99.88
50	100	100	100	99.97
100	100	100	100	99.95

Estimaciones para el cuantil de 99%				
tamaño de muestra	Distribuciones de la familia Pareto			
	ksi = .05	ksi = .1	ksi = .3	ksi = .5
25	70.69	79.16	95.3	98.79
50	97.52	99.19	100	99.97
100	100	100	100	99.99

Distribuciones de la familia Loggamma				
tamaño de muestra	ksi = .05			
	ksi = .1	ksi = .3	ksi = .5	
25	98.62	99.39	99.89	99.95
50	100	100	100	99.97
100	100	100	100	99.98

Distribuciones de la familia Loggamma				
tamaño de muestra	ksi = .05			
	ksi = .1	ksi = .3	ksi = .5	
25	69.22	77.9	95.23	99.14
50	97.51	98.97	100	99.99
100	100	100	100	100

Distribuciones de la familia Burr				
tamaño de muestra	ksi = .05			
	ksi = .1	ksi = .3	ksi = .5	
25	98.61	99.41	99.97	99.81
50	100	100	100	99.98
100	100	100	100	99.96

Distribuciones de la familia Burr				
tamaño de muestra	ksi = .05			
	ksi = .1	ksi = .3	ksi = .5	
25	70.72	79.26	95.74	98.43
50	97.8	99.3	99.99	99.96
100	99.98	100	100	99.99

Distribuciones de la familia Fréchet				
tamaño de muestra	ksi = .05			
	ksi = .1	ksi = .3	ksi = .5	
25	98.81	99.43	99.98	99.88
50	100	100	100	99.98
100	100	100	100	99.96

Distribuciones de la familia Fréchet				
tamaño de muestra	ksi = .05			
	ksi = .1	ksi = .3	ksi = .5	
25	70.29	78.86	95.16	98.71
50	97.78	98.99	99.98	99.98
100	99.99	100	100	100

Cuadro 3.14: Porcentaje de casos en que bajo los dos enfoques se pudo calcular los intervalos de verosimilitud para distribuciones en el DA Fréchet.

los casos, tienen el siguiente comportamiento: dado un tamaño de muestra, toman valores menores cuando consideramos el cuantil más alto de 99%, mientras que con un cuantil dado, los valores que aparecen tienden a “acercarse” al 1 conforme aumenta el tamaño de muestra. Por lo tanto, para cuantiles altos y tamaños de muestra pequeños se tiene que los intervalos basados en las *DVE* son más pequeños comparados con los obtenidos a partir de la *DGVE*.

Para las distribuciones que pertenecen al DA Fréchet, los box-plots de los cocientes aparecen de la Figura 3.8 a la 3.11. En esas figuras aparecen por “renglones” las estimaciones correspondientes a cada cuantil, mientras que por “renglón”, los resultados para cada tamaño de muestra. Para las cuatro familias de distribuciones que consideramos obtenemos el siguiente comportamiento. Cuando el parámetro de forma aumenta, aparece una tendencia a que las cajas y los bigotes se hagan más estrechos, llegando incluso a aparecer como una sola

Estimaciones para el cuantil de 95%							
tamaño de muestra	Weibull CP	Lognormal	Weibull CL	Exponencial	Gamma	Logística	Normal
25	98.95	99.67	96.59	97.67	99.58	97.57	96.28
50	100	100	100	100	100	100	100
100	100	100	100	100	100	100	100

Estimaciones para el cuantil de 99%							
tamaño de muestra	Weibull CP	Lognormal	Weibull CL	Exponencial	Gamma	Logística	Normal
25	67.1	83.16	58.75	61	98.78	59.99	56.78
50	96.74	99.61	92.99	94.51	99.45	94.82	92.37
100	99.99	100	99.96	99.94	100	99.98	99.93

Cuadro 3.15: Porcentaje de casos en que bajo los dos enfoques se pudo calcular los intervalos de verosimilitud para distribuciones en el DA Gumbel.

Estimaciones para el cuantil de 95%					Estimaciones para el cuantil de 99%				
tamaño de muestra	Distribuciones de la familia Beta				tamaño de muestra	Distribuciones de la familia Beta			
	ksi = -.05	ksi = -.1	ksi = -.3	ksi = -.5		ksi = -.05	ksi = -.1	ksi = -.3	ksi = -.5
25	99.99	99.99	99.99	98.61	25	99.95	99.98	99.66	95.05
50	100	100	100	99.46	50	100	100	99.98	95.79
100	100	100	100	99.74	100	100	100	100	93.48

tamaño de muestra	Distribuciones de la familia Weibull (DVE)				tamaño de muestra	Distribuciones de la familia Weibull (DVE)			
	ksi = -.05	ksi = -.1	ksi = -.3	ksi = -.5		ksi = -.05	ksi = -.1	ksi = -.3	ksi = -.5
25	99.99	100	99.91	98.87	25	99.93	100	99.8	95.73
50	99.94	100	99.99	99.59	50	99.98	100	99.99	96.17
100	99.94	100	100	99.78	100	99.95	100	100	93.14

Cuadro 3.16: Porcentaje de casos en que bajo los dos enfoques se pudo calcular los intervalos de verosimilitud para distribuciones en el DA Weibull.

recta. Por lo tanto, cuando aumenta el parámetro de forma, las longitudes de los intervalos obtenidos bajo cada enfoque tienden a ser más parecidas

En lo que se refiere a las familias que pertenecen al DA Gumbel, el comportamiento de los cocientes nos sugiere dividir las distribuciones en dos grupos, los cuales coinciden con los encontrados con el análisis de la cobertura. Uno de esos grupos está constituido por las distribuciones Weibull C.P. y Lognormal las cuales, como aparece en la Figura 3.12, en casi todos los casos tuvieron valores más "cercaños" a uno si consideramos que presentan, comparado con el resto de las distribuciones con DA Gumbel, tamaños más pequeños de las cajas y/o una posición cercana a 1 de la recta que denota la mediana. Por lo tanto, para esas distribuciones los enfoques considerados presentan menores diferencias en lo que se refiere a la longitud de los intervalos de verosimilitud para periodos de retorno.

Finalmente, para las distribuciones en el DA Weibull, los cocientes presentan valores más cercanos a uno conforme disminuye el parámetro de forma (véanse las Figuras 3.13 y 3.14). Por lo tanto, al igual que en el caso Fréchet, las diferencias en términos de las longitudes de los intervalos de verosimilitud para niveles de retorno calculadas con cada enfoque, son mayores cuando consideramos distribuciones con parámetro de forma "cercano" a cero.



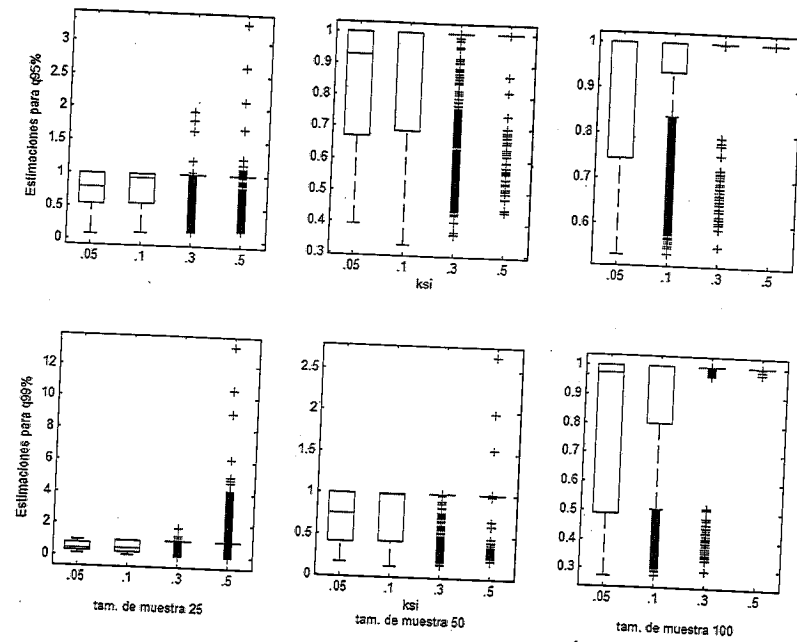


Figura 3.8: Box plots de los cocientes entre las longitudes de los intervalos de verosimilitud obtenidos con cada enfoque (DVE/DGVE) para las distribuciones de la familia Pareto.

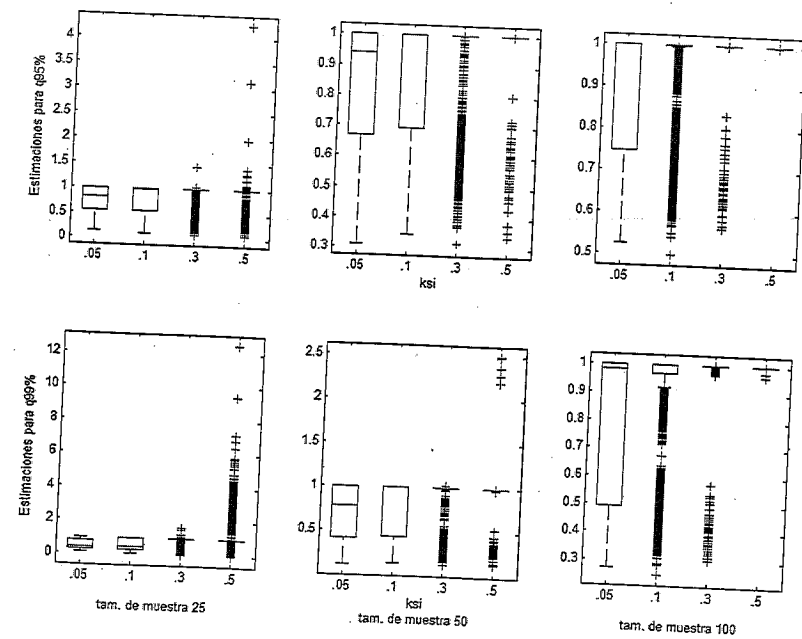


Figura 3.9: Box plots de los cocientes entre las longitudes de los intervalos de verosimilitud obtenidos con cada enfoque (DVE/DGVE) para las distribuciones de la familia Loggamma.

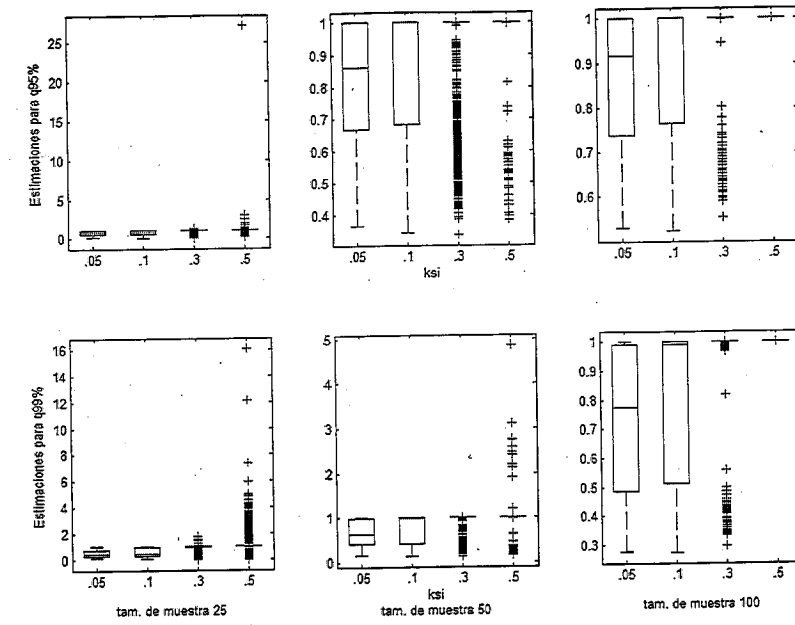


Figura 3.10: Box plots de los cocientes entre las longitudes de los intervalos de verosimilitud obtenidos con cada enfoque (DVE/DGVE) para las distribuciones de la familia Burr.

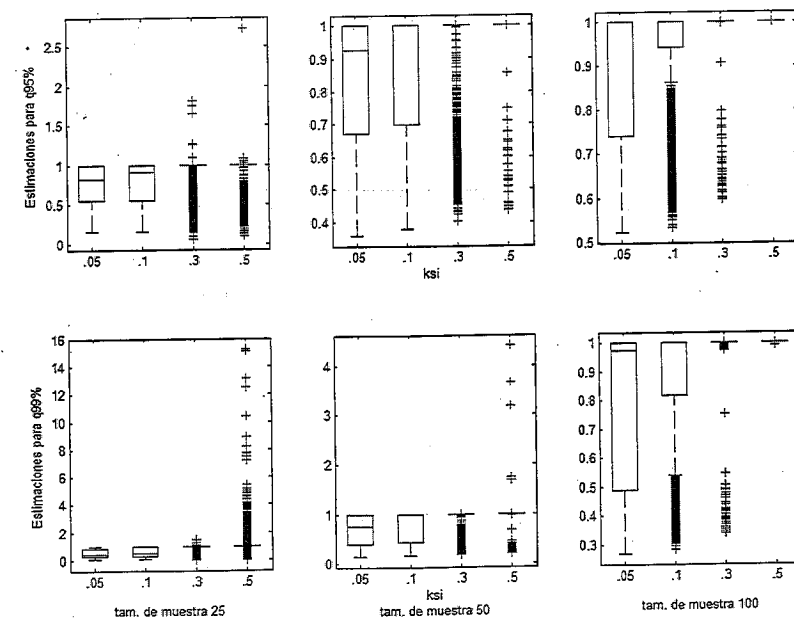


Figura 3.11: Box plots de los cocientes entre las longitudes de los intervalos de verosimilitud obtenidos con cada enfoque (DVE/DGVE) para las distribuciones de la familia Fréchet.

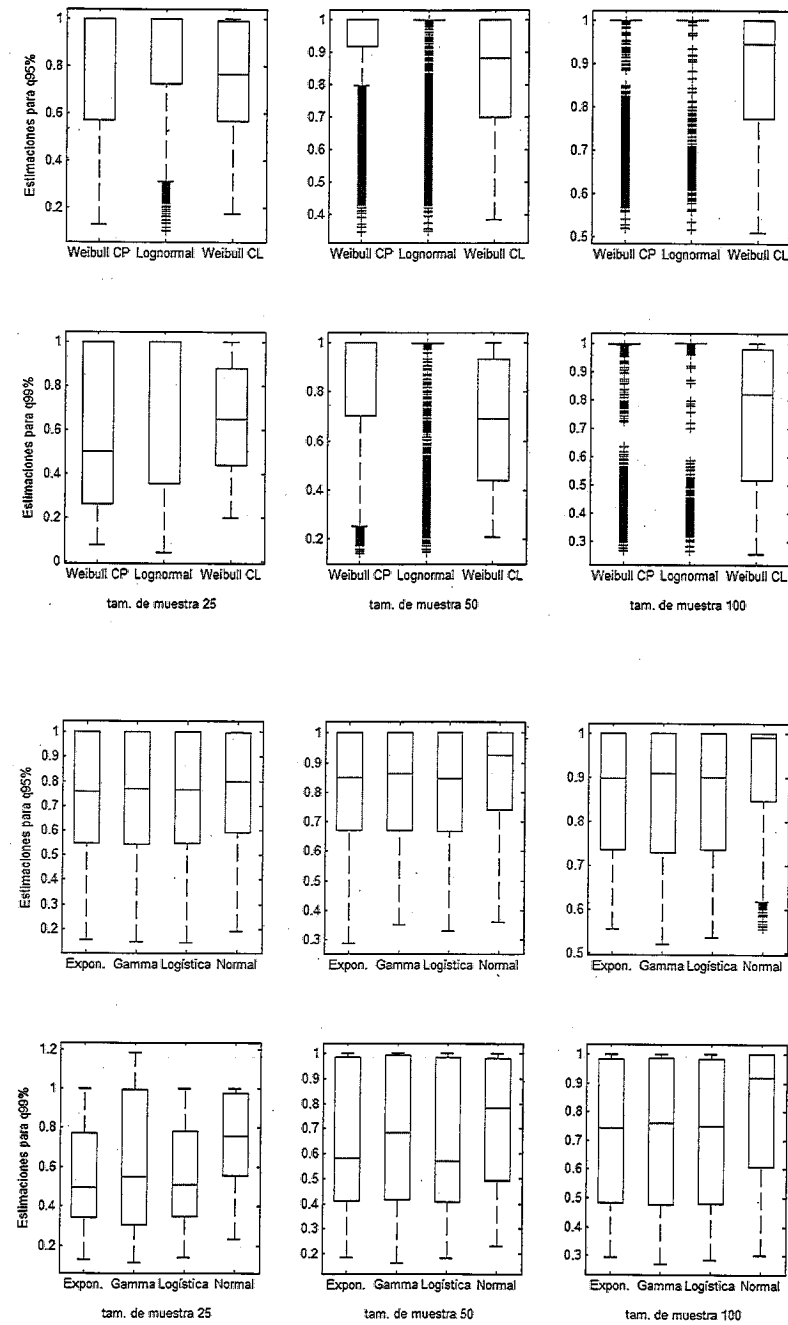


Figura 3.12: Box plots de los cocientes entre las longitudes de los intervalos de verosimilitud obtenidos con cada enfoque (DVE/DGVE) para las distribuciones con DA Gumbel.

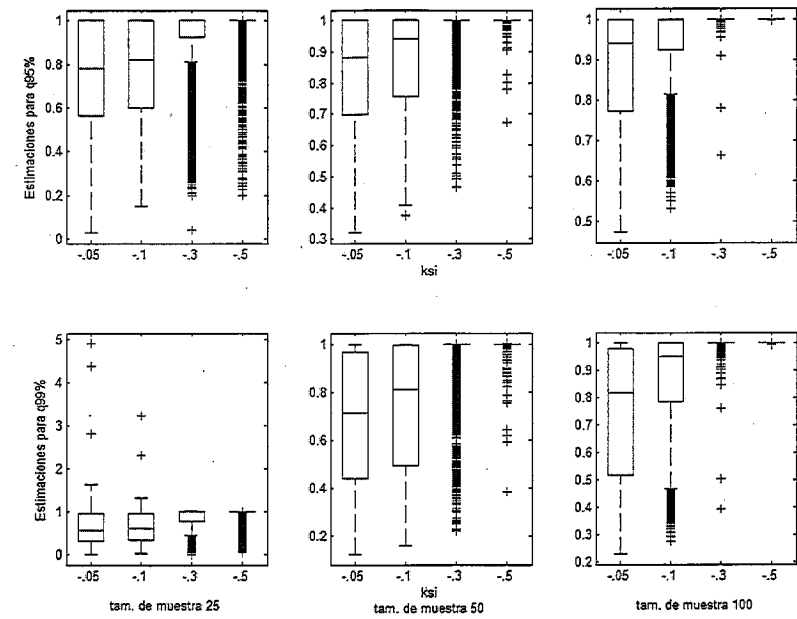


Figura 3.13: Box plots de los cocientes entre las longitudes de los intervalos de verosimilitud obtenidos con cada enfoque (DVE/DGVE) para las distribuciones de la familia Beta.

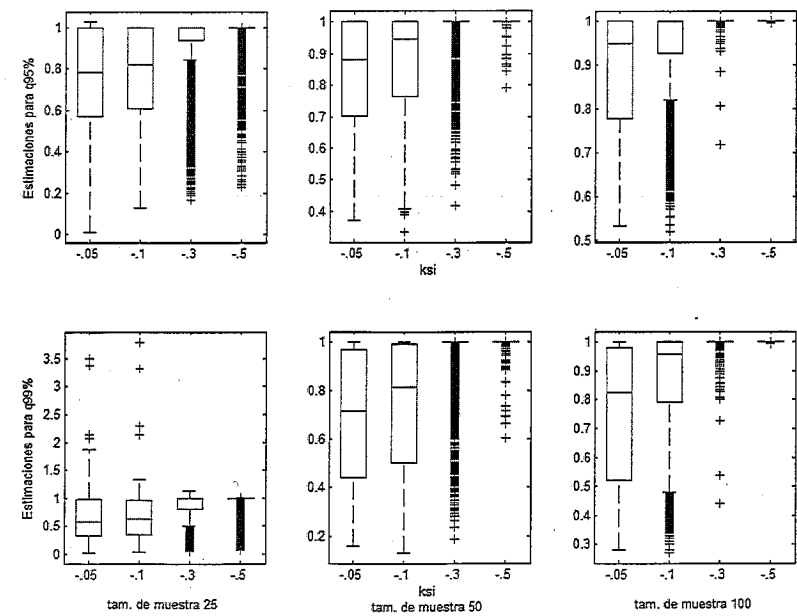


Figura 3.14: Box plots de los cocientes entre las longitudes de los intervalos de verosimilitud obtenidos con cada enfoque (DVE/DGVE) para las distribuciones de la familia Weibull (DVE).

### 3.2.3. Análisis del Extremo Derecho

Las funciones de verosimilitud perfil para los cuantiles pueden presentar asimetría, especialmente cuando se consideran muestras pequeñas. Específicamente, a la derecha del estimador de máxima verosimilitud del cuantil que estemos considerando, la función de verosimilitud perfil puede decaer muy lentamente a cero al compararse con lo que ocurre en el lado izquierdo. Por lo tanto, un análisis detallado del extremo derecho de los intervalos de verosimilitud obtenidos bajo cada enfoque, es valioso pues da información sobre qué tan anchos son dichos intervalos. A continuación explicamos el análisis que aplicamos.

Supongamos que, bajo las condiciones de alguno de los escenarios incluidos en este trabajo, hemos calculado, con los dos enfoques, los intervalos de verosimilitud que corresponden al cuantil del 99% ( $q99$ ). Para obtener una aproximación a la forma en que se comportó el extremo derecho de esos intervalos en términos de su cercanía al valor real  $q99$ , podemos calcular el porcentaje de los casos que el extremo derecho perteneció a alguno de los intervalos  $[q99, x_1), [x_1, x_2), \dots, [x_{n-1}, x_n)$  o  $[x_n, \infty)$ , donde  $x_1 \leq x_2 \leq \dots \leq x_n$  son números reales mayores que  $q99$ . Comparando esos porcentajes para los intervalos obtenidos con cada enfoque, podemos determinar si en algún caso obtuvimos estimaciones más precisas. Más aún, si usamos el razonamiento anterior calculando porcentajes pero para los intervalos  $[q99, x_1), [q99, x_2), \dots, [q99, x_n)$  y  $[q99, \infty)$ , entonces es más fácil apreciar cómo se ubicó el extremo derecho de los intervalos a la derecha de  $q99$ , pues podemos apreciar cómo se incrementan dichos porcentajes al aumentar la longitud de los intervalos  $[q99, x_i)$ .

