



**DISTRIBUCIONES MAX-ESTABLES.
TEORÍA Y APLICACIONES**

José Manuel Sanguino Sánchez



DISTRIBUCIONES MAX-ESTABLES. TEORÍA Y APLICACIONES

José Manuel Sanguino Sánchez

Memoria presentada como parte de los requisitos
para la obtención del título de Grado en Matemá-
ticas por la Universidad de Sevilla.

Tutorizada por

Inmaculada Barranco Chamorro

Índice general

Resumen	1
English Abstract	3
1. Herramientas básicas	5
1.1. Conceptos elementales en distribuciones de tiempos de vida	5
1.2. Inferencia en muestras completas	10
1.3. Inferencia en muestras censuradas	13
1.4. Distribución Exponencial	15
2. Una clase de familias max-estables	19
2.1. Introducción	19
2.2. Origen de las distribuciones max-estables y algunos modelos	21
2.2.1. Modelos generales	21
2.2.2. Modelos basados en tiempos de vida y distribuciones extremas	23
2.2.3. Modelos basados en distribuciones de ingresos	24
2.2.4. Propiedades probabilísticas	26
2.2.5. Medidas de desigualdad	32

2.2.6.	Inferencia	34
2.2.7.	Generalizaciones	38
3.	Modelo Exponencial Exponenciado	41
3.1.	Definición del modelo	41
3.2.	Algunas propiedades básicas	43
3.2.1.	Comparación del modelo EE con las distribuciones Gamma y Weibull	45
3.3.	Inferencia en la distribución EE	49
3.4.	Análisis de datos	52
3.5.	Experimentos numéricos y discusiones	54
3.6.	Conclusiones	58
	Bibliografía	59

Resumen

El objetivo de este trabajo es el estudio de las distribuciones max-estables. Son modelos paramétricos en los que se incorpora un parámetro de forma. En ellos se darán propiedades que permiten estudiar la forma de las distribuciones, obtener características de interés en teoría de la fiabilidad y análisis de supervivencia, y hacer inferencia. Estas distribuciones tienen aplicaciones en distribuciones de ingresos y distribuciones de tiempo de vida. Específicamente, nos centramos en el estudio de una distribución max-estable, la distribución exponencial exponenciada, que tiene un especial interés como modelo de tiempo de vida. Se han estudiado la estimación de los parámetros del modelo. Cabe destacar que en la mayoría de los casos, para el proceso de estimación habrá que recurrir a técnicas numéricas. El estudio teórico de esta distribución se ha completado con aplicaciones a datos reales.

La memoria se ha estructurado en tres capítulos. En el Capítulo 1 se introducen herramientas básicas en modelos de tiempo de vida, como son la función de supervivencia, función de riesgo o hazard, cuantiles, curva de Lorenz o índice de Gini. Para la estimación de parámetros se propone el método de máxima verosimilitud, introduciendo las modificaciones adecuadas para tratar con muestras censuradas. El capítulo se completa con una recopilación de resultados de la distribución exponencial que será clave para capítulos posteriores.

El Capítulo 2 se dedica a la teoría general de las distribuciones max-estables. En este capítulo vamos a estudiar varios aspectos de la clase de distribuciones max-estables. Se estudian diferentes modelos, distinguiendo entre modelos generales, modelos basados en distribuciones de vida y distribuciones extremas y modelos basados en distribuciones de ingresos. Se recogen varias propiedades probabilísticas de las distribuciones como la función de densidad, la función de hazard, la función de supervivencia, relaciones de orden, momentos y estadísticos de orden. Se dan expresiones explícitas para las correspondientes curvas de Lorenz y los índices de Gini. Finalmente, se hace

inferencia en estas distribuciones.

Finalmente, el Capítulo 3 se dedica a la distribución exponencial exponenciada. Se obtienen su función de supervivencia, función de hazard y momentos, los cuáles nos serán útiles para estudiar la forma de la distribución y propiedades de los estimadores. Se realiza una comparación de esta distribución con la distribución Weibull y Gamma. Se hace inferencia sobre sus parámetros y, finalmente, se estudiarán aplicaciones con datos reales y experimentos numéricos.

English Abstract

The aim of this dissertation is to study max-stable class of distributions. These are parametric models in which a shape parameter is incorporated into the model. We will give properties that allow us to study the shape of the distribution, to obtain summaries of interest in Reliability Theory and Survival Analysis, and to carry out Inference. These distributions have applications in income distributions and lifetime distributions. Specifically, we will focus on the study of a max-stable distribution, the exponentiated exponential distribution, which is of particular interest as a lifetime model. The issue of estimation of parameters in the model is studied. It is important to note that, quite often, numerical techniques are necessary in the process of estimation. The theoretical study of this distribution has been completed with applications to real data.