Para implementar el análisis esbozado en el párrafo anterior es necesario determinar para cada distribución valores "adecuados" para los reales  $x_1 \leq x_2 \leq \dots \leq x_n$ , los cuales también varían según el cuantil que estemos estimando. Para ello, consideraremos cuantiles altos de la distribución de máximos que corresponda según el escenario en el que estemos trabajando. Un aspecto que será común a todos ellos es que tomaremos  $n = 3$ . Para las distribuciones incluidas en este estudio que pertenecen al DA Fréchet o Gumbel tenemos dos casos. Cuando consideramos intervalos que estiman el cuantil del 99%, definimos  $x_1 = q99.6$ ,  $x_2 = q99.9$  y  $x_3 = q99.99$ , donde  $q99.6$  se refiere al cuantil del 99.6%,  $q99.9$  al cuantil del 99.9% y  $q99.99$  corresponde al cuantil de 99.99%. El segundo caso ocurre cuando las estimaciones se refieren al cuantil del 95% ( $q95$ ), y usando la notación anterior, definimos  $x_1 = q98$ ,  $x_2 = q99.5$  y  $x_3 = q99.95$ . Por otro lado, para las distribuciones que pertenecen al DA Weibull tomamos cuantiles más altos que en el caso anterior. Para las estimaciones del cuantil del 95% definimos  $x_1 = q98$ ,  $x_2 = q99.9$  y  $x_3 = 99.999$ , mientras que para las que se

refieren al  $q99$ , consideraremos  $x_1 = q99.6$ ,  $x_2 = q99.98$  y  $x_3 = 99.9998$ . Una vez definidos para cada caso los valores que toman  $x_1, x_2$  y  $x_3$ , calculamos los porcentajes, como se explicó en el párrafo anterior, para los intervalos  $[y, x_1), [y, x_2), [y, x_3)$  y  $[y, \infty)$ , donde  $y = q99$  o  $q95$  según las estimaciones que estemos considerando. A continuación comentamos los resultados obtenidos.

Para las distribuciones del dominio de atracción Fréchet consideremos primero las estimaciones que se refieren al cuantil de 99% con los diferentes tamaños de muestra incluidos en el estudio. Las gráficas correspondientes aparecen de la Figura 3.15 a la 3.26. Para las distribuciones con  $\xi$  igual a .05 o .1 obtuvimos que los intervalos estimados con el enfoque basado en las *DVE*, tienen asociados porcentajes más altos en los intervalos  $[q99, x_i)$  más angostos, pero esa "ventaja" con respecto a los intervalos obtenidos con el otro enfoque se reduce conforme aumenta el tamaño de la muestra. Por otro lado, para las distribuciones con  $\xi = .3$  o .5 y considerando un tamaño de muestra igual a 25, obtuvimos porcentajes mayores para los intervalos obtenidos usando el enfoque basado en la *DGVE*. Para esos valores de  $\xi$  pero con los otros tamaños de muestra, apreciamos una disminución de la diferencia entre los porcentajes que corresponden a los intervalos calculados bajo cada enfoque. Finalmente, para las estimaciones que se refieren al cuantil de 95%, en general son válidas las observaciones anteriores, aunque con menores diferencias entre los porcentajes asociados a los diferentes enfoques. Las gráficas para este caso aparecen de la Figura 3.27 a la 3.38

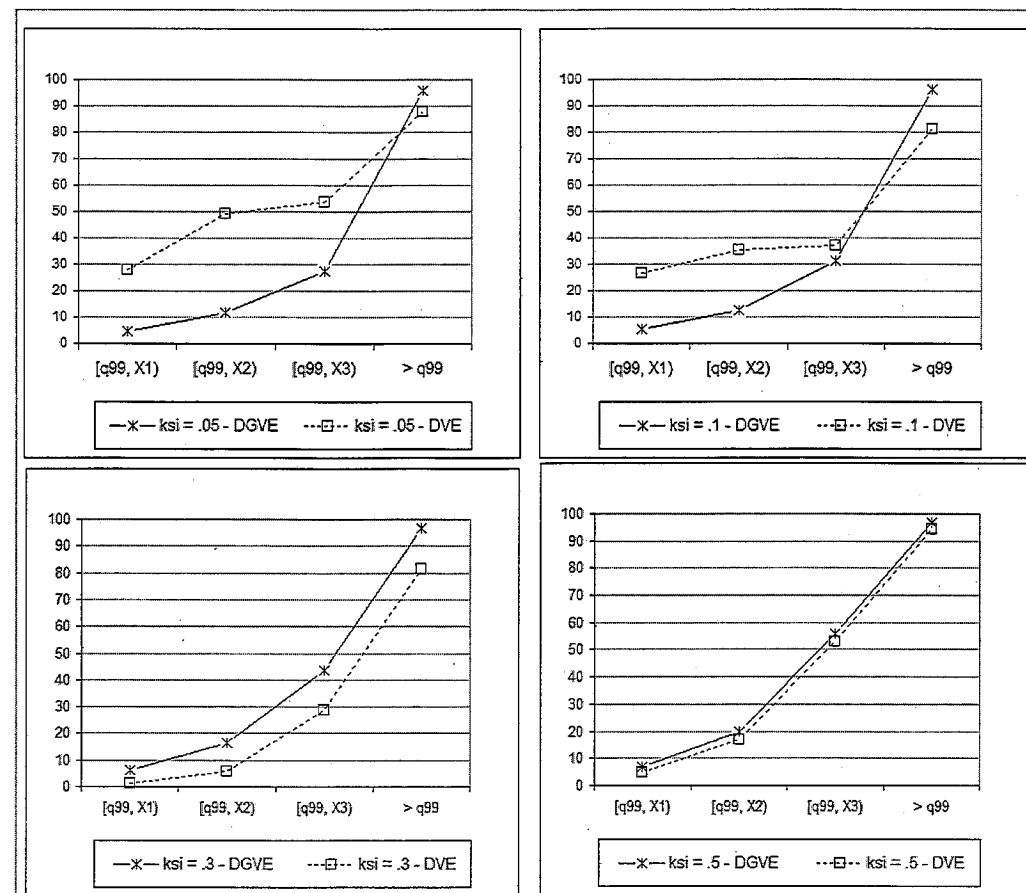


Figura 3.15: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Pareto. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 25.

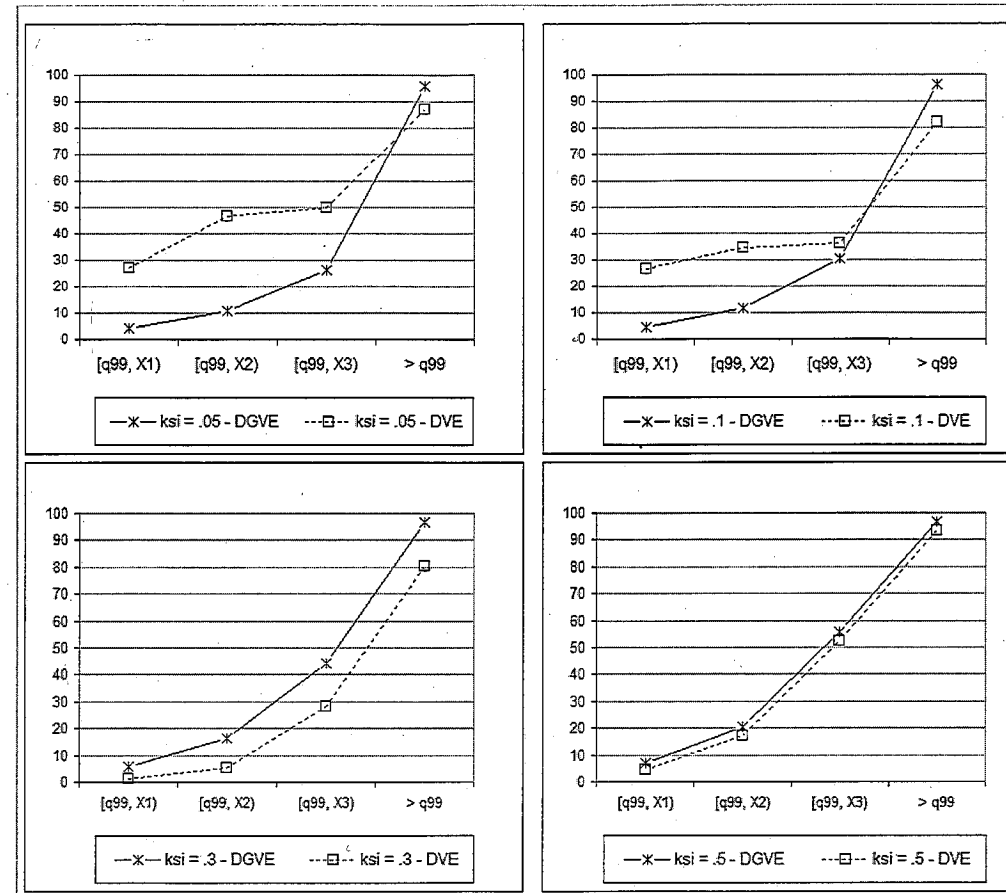


Figura 3.16: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Loggamma. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 25.

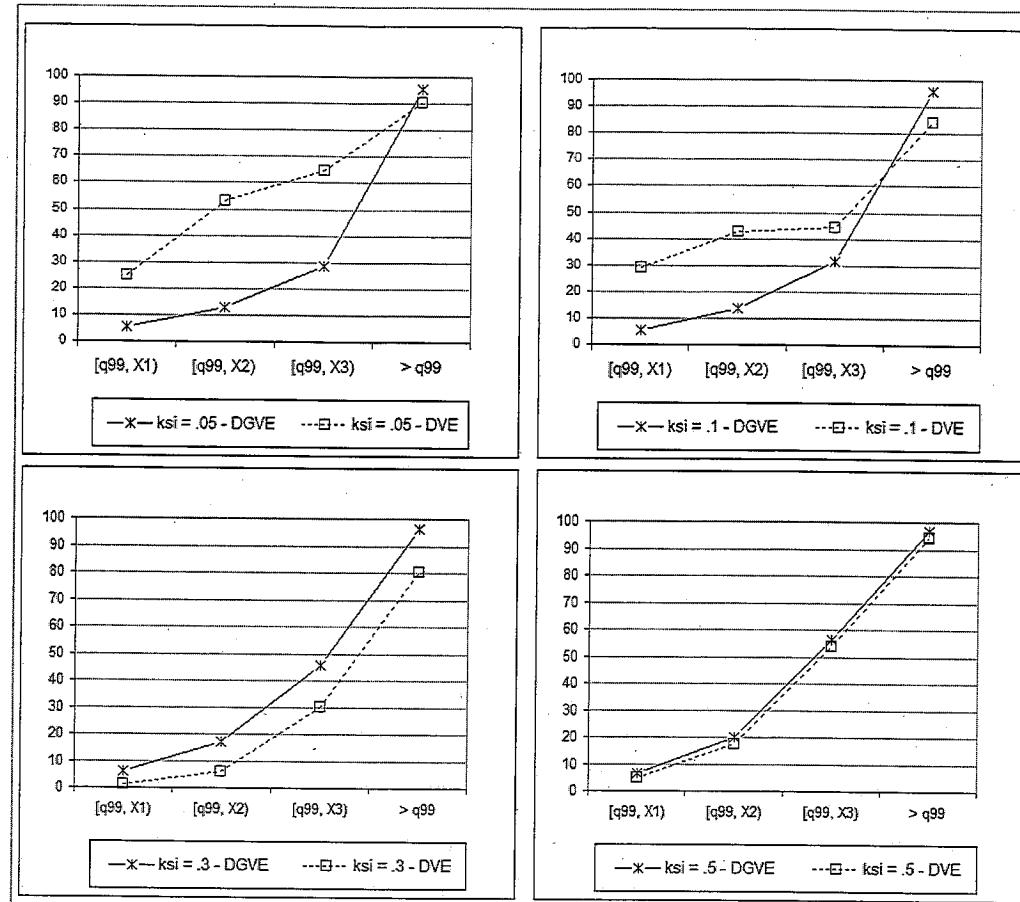


Figura 3.17: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Burr. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 25.

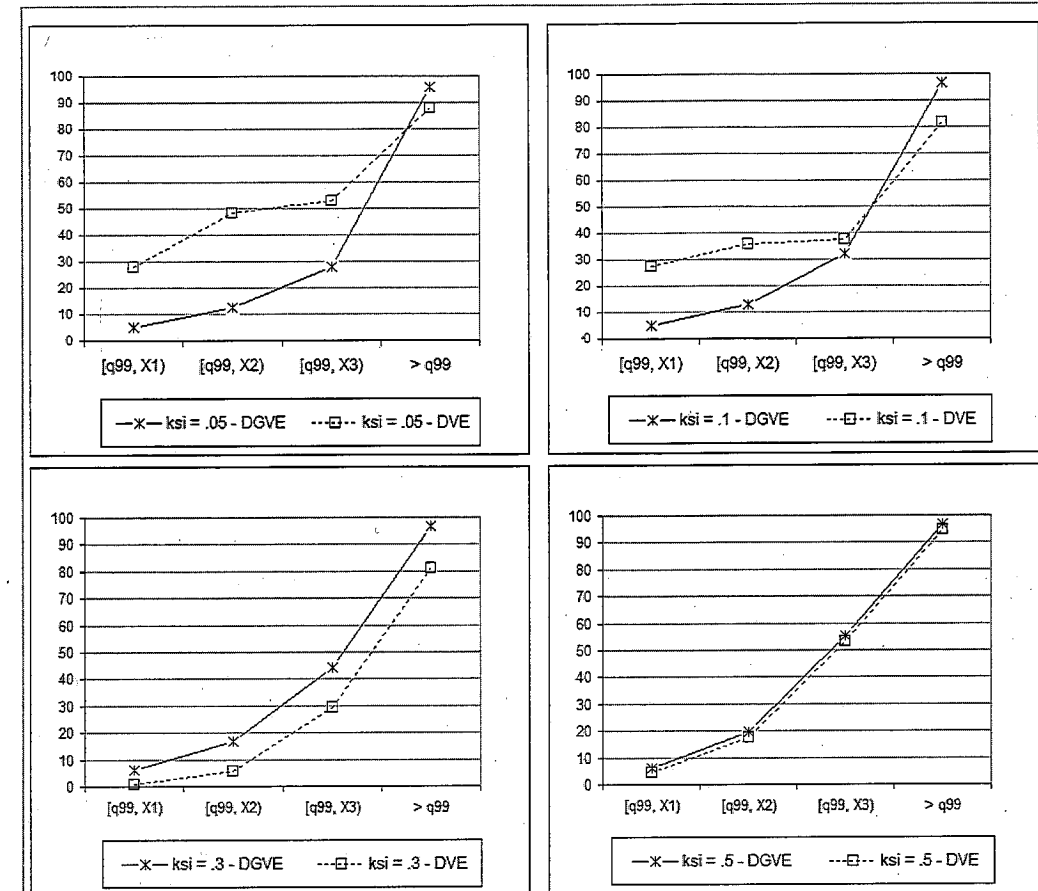


Figura 3.18: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Fréchet. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 25.

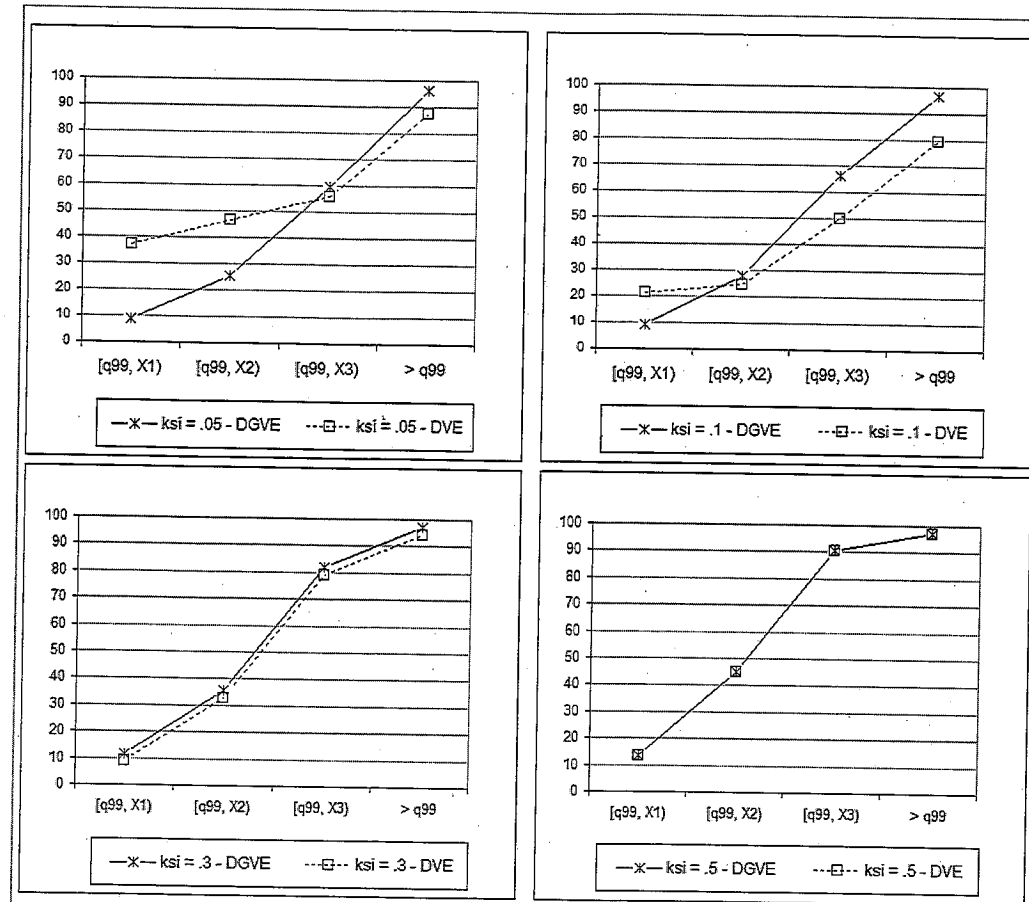


Figura 3.19: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Pareto. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 50.

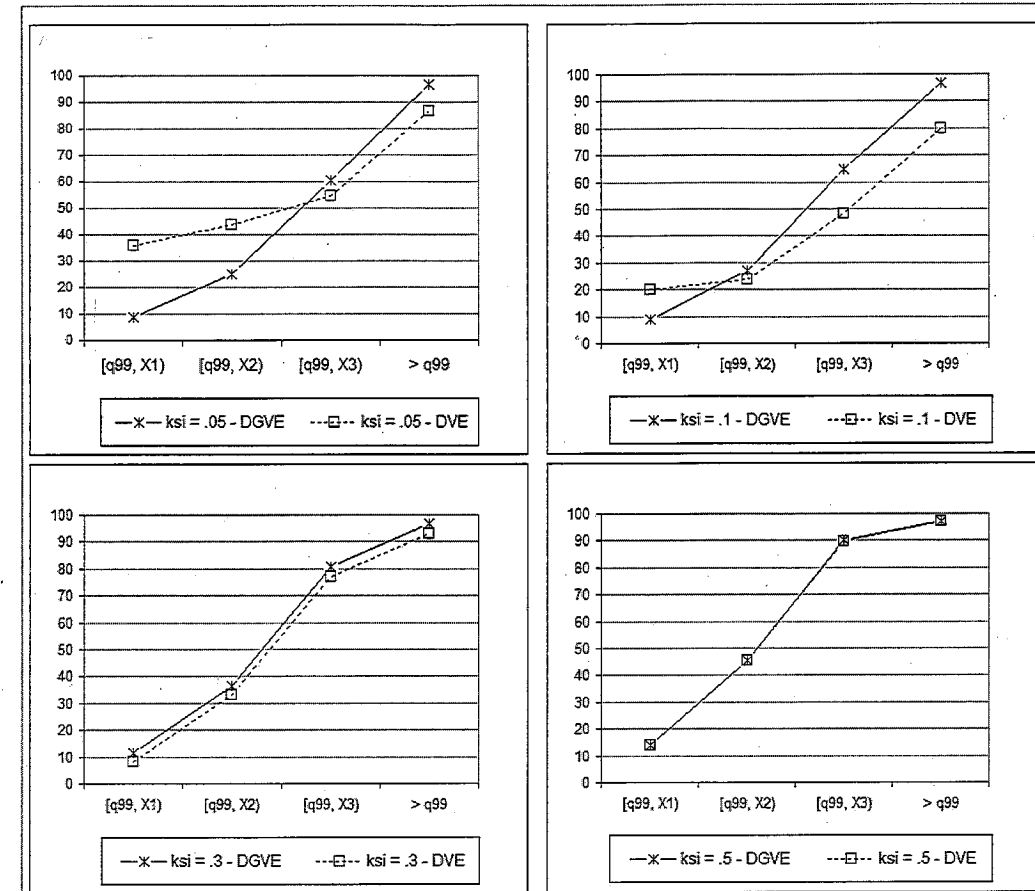


Figura 3.20: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Loggamma. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 50.

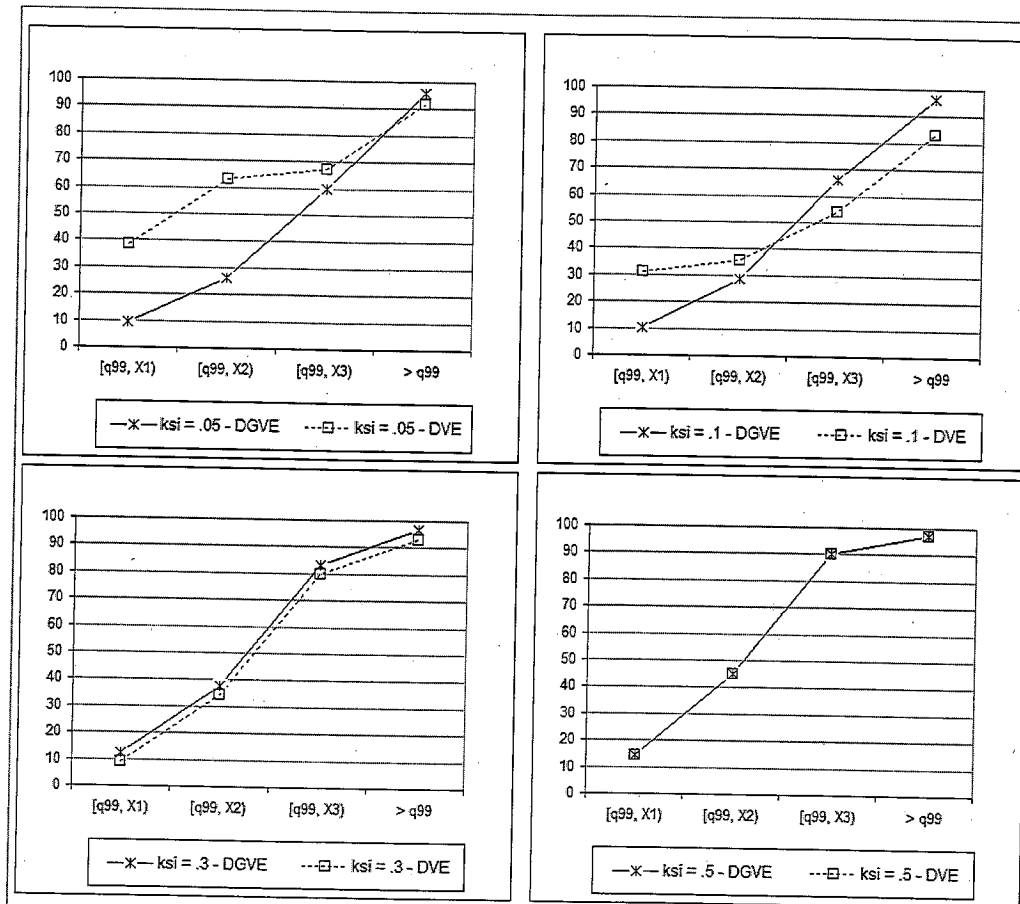


Figura 3.21: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Burr. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 50.

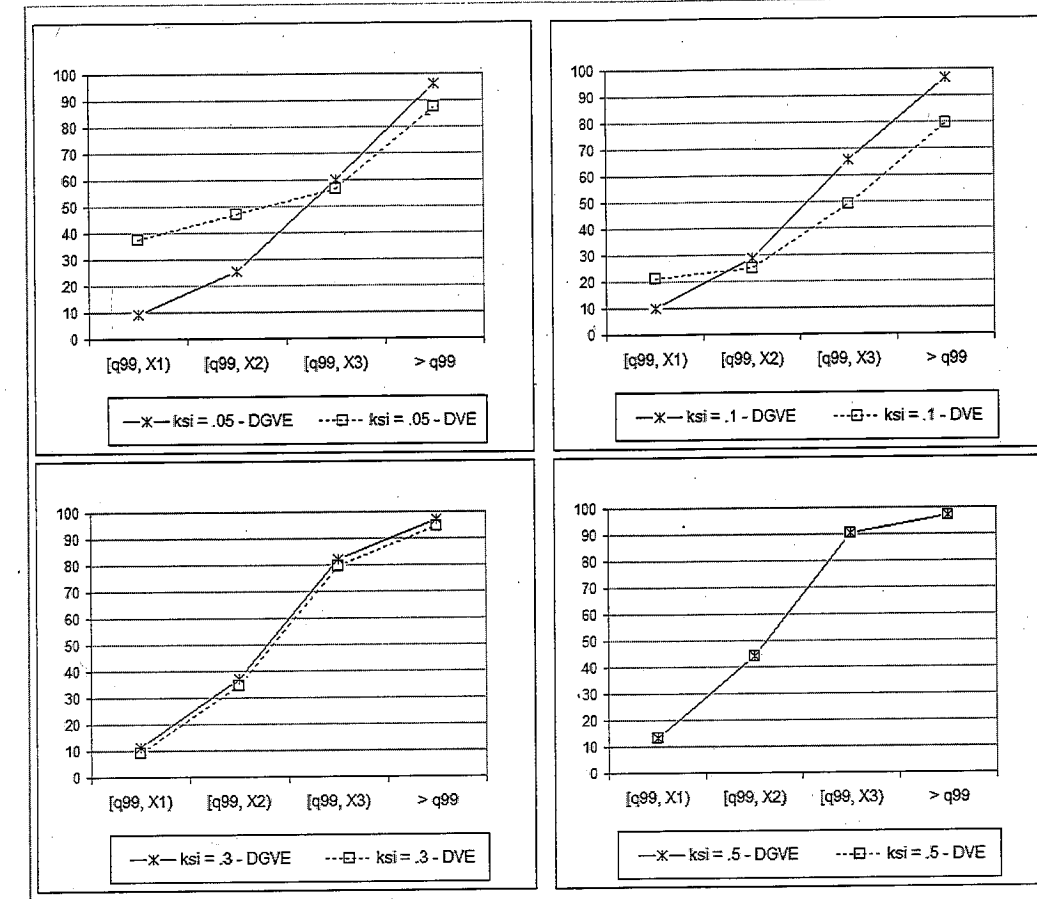


Figura 3.22: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Fréchet. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 50.

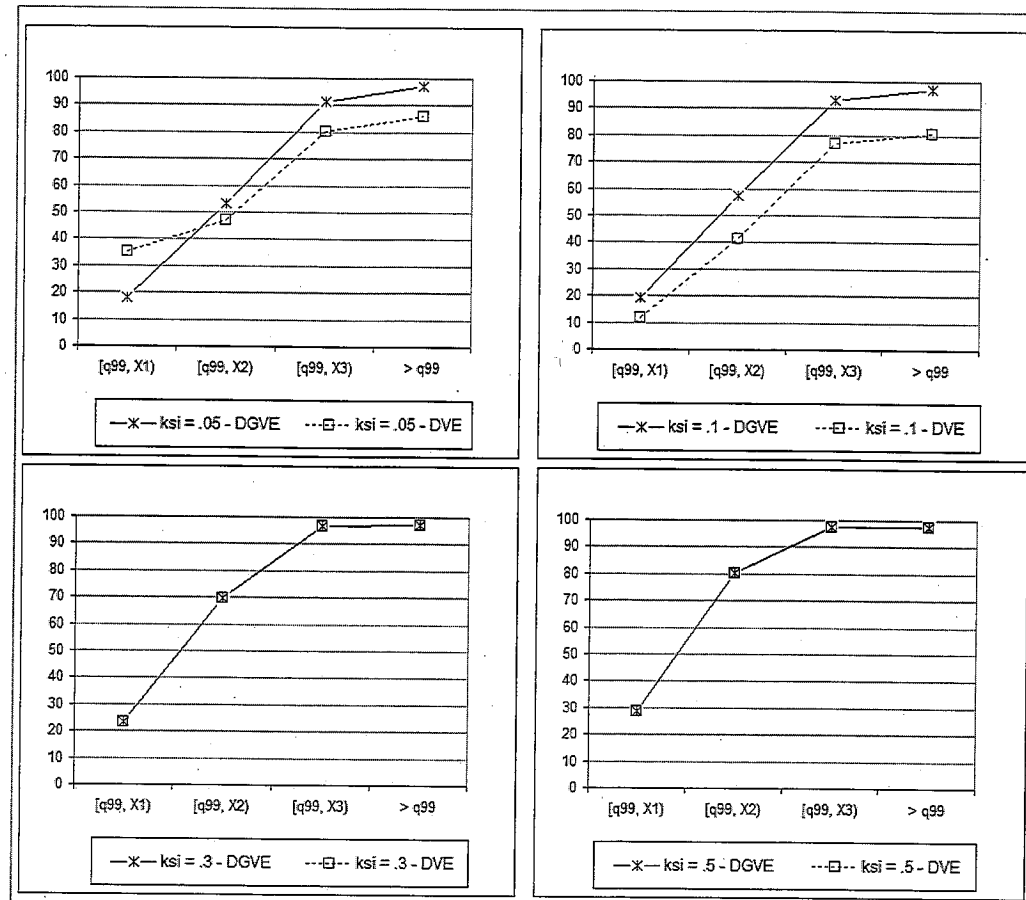


Figura 3.23: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Pareto. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 100.

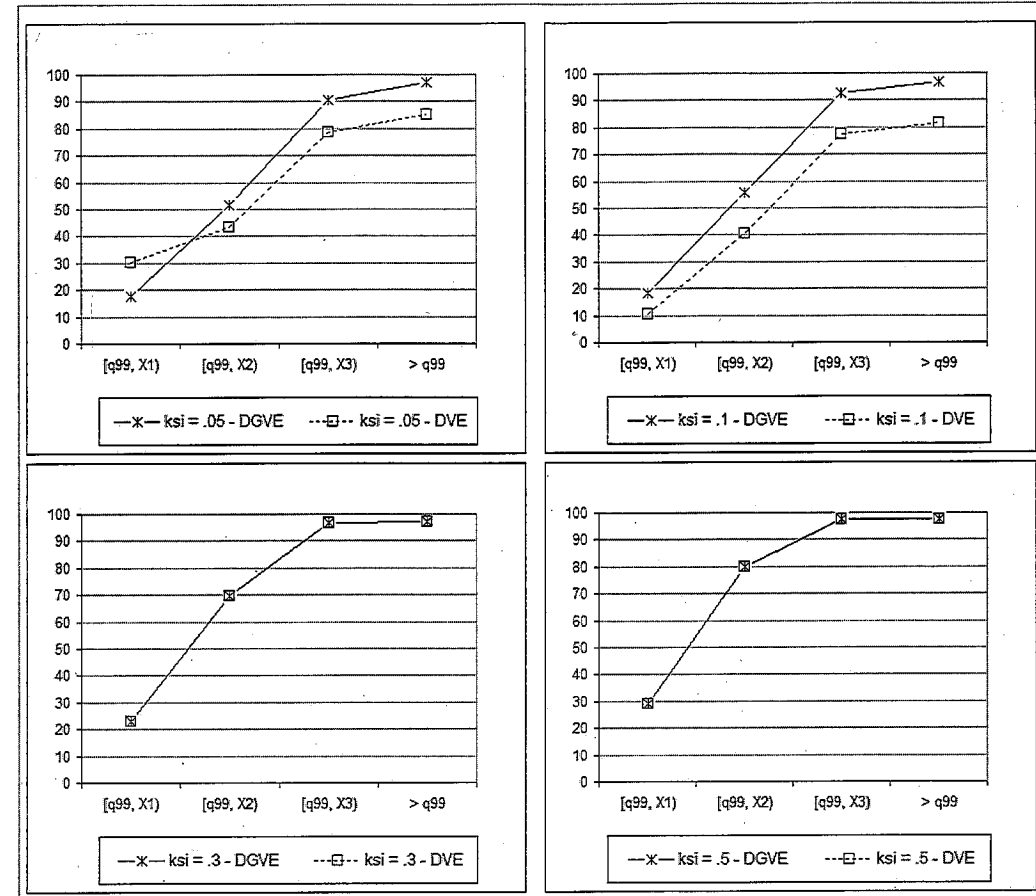


Figura 3.24: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Loggamma. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 100.



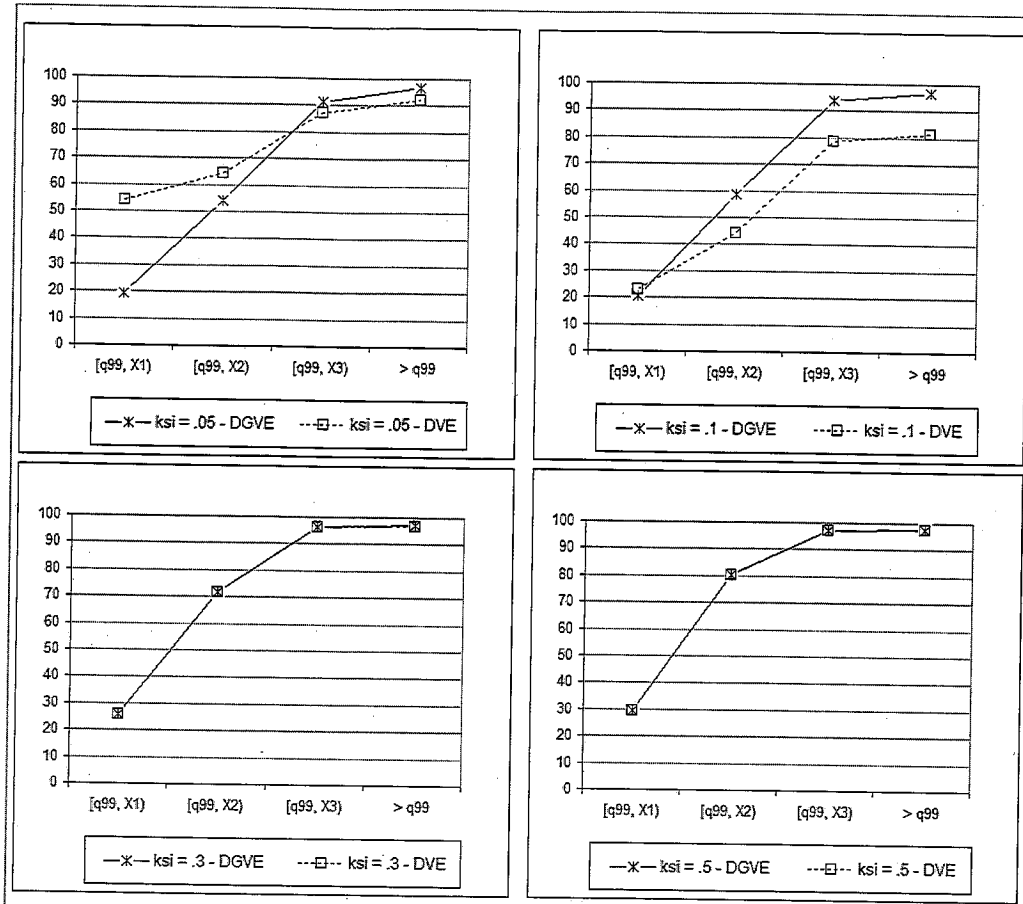


Figura 3.25: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Burr. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 100.

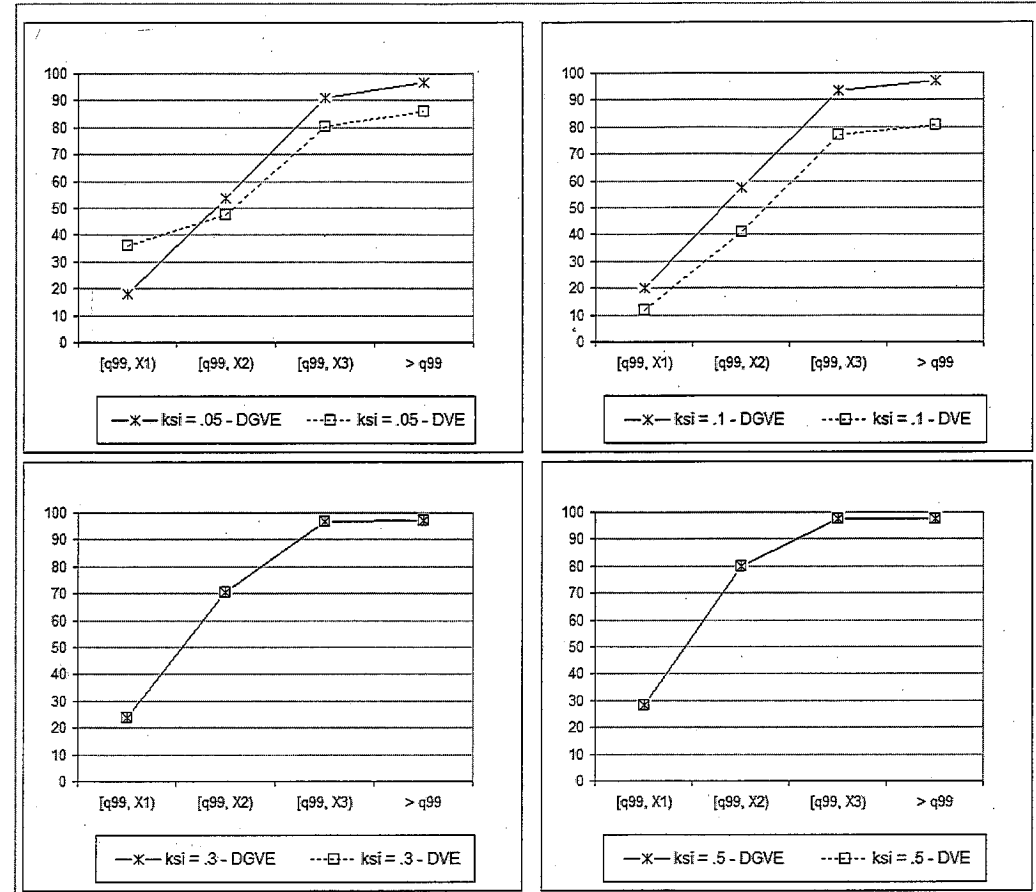


Figura 3.26: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Fréchet. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 100.

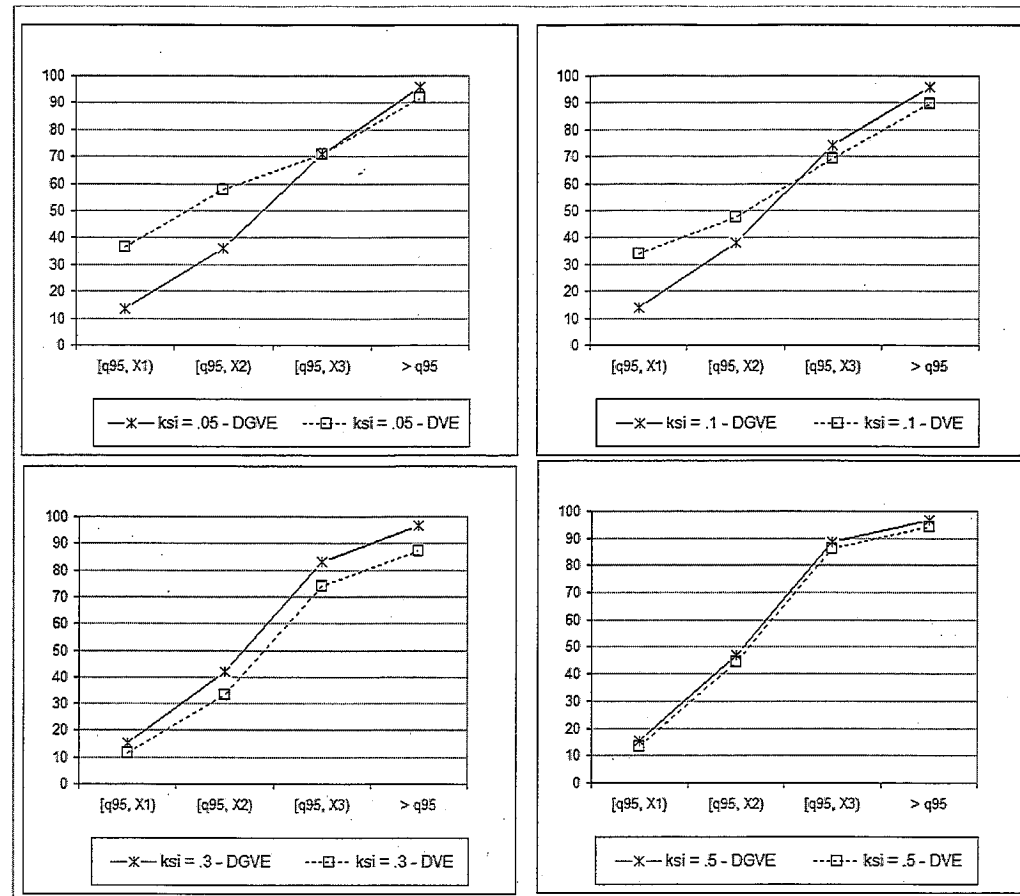


Figura 3.27: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Pareto. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 25.

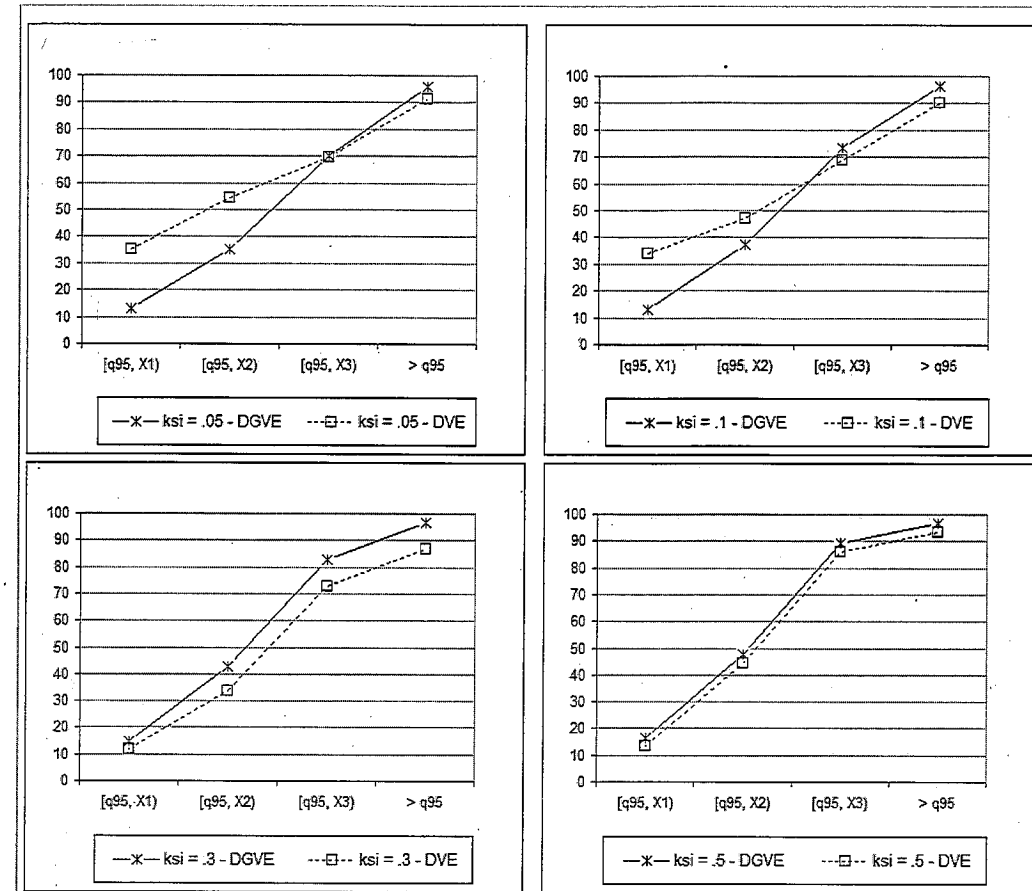


Figura 3.28: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Loggamma. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 25.

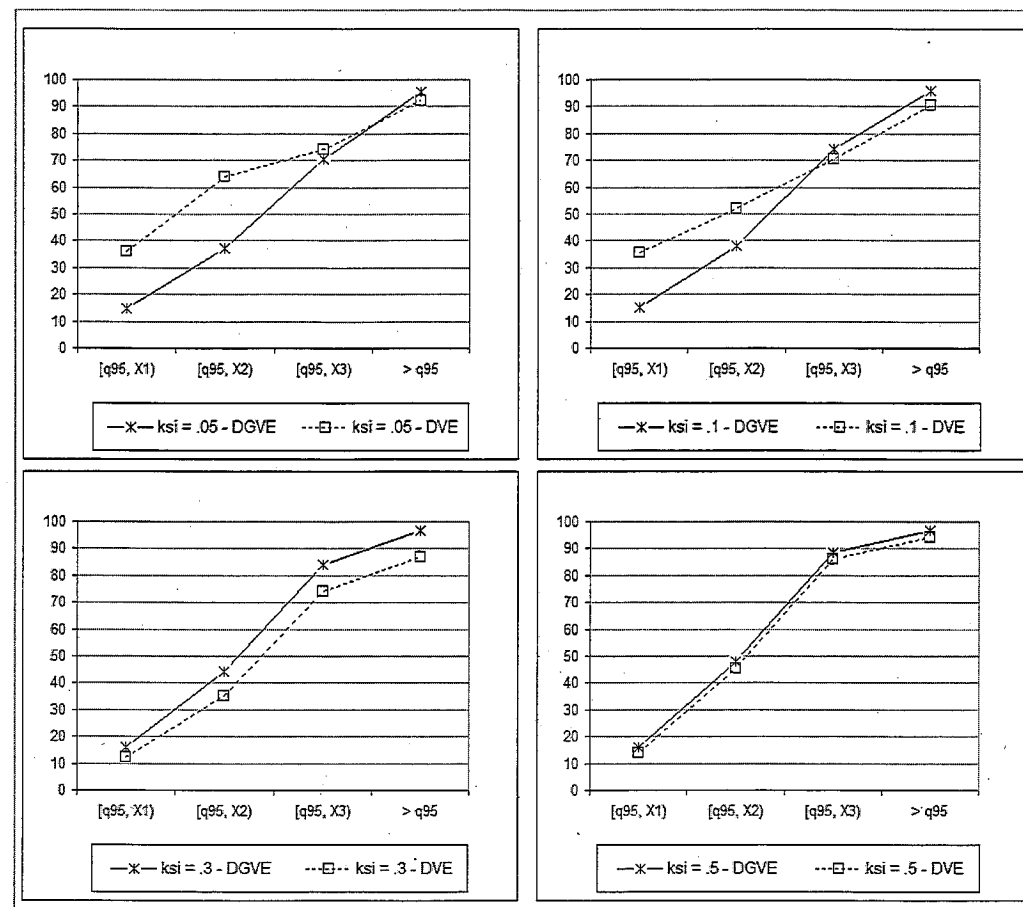


Figura 3.29: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Burr. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 25.