The document is divided into three chapters. In Chapter 1, the basic tools in lifetime models are introduced. These are: survival function, hazard function, quantiles, Lorenz curve and Gini index. Estimation of parameters is carried out by using maximum likelihood. Specific results for censored samples are given. To conclude the chapter, a section is included dealing with relevant results in the exponential distribution, which will be used later.

Chapter 2 is devoted to the general theory of max-stable distributions. In this chapter we study several aspects of the max-stable class of distributions. A revision of the different models proposed in the literature is done, distinguishing among general models, models based on lifetime and extreme distributions and models based on income distributions. Several probabilistic properties of these families are studied including density, hazard, survival function, order relations, moments and order statistics. Explicit expressions for the corresponding Lorenz curves and the Gini indices are provided. Finally, some inference procedures are discussed.

Finally, Chapter 3 is devoted to the exponentiated exponential distribution. Prop-

erties, moment and moment generating functions, and results, which allow us to interpret the parameters in this model, are given. A comparison of this distribution with the Weibull and Gamma distribution is carried out. Some inference procedures are discussed and, finally, applications with real data and numerical experiments will be studied.

1 | Herramientas básicas

Este trabajo es una introducción a la Teoría de Fiabilidad y Análisis de Supervivencia, en la que estudiaremos distribuciones de probabilidad relacionadas con este campo. Estas distribuciones se suelen utilizar para modelar tiempos de vida sobre muestras completas y sobre muestras censuradas, por lo que comenzaremos viendo las herramientas básicas para trabajar con este tipo de distribuciones. A lo largo de todo el trabajo, se utilizarán variables continuas no negativas T para representar los tiempos de vida.

En este primer capítulo, nos centraremos en conceptos y resultados necesarios para entender la utilidad de las distribuciones de probabilidad que aquí se estudian, además de recordar propiedades básicas de la distribución exponencial, sobre la cual se basa la distribución que exponemos en el tercer capítulo.

1.1 Conceptos elementales en distribuciones de tiempos de vida

El término **fiabilidad** describe la probabilidad de que un elemento funcione correctamente durante un cierto período de tiempo.

En este trabajo abordaremos el estudio de **tiempos de vida** de los elementos. Los elementos que estudiaremos se agrupan en dos familias:

- **Productos reparables:** Aquellos en los que puede ocurrir más de un fallo, es decir, se pueden arreglar. Por ejemplo los móviles, el motor de un coche, etc.
- **Productos no reparables:** Aquellos que no sirven una vez que han fallado, es decir, no tienen arreglo. Por ejemplo una bombilla eléctrica, fusibles, etc.

Los elementos que estudiaremos en este trabajo son los que no pueden repararse. Las distribuciones que expondremos más adelante estudian su tiempo de vida. Estas distribuciones usan datos independientes e idénticamente distribuidos (lo denotaremos i.i.d.).

Llamaremos T a la variable aleatoria continua y no negativa que describe el tiempo de vida del individuo o ítem en nuestro estudio. Denotaremos por $F(t)$ a la Función de Distribución (FdD) de T y por $f(t)$ a la función de densidad (fdd). Estas dos funciones cumplen la siguiente propiedad.

$$F(t) = P(T \leq t) = \int_0^t f(x) dx \quad \text{y} \quad F'(t) = f(t), \quad t > 0.$$

| Definición 1.1. Llamamos *función de supervivencia* a

$$S(t) = P(T \geq t).$$

Su valor indica la probabilidad de que el tiempo de vida del producto sea mayor que el instante t , es decir, la probabilidad de que el producto no falle antes de ese instante.