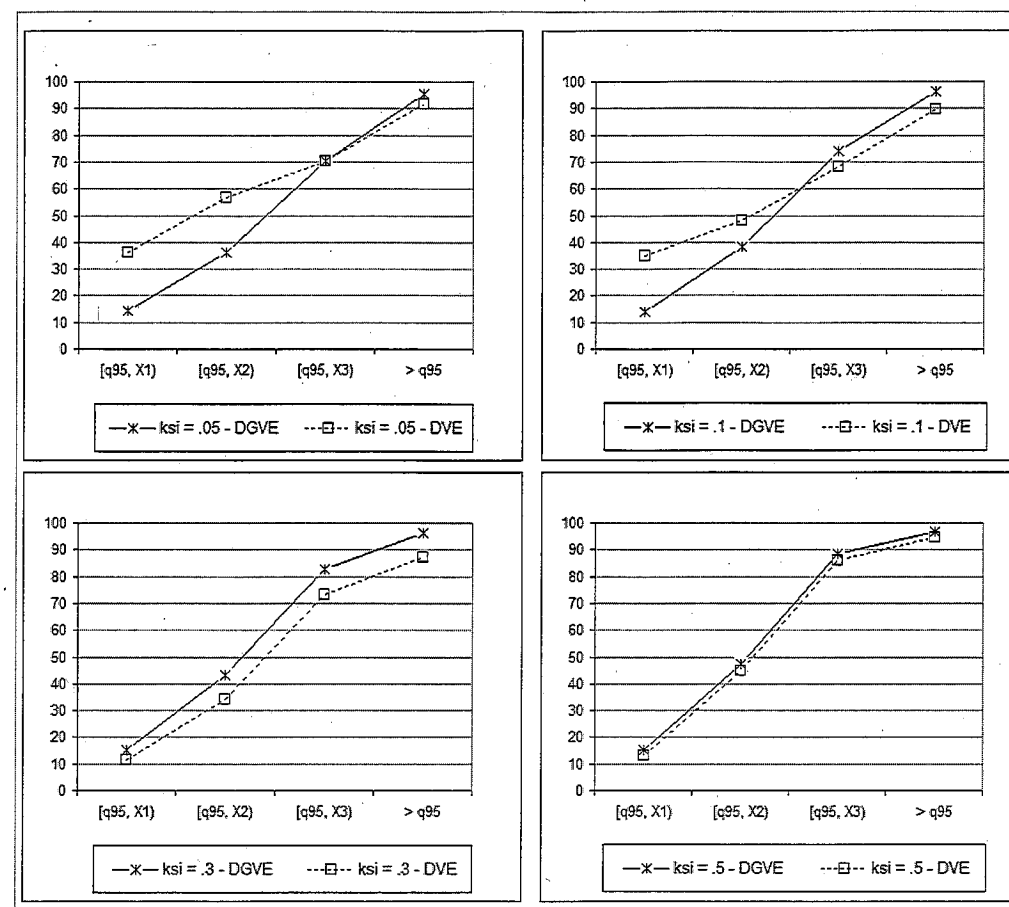


Figura 3.30: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Fréchet. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 25.

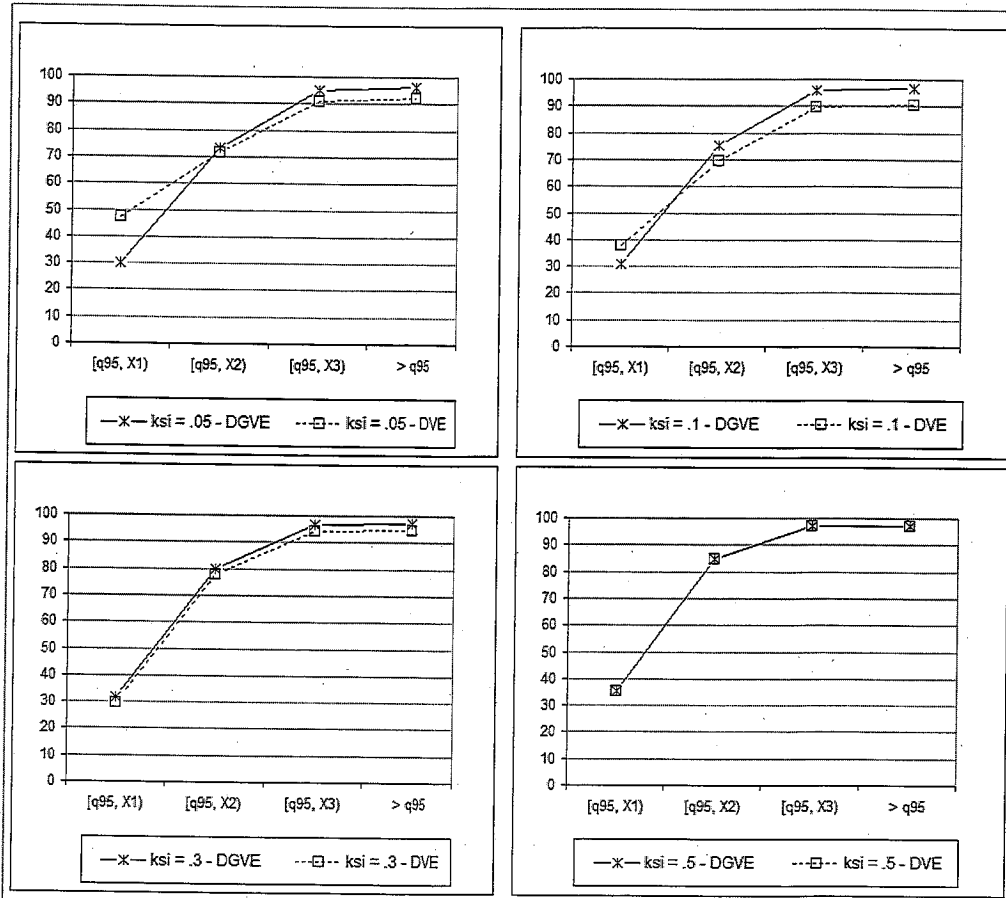


Figura 3.31: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Pareto. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 50.

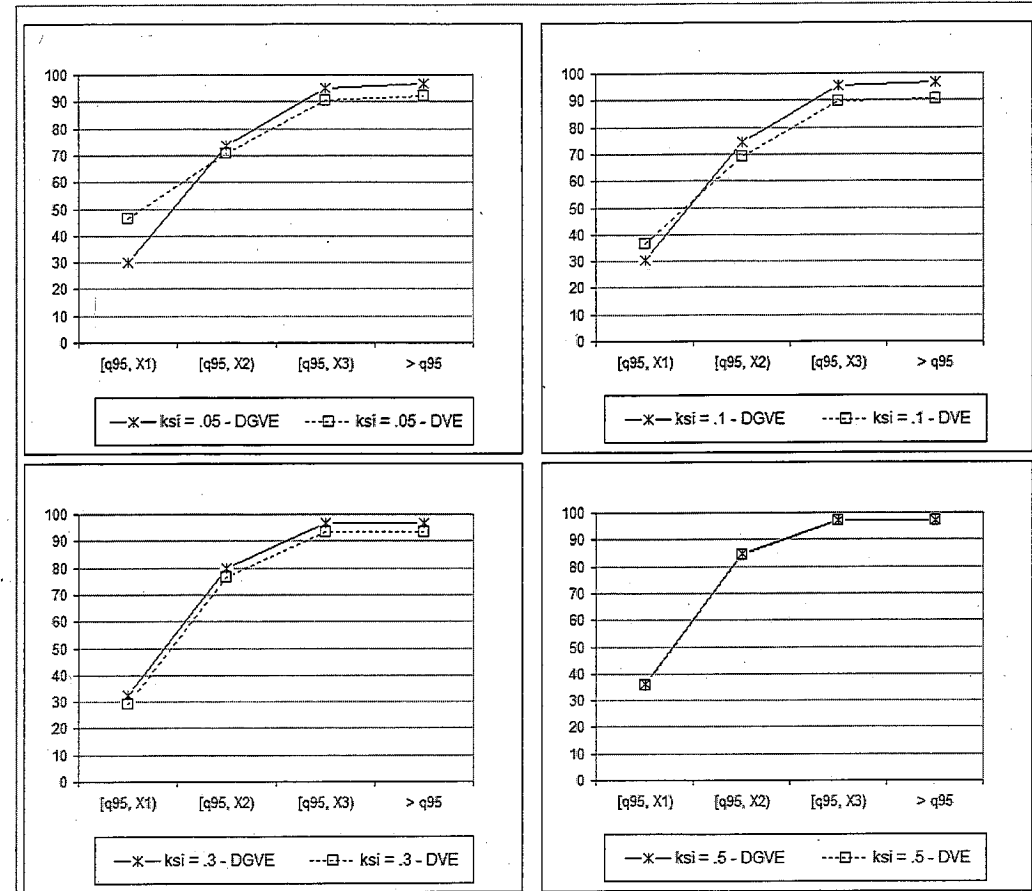


Figura 3.32: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Loggamma. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 50.

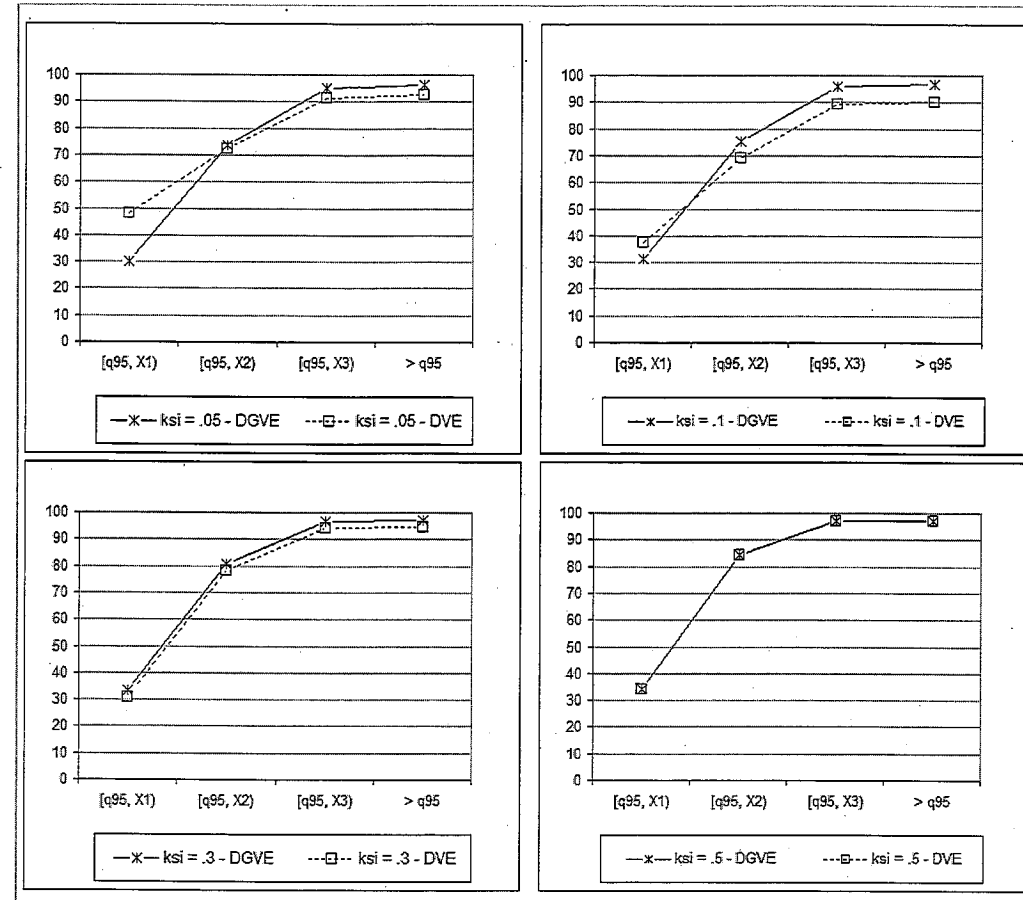


Figura 3.33: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Burr. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 50.

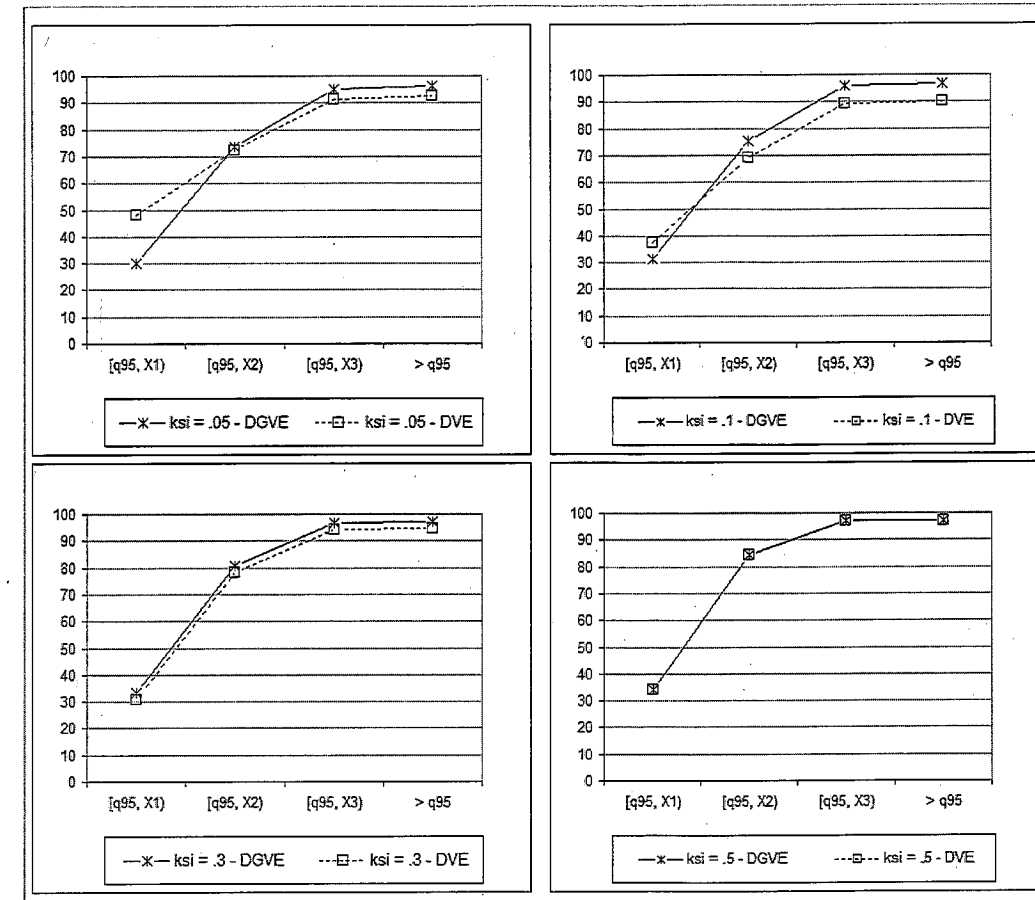


Figura 3.34: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Fréchet. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 50.

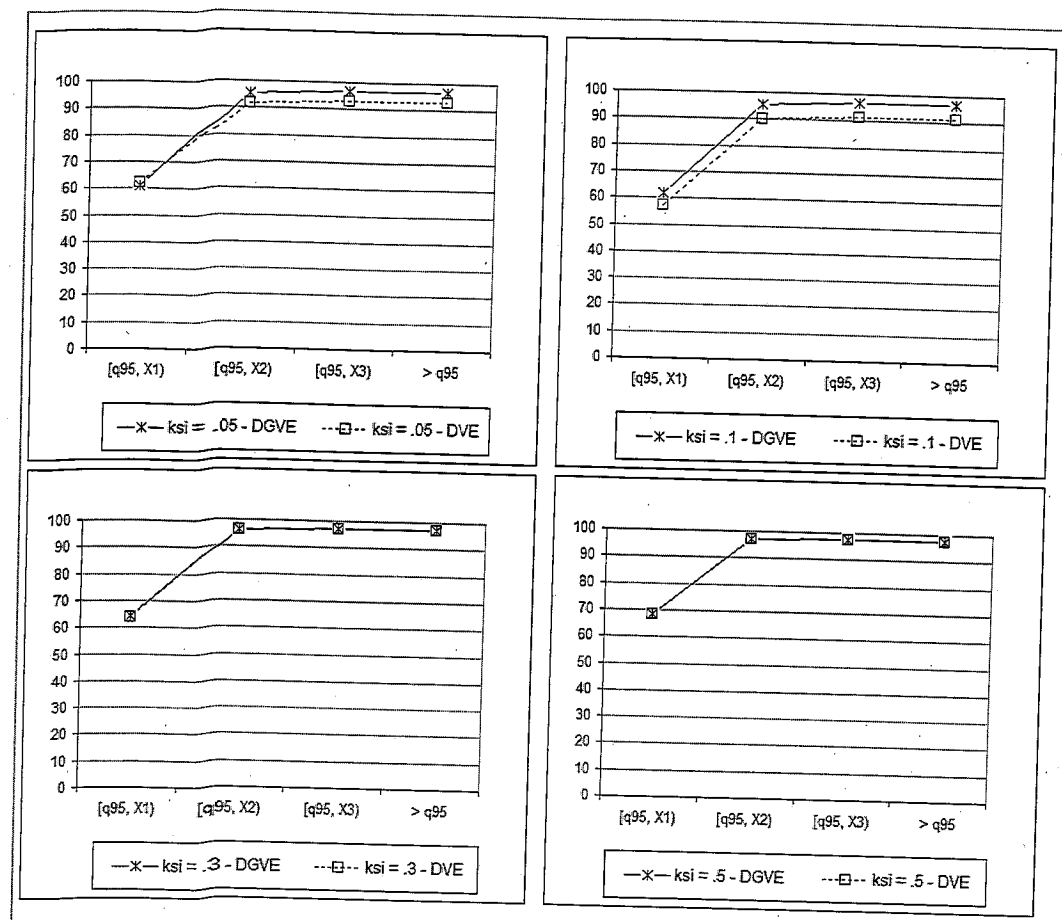


Figura 3.35: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Pareto. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 100.

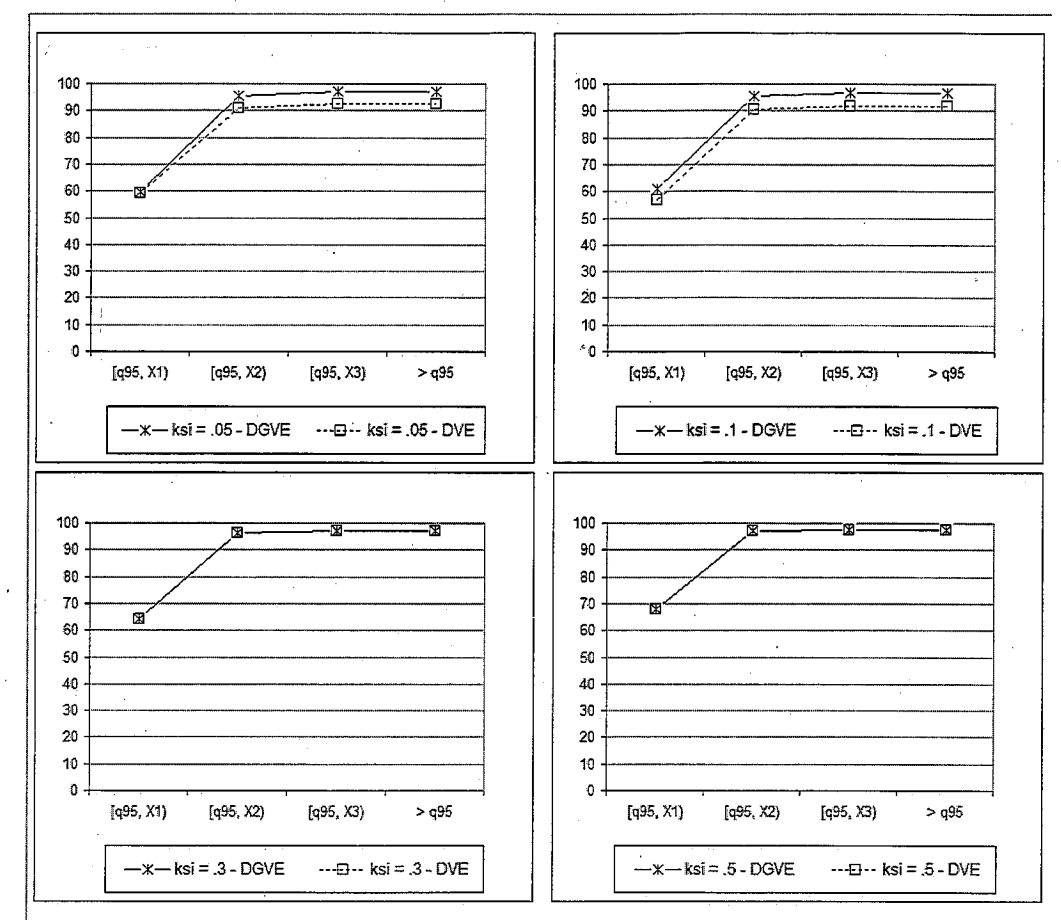


Figura 3.36: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Loggamma. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 100.

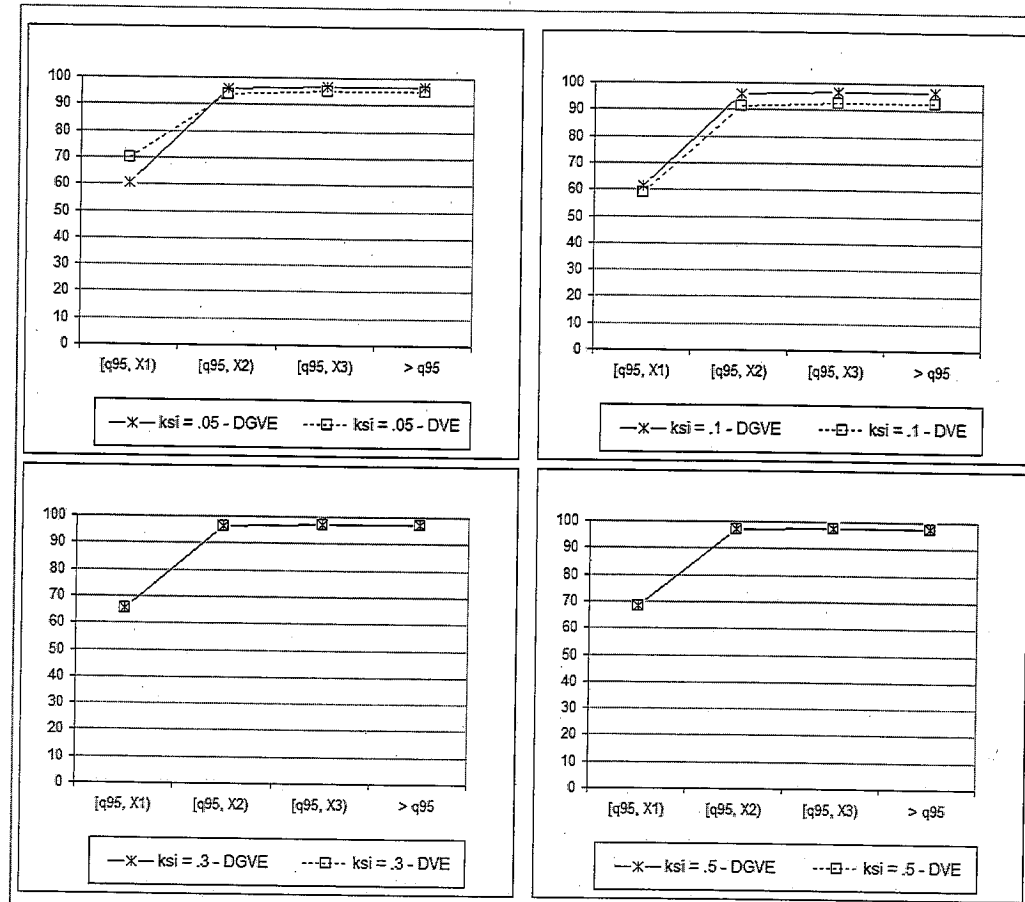


Figura 3.37: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Burr. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 100.

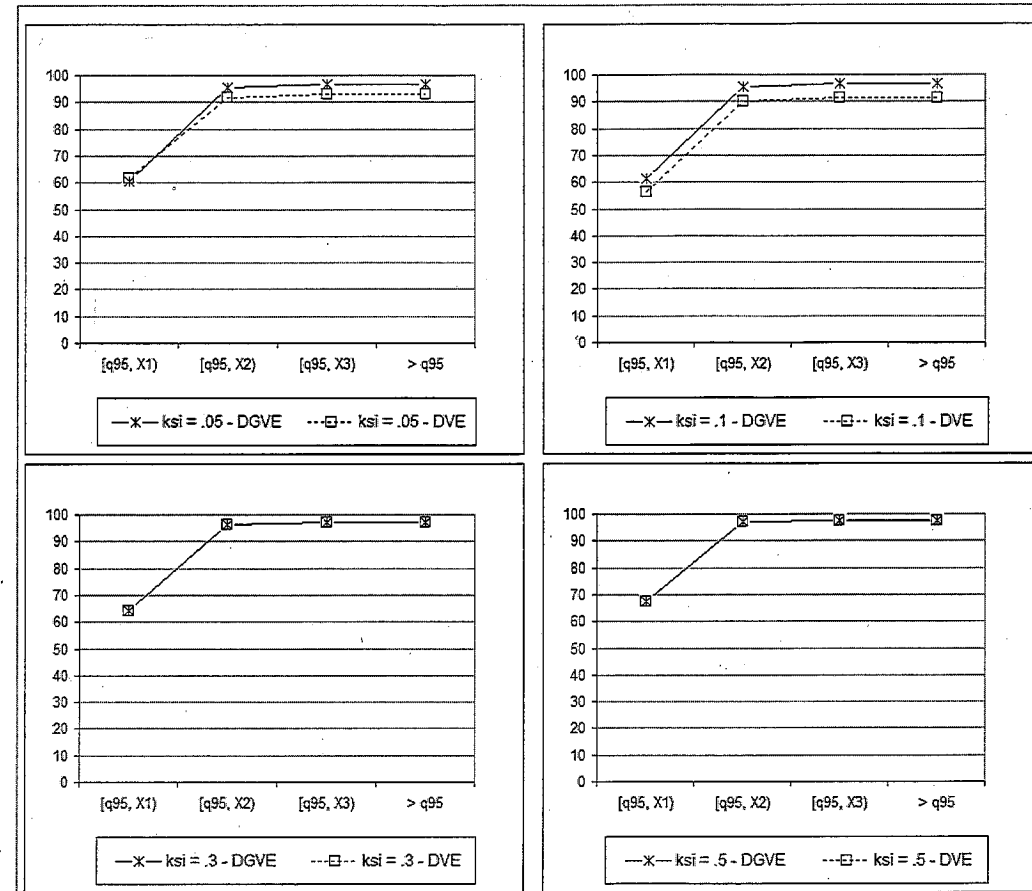


Figura 3.38: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Fréchet. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 100.

De la Figura 3.39 a la 3.44 aparecen los resultados para el extremo derecho de los intervalos obtenidos con las distribuciones que pertenecen al DA Weibull y que estiman el cuantil de 99%. Se puede apreciar que existen dos grupos que coinciden con los encontrados al analizar la cobertura o el cociente de las longitudes de los intervalos. Por un lado, los resultados para las distribuciones Weibull C.P y Lognormal muestran que, en la mayoría de los casos, son mayores los porcentajes asociados al extremo derecho de los intervalos calculados usando la *DGVE*. Y por el otro, para el resto de las distribuciones consideradas ocurre lo contrario, es decir, los porcentajes más altos pertenecen a las estimaciones obtenidas con el enfoque basado en las *DVE*. El comportamiento descrito también se aplica a los intervalos de verosimilitud para el cuantil de 95%, aunque con diferencias menores entre los porcentajes asociados a las estimaciones obtenidas con cada enfoque. Las gráficas para ese caso las presentamos de la Figuras 3.45 a la 3.50.

Finalmente, para las distribuciones que pertenecen al DA Weibull, cuya información aparece de la Figura 3.51 a la 3.62, los resultados obtenidos considerando las estimaciones que corresponden al cuantil de 99%, varían cuando se refieren a distribuciones con diferente parámetro de forma. Con  $\xi = -.05$  o  $-.1$  obtuvimos, con cualquiera de los tamaños de muestra considerados, porcentajes más elevados para las estimaciones obtenidas con el enfoque basado en las *DVE*. En contraste, para las distribuciones con  $\xi = -.3$  o  $-.5$ , los porcentajes fueron muy similares. El comportamiento anterior también apareció para las estimaciones que se refieren al cuantil de 95%, aunque en este caso fueron menores las diferencias entre los porcentajes.

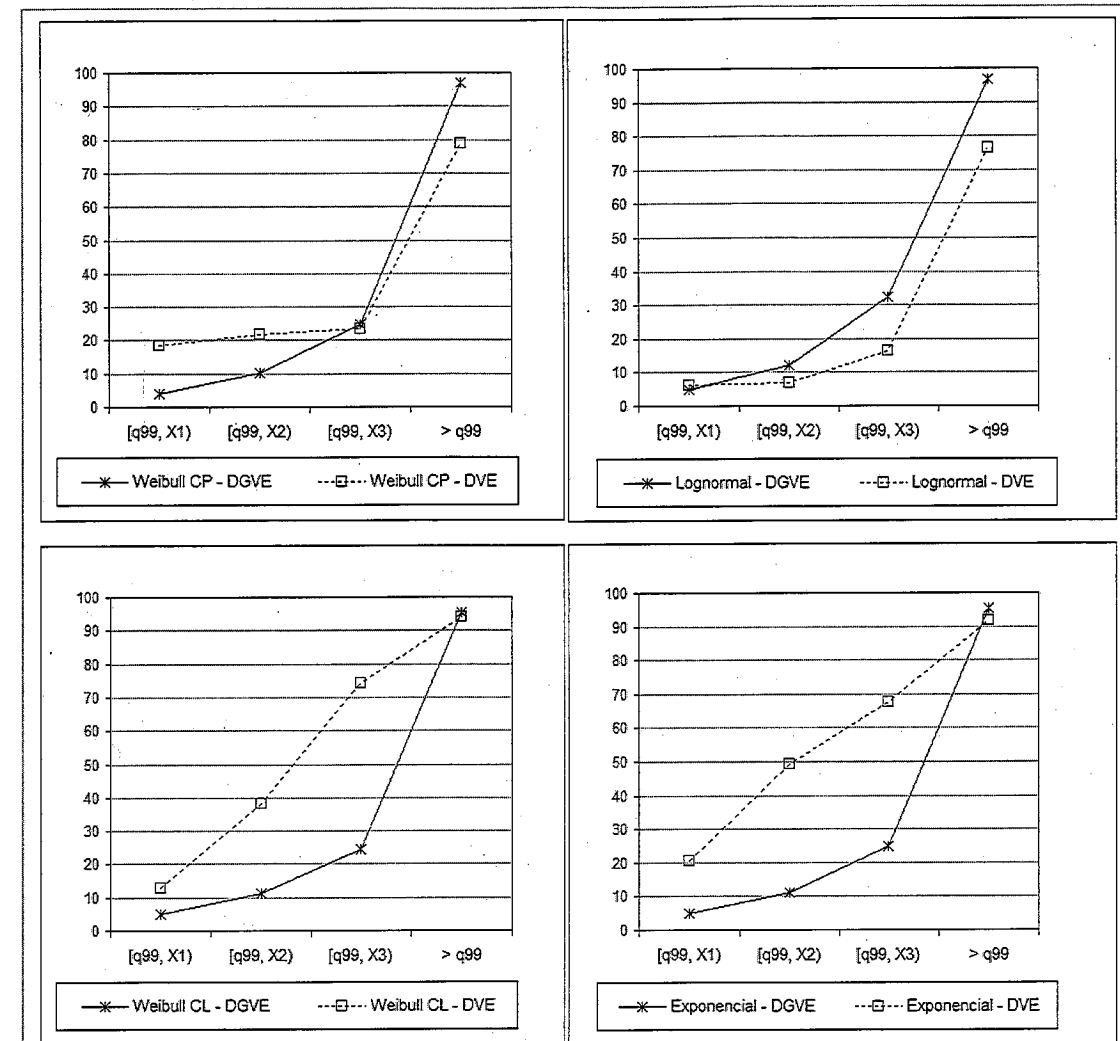


Figura 3.39: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 25.



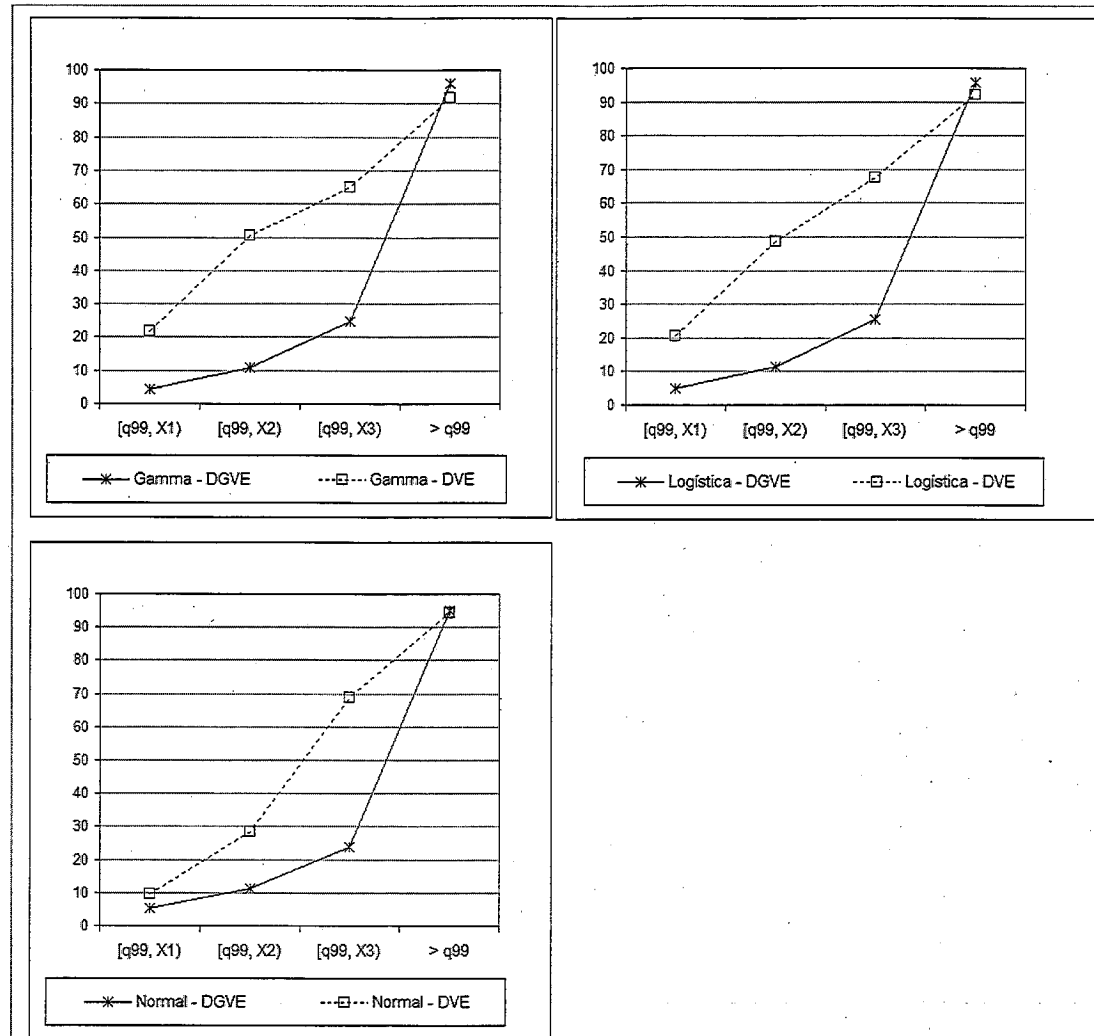


Figura 3.40: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 25.

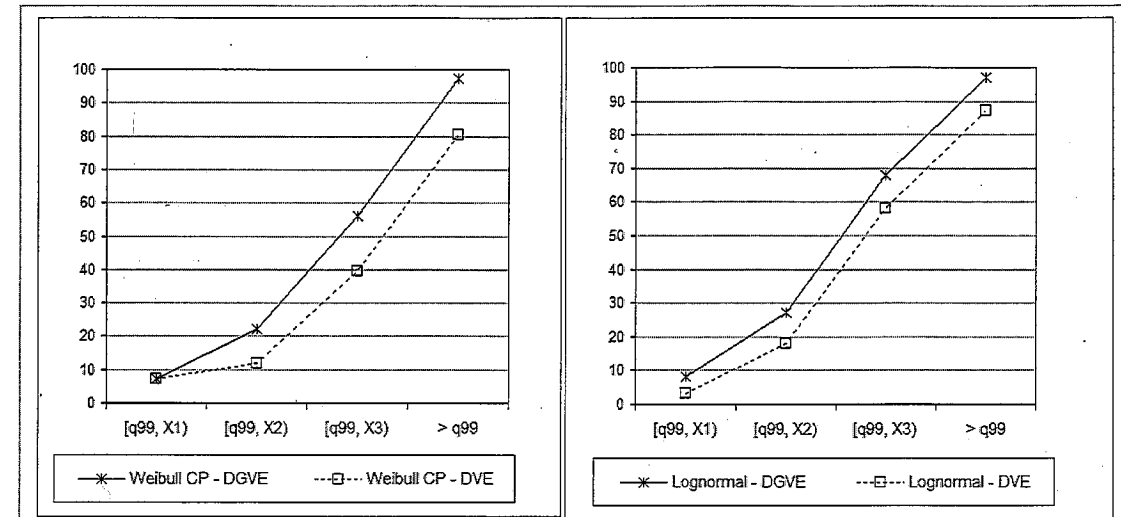


Figura 3.41: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 50.

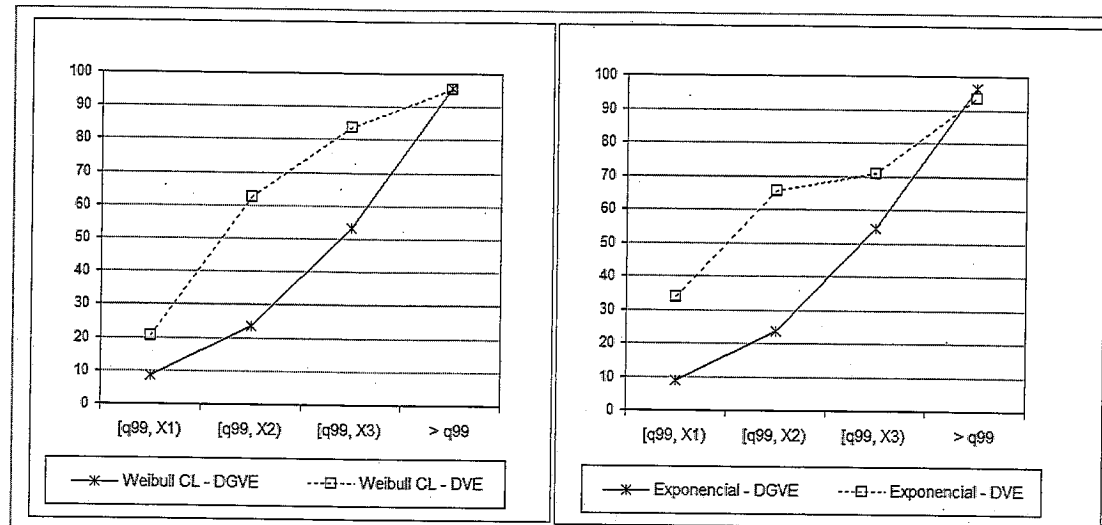


Figura 3.42: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 50.

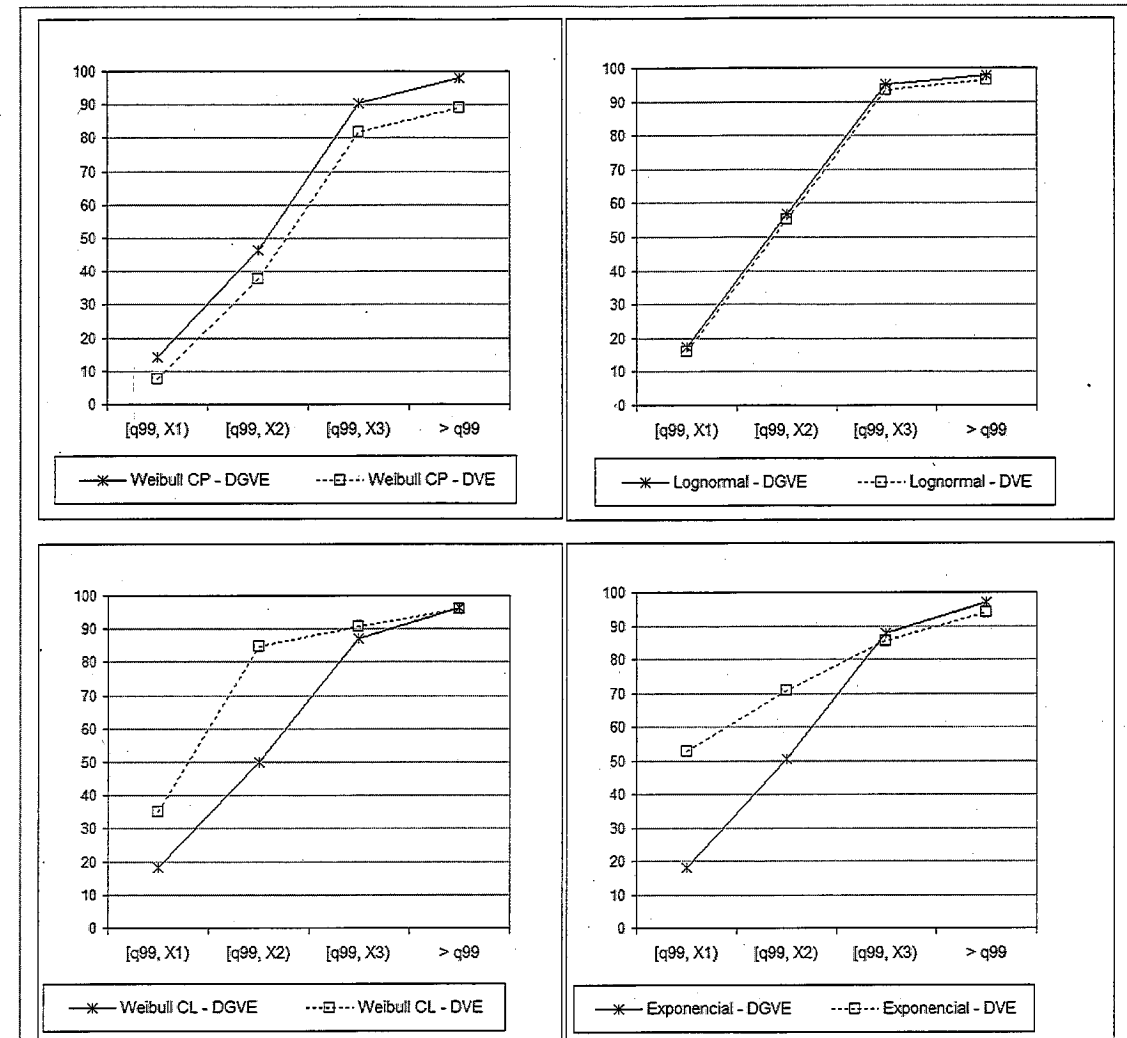


Figura 3.43: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 100.

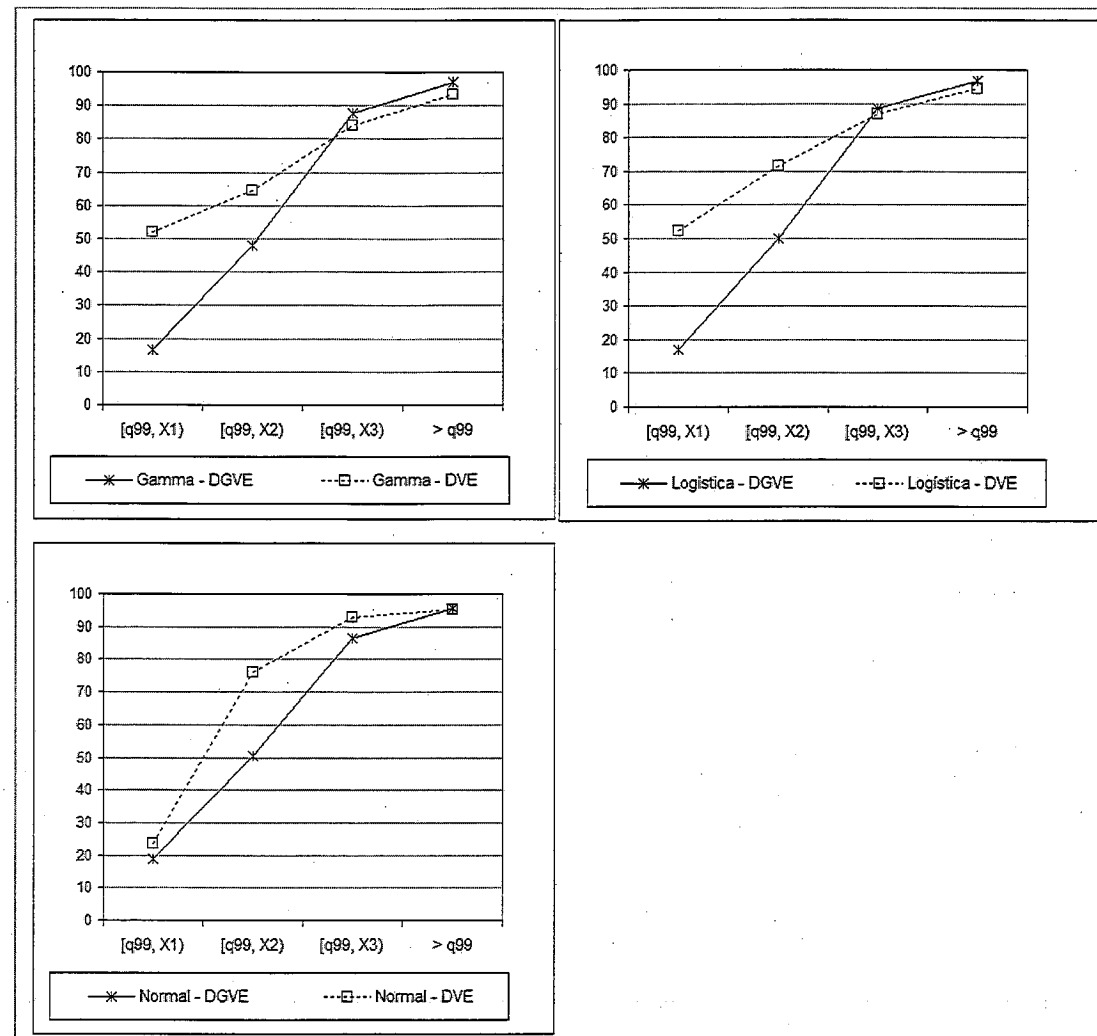


Figura 3.44: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 100.

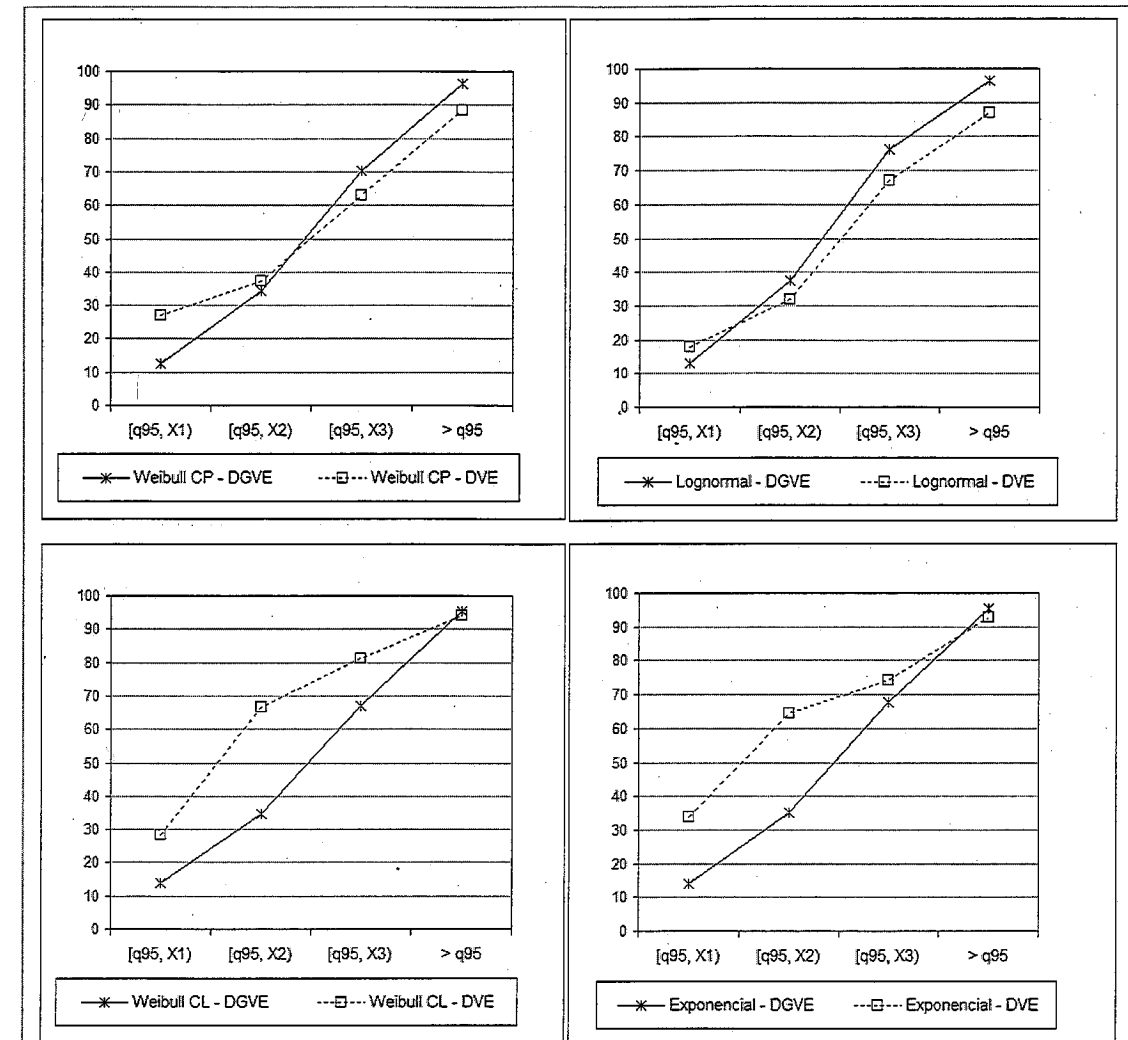


Figura 3.45: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 25.

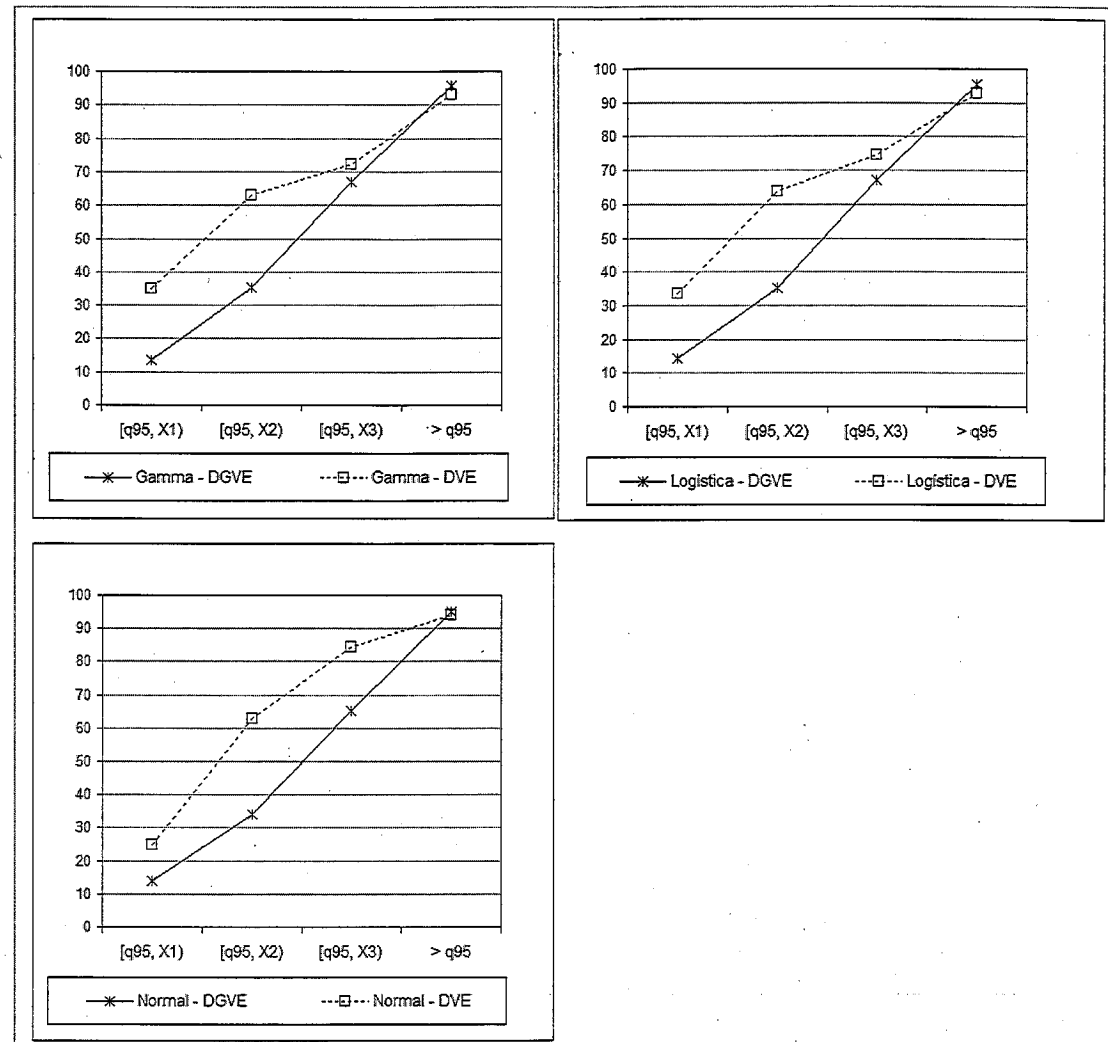


Figura 3.46: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 25.

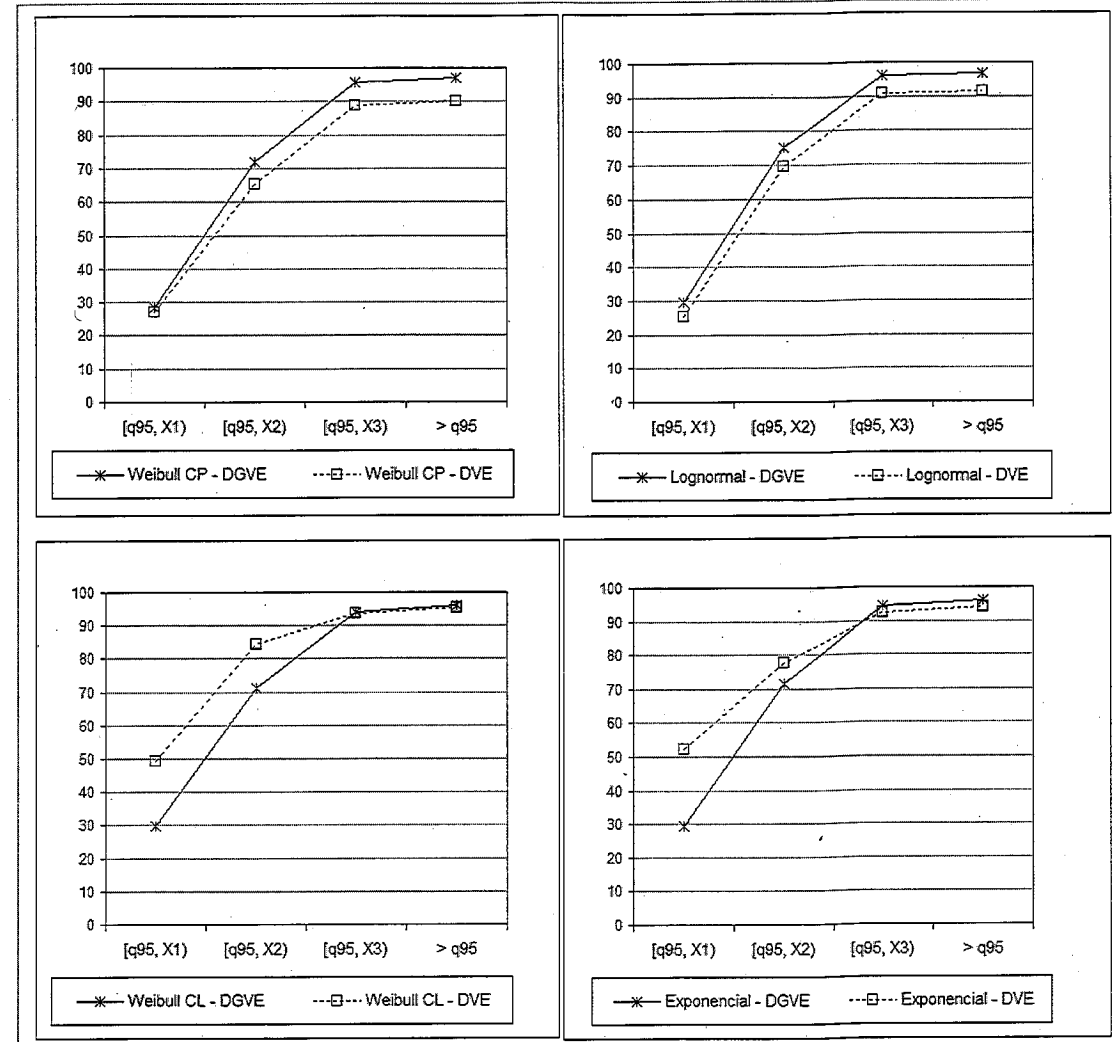


Figura 3.47: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 50.

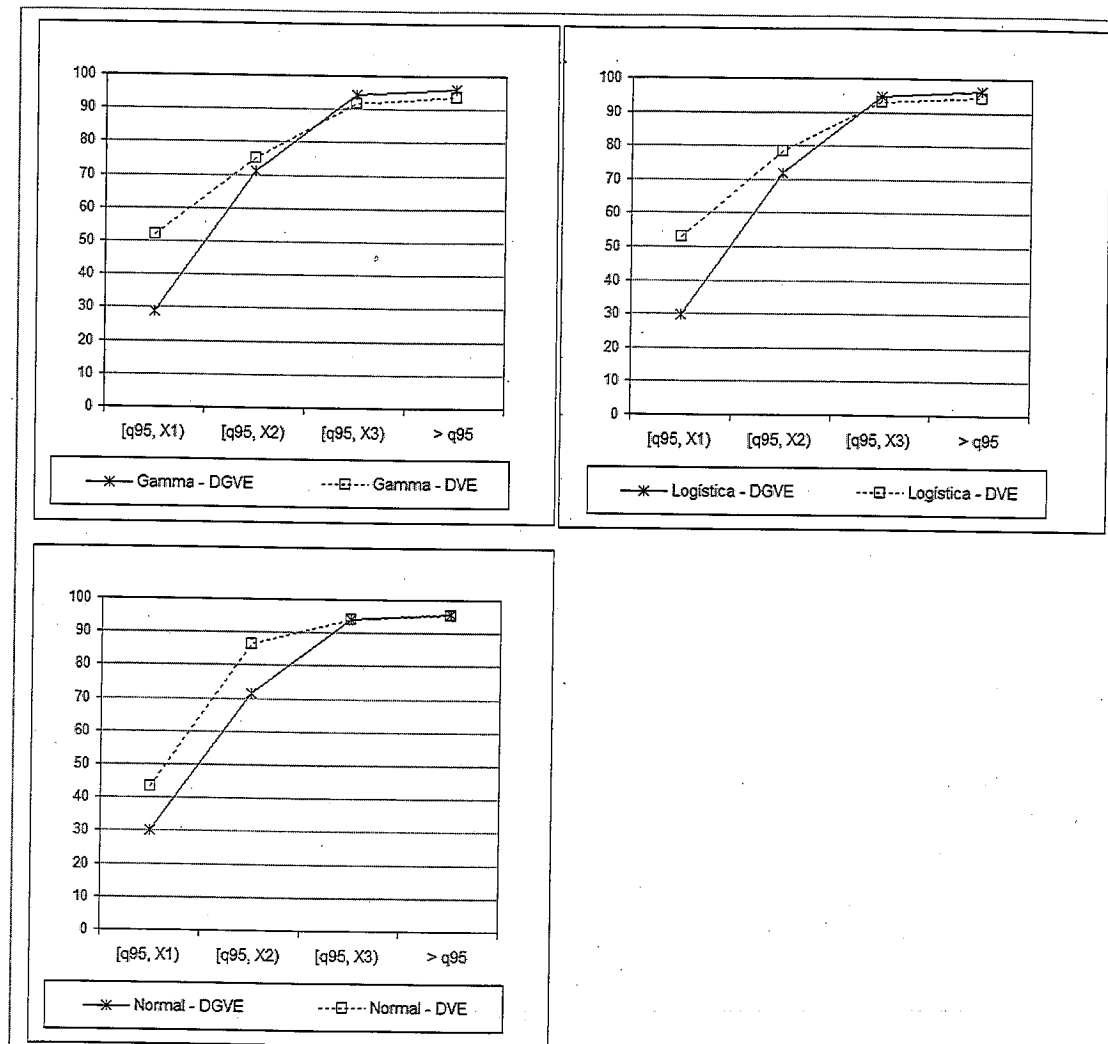


Figura 3.48: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 50.

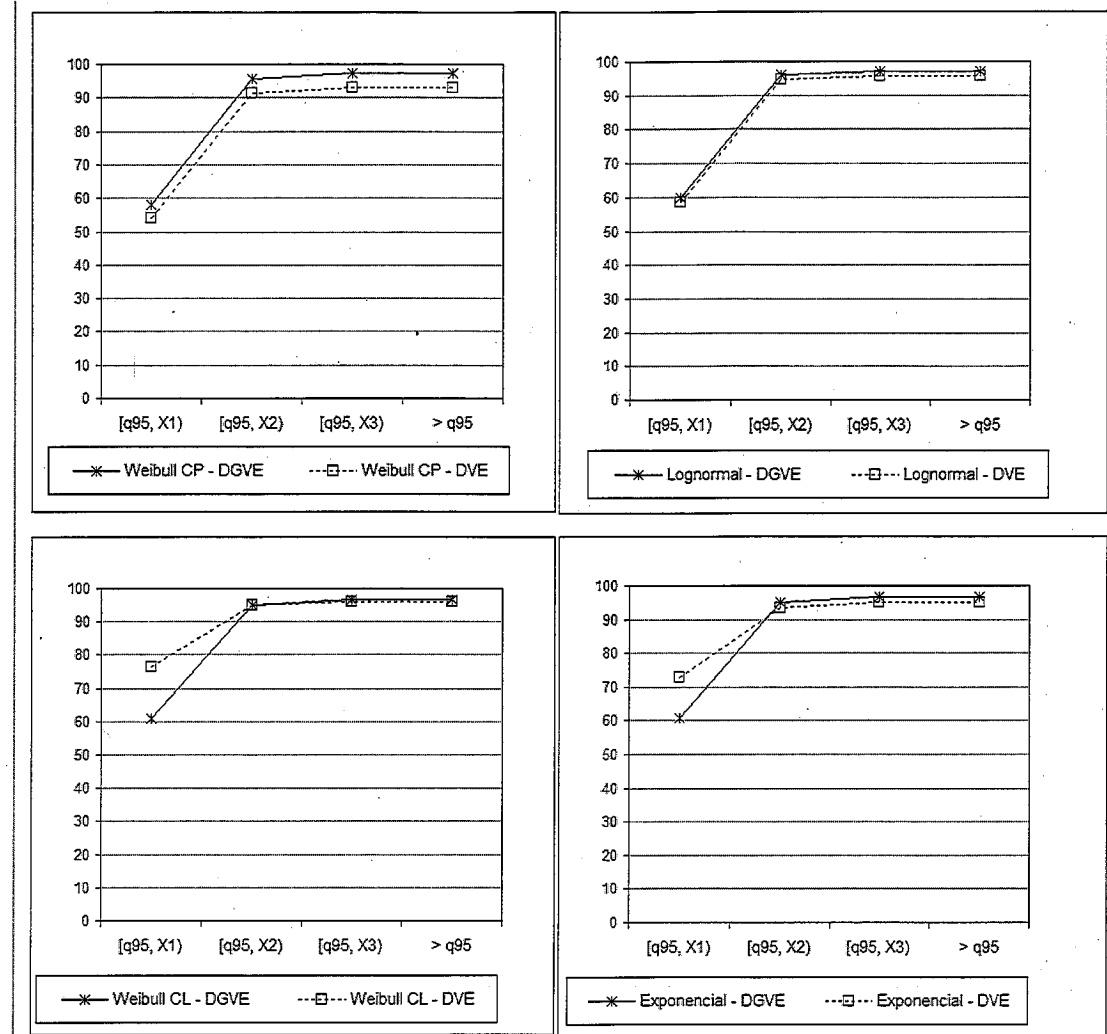


Figura 3.49: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 100.

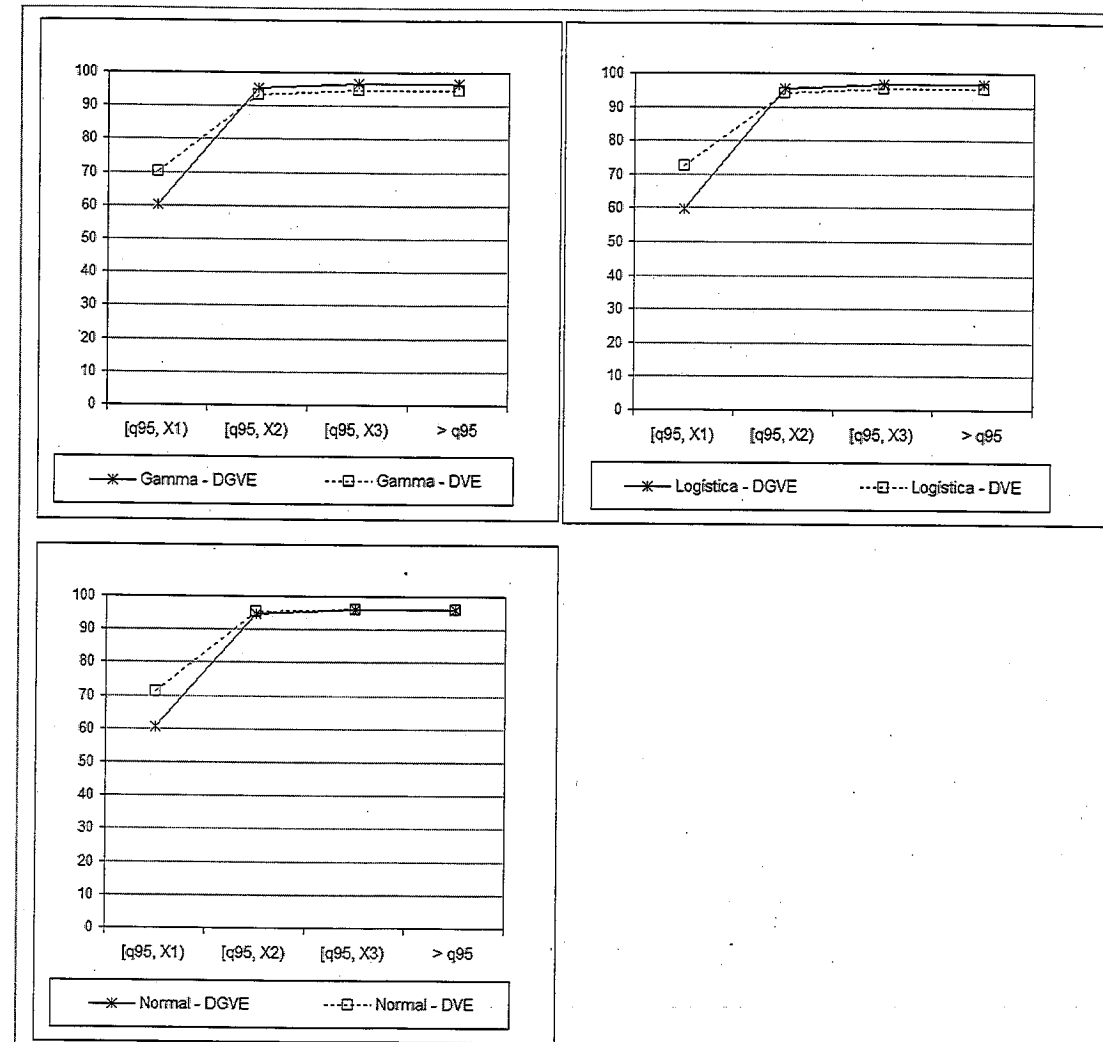


Figura 3.50: Colocación del extremo derecho de los intervalos de verosimilitud para distribuciones que pertenecen al DA Gumbel. Estimaciones para el cuantil de 95% con un tamaño de muestra igual a 100.

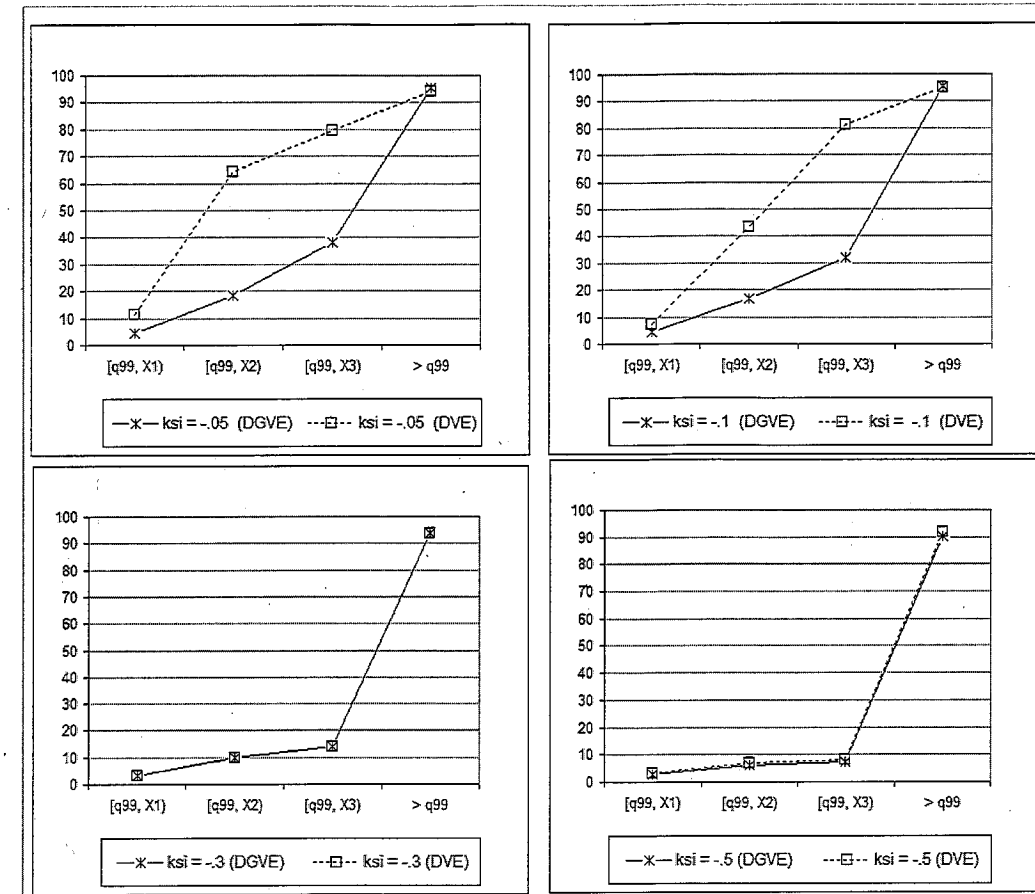


Figura 3.51: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Beta. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 25.

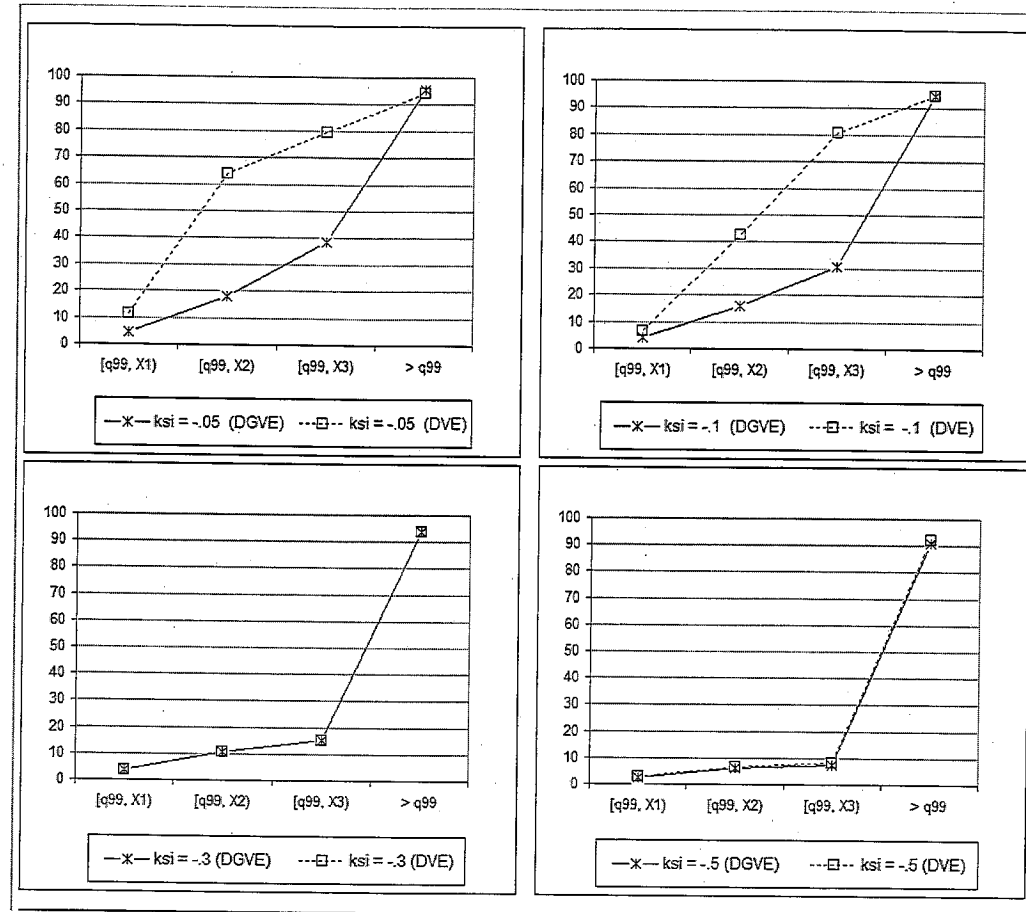


Figura 3.52: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Weibull (DVE). Estimaciones para el cuantil de 99 % con un tamaño de muestra igual a 25.

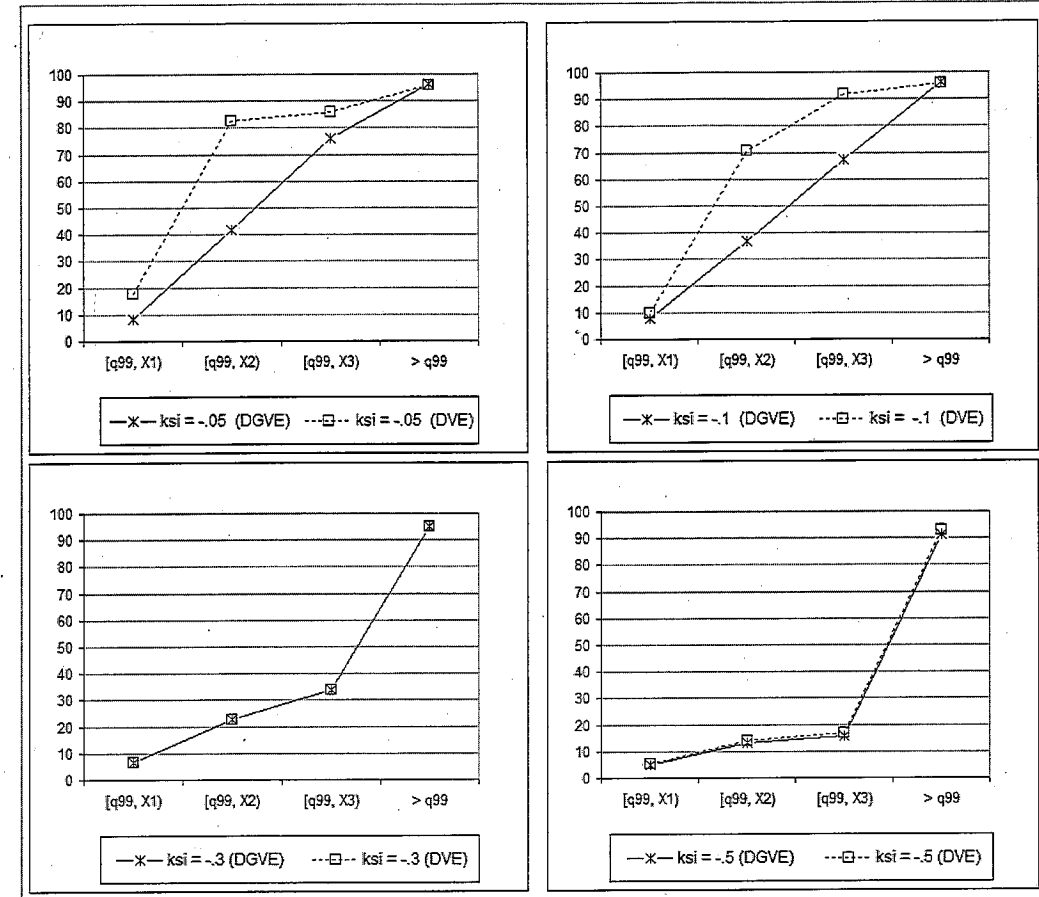


Figura 3.53: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Beta. Estimaciones para el cuantil de 99 % con un tamaño de muestra igual a 50.

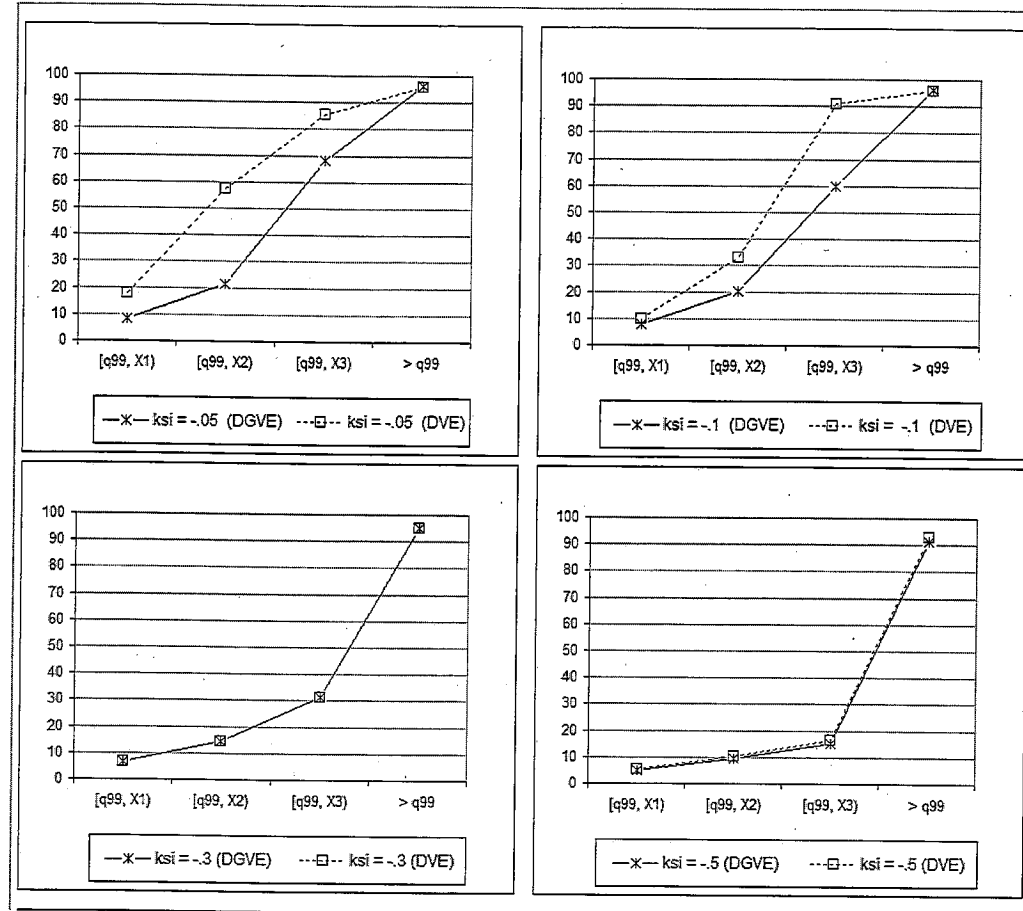


Figura 3.54: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Weibull (DVE). Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 50.

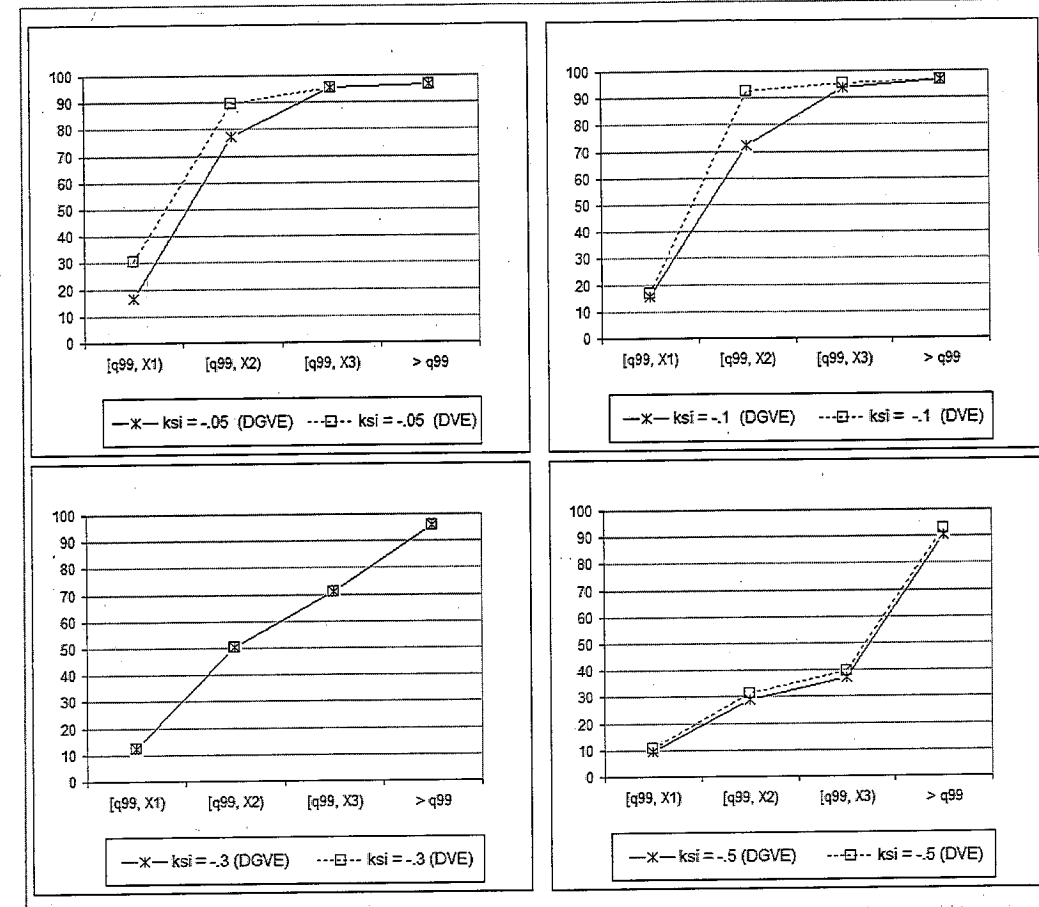


Figura 3.55: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Beta. Estimaciones para el cuantil de 99% con un tamaño de muestra igual a 100.



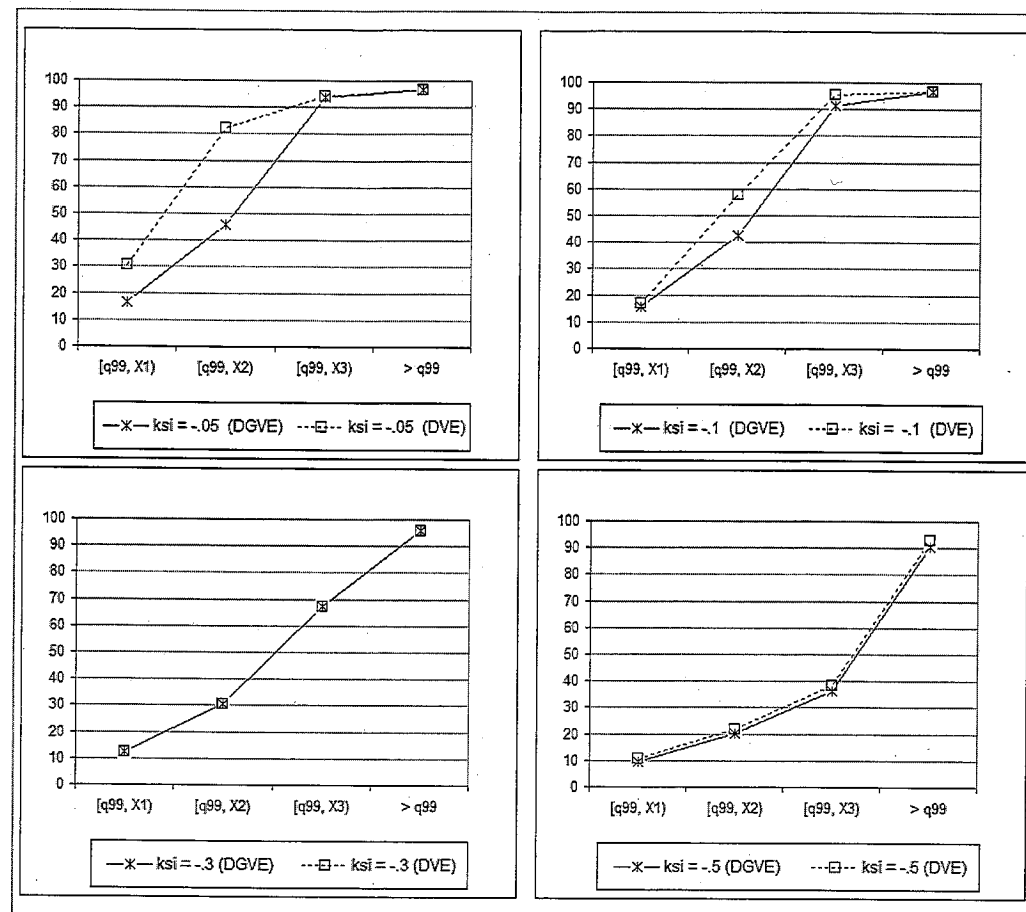


Figura 3.56: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Weibull (DVE). Estimaciones para el cuantil de 99 % con un tamaño de muestra igual a 100.

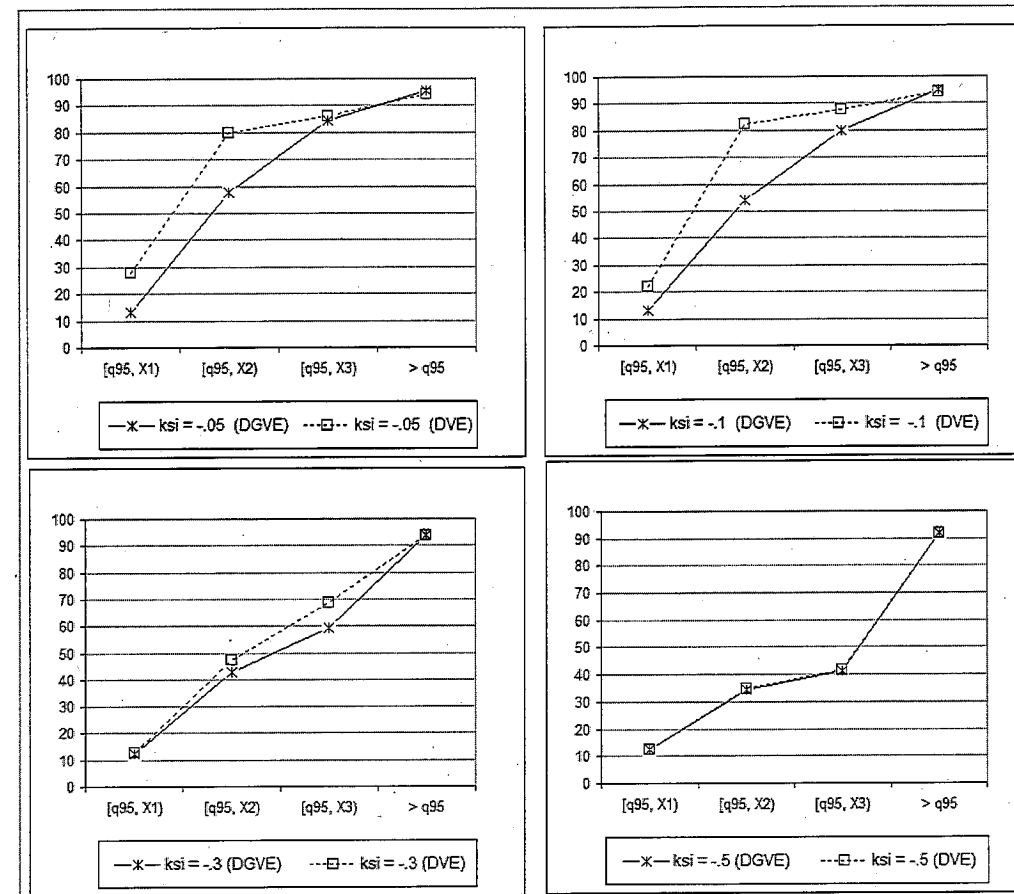


Figura 3.57: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Beta. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 25.

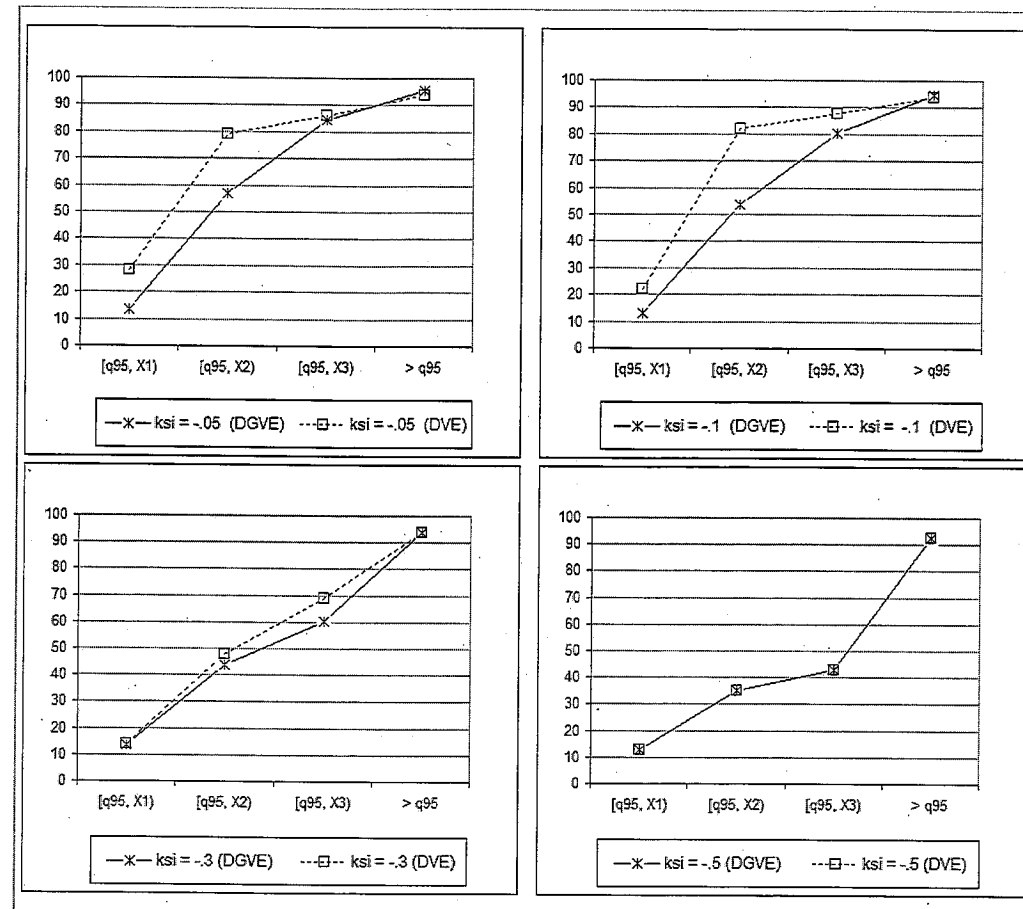


Figura 3.58: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Weibull (DVE). Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 25.

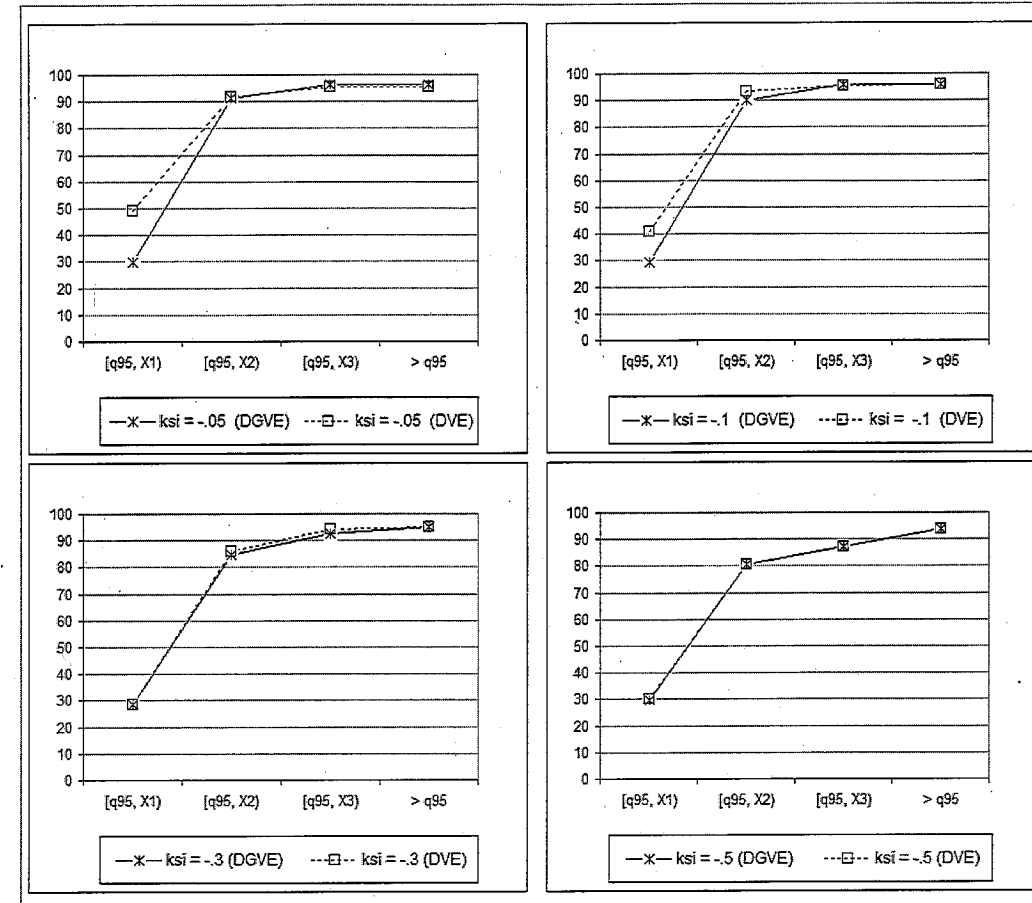


Figura 3.59: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Beta. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 50.

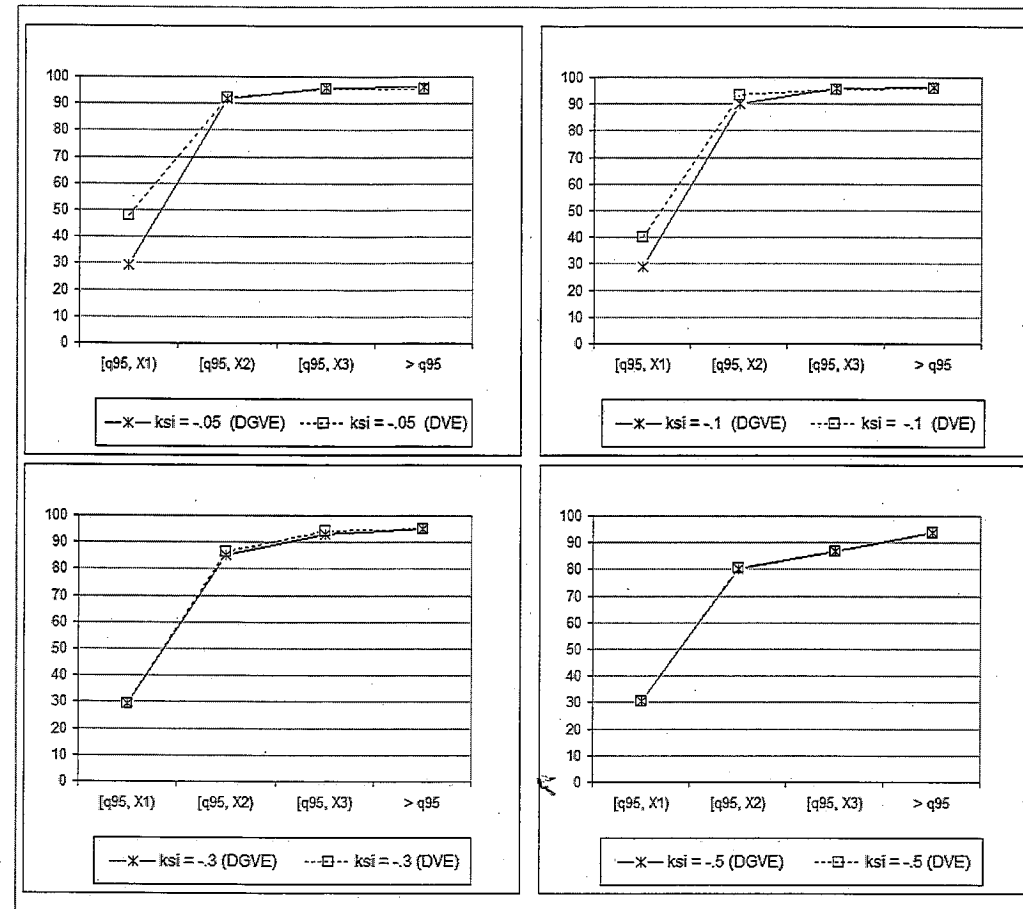


Figura 3.60: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Weibull (DVE). Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 50.

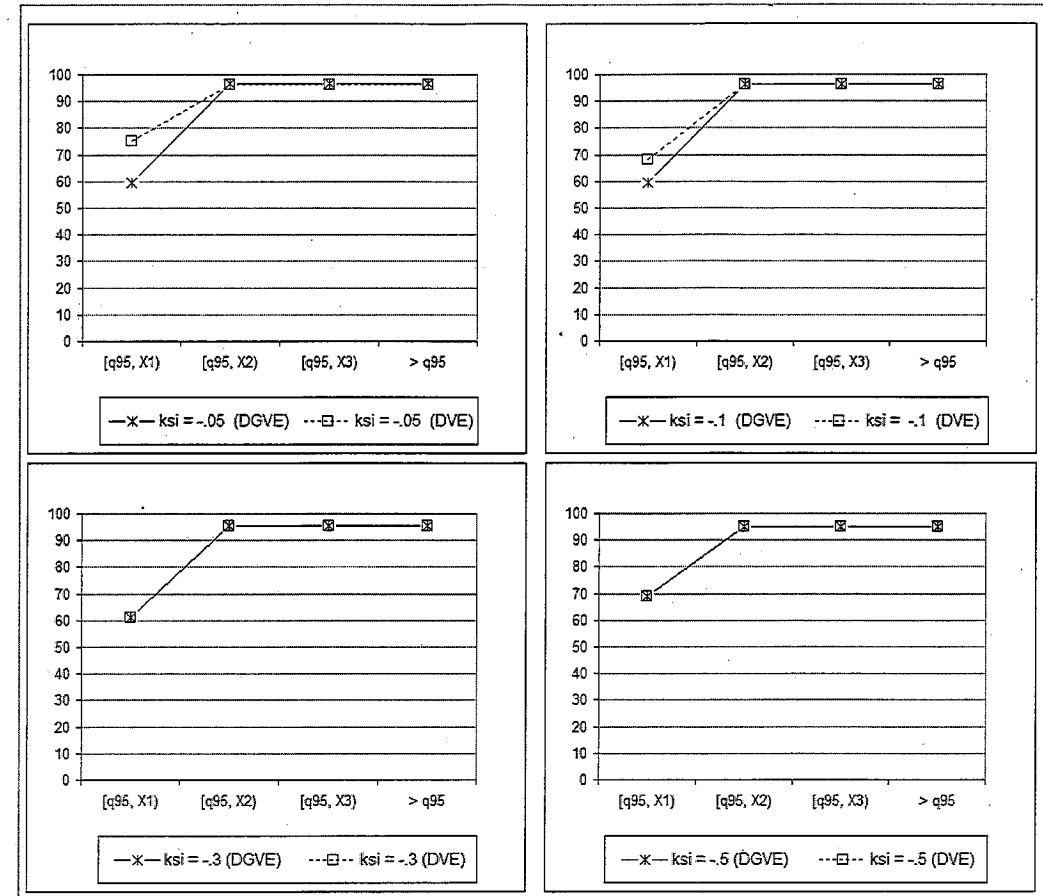


Figura 3.61: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Beta. Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 100.

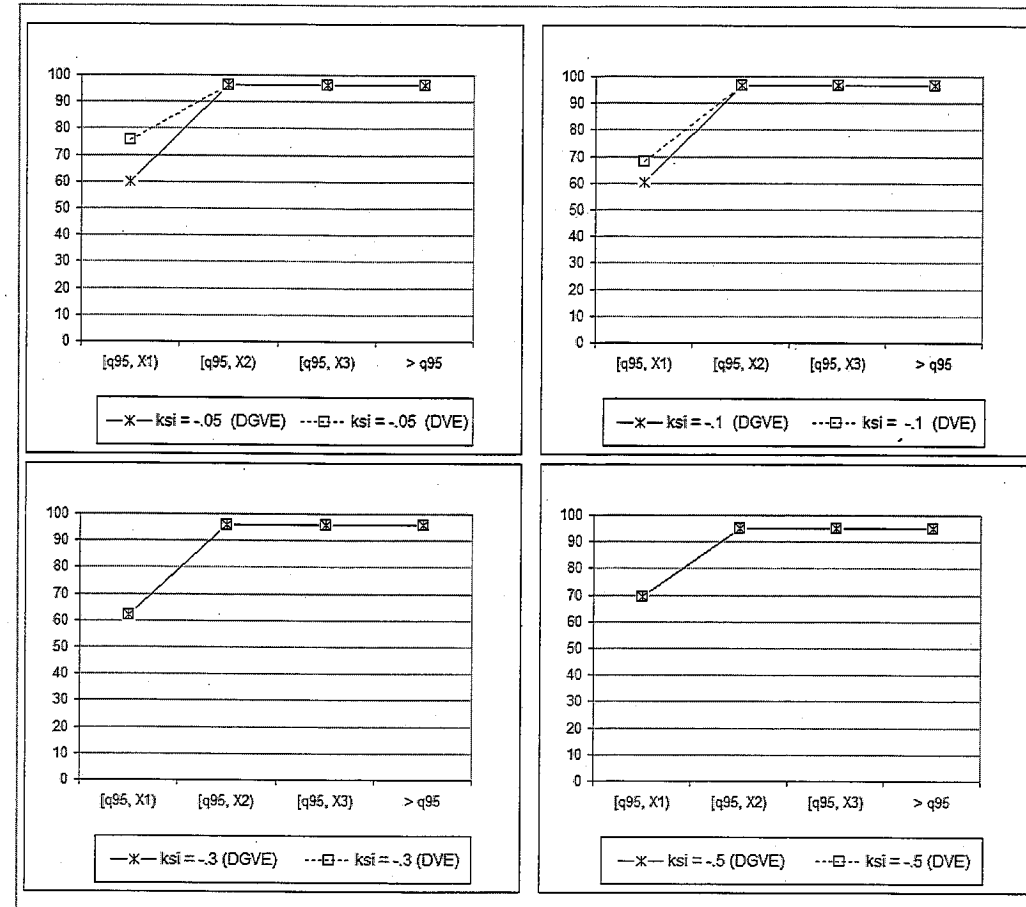


Figura 3.62: Colocación del extremo derecho de los intervalos de verosimilitud para las distribuciones de la familia Weibull (DVE). Estimaciones para el cuantil de 95 % con un tamaño de muestra igual a 100.

## Conclusiones

En este trabajo comparamos dos enfoques que podemos aplicar en el análisis estadístico de máximos por bloques. La diferencia más importante entre esos enfoques se refiere al modelo a partir del cual se hace inferencia. Uno de ellos considera como modelo a la distribución generalizada de valores extremos (*DGVE*), mientras que el otro propone usar algún tipo de las distribuciones de valores extremos (*DVE*). Para comparar esos enfoques calculamos los intervalos de verosimilitud que se obtienen, en cada caso, para los niveles de retorno con períodos de retorno de 20 y 100 años -o cuantiles del 95 % y 99 %-, respectivamente. A continuación resumimos los resultados obtenidos poniendo énfasis en los que corresponden al cuantil de 99 %, pues en ese caso es más difícil obtener estimadores por la extrapolación que se debe hacer debido a la menor cantidad de información disponible. Además, en términos generales, las observaciones que se hagan serán válidas también para las estimaciones que se refieren al cuantil de 95 %. Para comparar los intervalos de verosimilitud que obtuvimos bajo cada enfoque calculamos su cobertura, el cociente entre las longitudes de los intervalos -tomando como numerador la de los intervalos que obtuvimos usando la *DGVE*-, y por último, elaboramos un análisis de su extremo derecho.

En lo que se refiere a la cobertura, la que corresponde a los intervalos calculados con el enfoque basado en la *DGVE*, presentó, en casi todos los casos, valores cercanos al 95 %, y además, mayores a los obtenidos con las estimaciones basadas en las *DVE*. Sin embargo, las coberturas altas obtenidas con el enfoque basado en la *DGVE* también estuvieron acompañadas, en algunos casos, con intervalos más anchos comparados con los que obtuvimos con el otro enfoque. A continuación resumimos los principales resultados encontrados dividiendo a las distribuciones consideradas en el estudio según el tipo de dominio de atracción (*DA*) al cual pertenecen.

Para las distribuciones en el *DA* Fréchet con  $\xi = .05$  y  $.1$ ,<sup>1</sup> los intervalos de verosimilitud

<sup>1</sup>La asociación de un parámetro de forma con las distribuciones que pertenecen al *DA* Fréchet o Weibull

obtenidos con el enfoque basado en las *DVE* tienden a presentar menor longitud comparados con los obtenidos con el otro enfoque. Sin embargo, es también con esos intervalos que obtuvimos niveles de cobertura más bajos, lo cual se traduce en un mayor número de casos en que los intervalos simulados quedaron "por abajo" del valor real. Considerando las estimaciones para el cuantil de 99 %, las coberturas, para los diferentes tamaños de muestra, fueron de aproximadamente 85 % con  $\xi = .05$ , y de 80 % para  $\xi = .1$ . Mientras que para  $\xi = .3$  ó  $.5$  fueron superiores al 90 % en casi todos los casos, excepto con  $\xi$  igual a  $.3$  y con un tamaño de muestra igual a 25, pues en ese caso fue de aproximadamente 80 %.

Con base en los resultados obtenidos, las distribuciones consideradas en el estudio que pertenecen al DA Gumbel pueden ser divididas en dos grupos. El primero, constituido por las distribuciones Weibull de colas livianas (Weibull C.L.), Exponencial, Gamma, Logística y Normal, presenta, para las intervalos calculados con el enfoque basado en las *DVE*, niveles de cobertura de aproximadamente 90 %, además de que esos intervalos tienden a presentar una menor longitud que los calculados con el otro enfoque. Además, las observaciones anteriores son ciertas para los tres tamaños de muestra considerados. En contraste, con las distribuciones Weibull de colas pesadas (Weibull C.P.) y Lognormal, las estimaciones obtenidas con el enfoque basado en las *DVE* tienen asociadas niveles de cobertura bajos que están entre 74 % y 85 %, lo anterior a pesar de que no se puede apreciar una reducción importante en la longitud de dichos intervalos.

Finalmente, para las distribuciones que pertenecen al DA Weibull encontramos, con las estimaciones obtenidas usando las *DVE*, niveles de cobertura altos superiores al 87 %, así como menor longitud de los intervalos.

En resumen, asociando a las distribuciones que pertenecen al DA Gumbel un parámetro de forma igual a 0, tenemos que los resultados obtenidos muestran evidencia de que para las distribuciones con  $|\xi|$  menor o igual a  $.1$  podemos obtener intervalos de verosimilitud más angostos usando el enfoque basado en las *DVE*. Sin embargo, debemos tener presente que en ese caso obtenemos menores niveles de cobertura, situación que posiblemente se puede atribuir a la selección automática que se realizó, bajo ese enfoque, de algún tipo de las *DVE*.

Los resultados obtenidos no deben interpretarse como argumentos a favor o en contra de la aplicación del enfoque basado en las *DVE*, pues el objetivo que dió origen al estudio basado en simulaciones implementado, fue la comparación, por un lado, del enfoque basado en la *DGVE*, y por el otro, la propuesta de Díaz-Francés, Ortega y Vilchis (y con respecto a dicha propuesta, el enfoque basado en las *DVE* es una versión simplificada). Considerando

se hace siguiendo la convención mencionada al comienzo de la Sección 3.2.

el objetivo anterior, los resultados obtenidos deben tomarse como una guía para aplicar la propuesta de Díaz-Francés, Ortega y Vilchis. Es decir, si al considerar un problema específico se tiene información (por el contexto o estudios previos) sobre el tipo de distribución de valores extremos que le corresponde a la distribución de máximos de interés, entonces el estudio basado en simulaciones que elaboramos nos da una idea sobre cómo se comportarán las estimaciones que se obtienen con la propuesta de Díaz-Francés, Ortega y Vilchis. En particular, si es posible que el fenómeno de interés corresponda con alguno de los escenarios donde las estimaciones obtenidas con el enfoque basado en las *DVE* tuvieron asociadas coberturas por abajo del 85 %, entonces sabemos que debemos prestar especial atención a la selección del modelo, es decir, se debe considerar con detalle las características del fenómeno que estamos modelando, así como el ajuste de los modelos que resulten plausibles.

Finalmente, los resultados en este trabajo sugieren temas cuyo estudio es de interés. Por ejemplo, notamos que en el caso de las distribuciones que pertenecen al DA Fréchet las reducciones más grandes en la cobertura se obtuvieron cuando el parámetro de forma estaba "cerca" a cero. El estudio elaborado mostró que esa situación tiene más importancia con  $\xi = .1$ , sin embargo, es de interés considerar una mayor cantidad de valores para  $\xi$  a fin de conocer, de manera más precisa, el rango de valores de  $\xi$  en los cuales obtenemos ese comportamiento de la cobertura. Por otro lado, el comportamiento diferenciado obtenido en las estimaciones al considerar diferentes distribuciones refleja que la aproximación asintótica no se comporta igual para todas las distribuciones, por lo tanto, otro aspecto de interés consiste en analizar la velocidad de convergencia de la distribución de máximos normalizados a una distribución de valores extremos.

## Bibliografía

- [1] Coles, S. (2001) *An Introduction to Statistical Modelling of Extreme Values*. Springer, New York.
- [2] Castillo, E., Hadi, A.S., Balakrishnan, N. y Sarabia, J.M.(2005) *Extreme Value and Related Models with Applications in Engineering and Science*. Wiley, New York.
- [3] Embrechts, P., Klüpelberg, C. y Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer, New York.
- [4] Hogg, R.V. y Klugman, S. A. (1984) *Loss Distributions*. Wiley, New York.
- [5] Johnson, N.L. y Kotz, S. (1970) *Distributions in Statistics. Continuous Univariate Distributions I*. Wiley, New York.
- [6] Johnson, N.L. y Kotz, S. (1970) *Distributions in Statistics. Continuous Univariate Distributions II*. Wiley, New York.
- [7] Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.
- [8] Shao, J. (1999) *Mathematical Statistics*. Springer, New York.
- [9] Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics*. Wiley, New York.
- [10] Smith, R. L. (1985) Maximum likelihood estimation in a class of non-regular cases. *Biometrika* **42**, 67-90.
- [11] Vilchis, C. (2006) *Modelación Estadística de Máximos por Bloques*. Tesis de Maestría no publicada. Programa de Maestría en Ciencias en el CIMAT.