Por consiguiente, la función de supervivencia se puede definir como

$$S(t) = 1 - F(t).$$

La proposición que vamos a enunciar nos da algunas propiedades de la función de supervivencia, obtenidas a partir de su relación con la Función de Distribución.

| Proposición 1.1. Sea T una variable aleatoria no negativa y sea $S(t)$ su función de supervivencia. Entonces, se cumple:

1. $S(t)$ es continua y estrictamente decreciente
2. $S(0) = 1$
3. $S(+\infty) = 0$

Demostración. 1. Tenemos que $S(t) = 1 - F(t)$, donde F es la FdD de T . Al ser T una variable aleatoria continua, $F(t)$ debe ser una función continua y, por tanto, $S(t)$ también. Además $F(t)$ es estrictamente creciente, luego $S(t)$ es estrictamente decreciente.

$$2. S(0) = 1 - F(0) = 1 - \int_0^0 f(x) dx = 1 - 0 = 1.$$

$$3. S(+\infty) = \lim_{t \rightarrow \infty} (1 - F(t)) = 1 - \lim_{t \rightarrow \infty} \int_0^t f(x) dx = 1 - \int_0^{\infty} f(x) dx = 1 - 1 = 0.$$

Definición 1.2. Se llama **función de hazard o función de riesgo** a la función que se define como

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}. \quad (1.1)$$

De (1.1), se obtiene

$$h(t)\Delta t \simeq \lim_{\Delta t \rightarrow 0} P(t \leq T < t + \Delta t \mid T \geq t), \quad (1.2)$$

lo cual es una aproximación de la probabilidad de fallo en el intervalo $(t, \Delta t)$.

La función $h(t)$ se puede definir también como

$$h(t) = \frac{f(t)}{S(t)}, \quad (1.3)$$

lo que nos facilitará mucho su cálculo.

La función de hazard $h(t)$ nos ayuda a estudiar la fiabilidad de los ítems. Podemos dividir su comportamiento en tres grandes tipos:

- **Comportamiento creciente:** Nos indica que el producto es más probable que falle conforme pasa el tiempo.
- **Comportamiento decreciente:** Indica que es más probable que haya fallos al principio de la vida útil del producto.
- **Comportamiento constante:** La probabilidad de fallo es la misma a lo largo de toda la vida útil del producto.

Si en el comportamiento del individuo se observa al comienzo un comportamiento decreciente, luego constante y finalmente creciente, obtenemos la llamada **curva de bañera o bathtub curve**.

| Definición 1.3. La *función de riesgo acumulada* se define como

$$H(t) = \int_0^t h(x) dx, \quad t > 0.$$

Observación 1.1. Si $h(t)$ es constante, al representar $H(t)$ obtendremos una línea recta. Si $h(t)$ es creciente, obtendremos una función con un crecimiento más rápido que el de una recta, y si $h(t)$ es decreciente, una que crece más lento.

| Proposición 1.2. Las funciones de hazard y de supervivencia cumplen

1. $f(t) = -S'(t), \quad \forall t > 0.$
2. $h(t) = -\frac{\partial}{\partial t} \log S(t), \quad \forall t > 0.$
3. $S(t) = \exp(-H(t)), \quad \forall t > 0.$

Demostración. 1. $S(t) = 1 - F(t) \Rightarrow S'(t) = -F'(t) = -f(t).$

$$2. \frac{\partial}{\partial t} \log S(t) = \frac{S'(t)}{S(t)} = -\frac{f(t)}{S(t)} = -h(t).$$

3. Integrando en la expresión anterior y sabiendo que $S(0) = 1,$

$$\begin{aligned} \frac{\partial}{\partial t} \log S(t) = -h(t) &\Rightarrow \int_0^x \frac{\partial}{\partial t} \log S(t) dt = \log S(t) \Big|_{t=0}^{t=x} = \log S(x) = - \int_0^x h(t) dt \Rightarrow \\ &\Rightarrow \log S(x) = -H(x) \Rightarrow S(x) = \exp(-H(x)). \end{aligned}$$

|

Cuantiles.

Los cuantiles son una medida interesante para el estudio de tiempos de vida. Concretamente, serán de utilidad la mediana y los cuantiles de orden muy bajo y muy alto, ya que nos pueden aportar detalles importantes para la interpretación de los datos. Para cualquier distribución, se definen como sigue

| Definición 1.4. Sea T una variable aleatoria. Se define el cuantil de orden p , al valor t_p tal que

$$P(T \leq t_p) = p, \quad 0 < p < 1.$$

Y, por tanto,

$$t_p = F_T^{-1}(p), \quad 0 < p < 1.$$

| Definición 1.5. La curva de Lorenz es una característica muy importante de las distribuciones de ingresos. Si la distribución de la renta tiene función de densidad $f_r(p)$ la curva de Lorenz puede obtenerse como sigue. Para cada $p \in (0, 1)$

$$R_F(p) = \frac{\int_0^p u f_r(u) du}{\int_0^\infty u f_r(u) du} = \frac{1}{R_m} \int_0^p u f_r(u) du,$$

donde R_m es la renta media.

| Proposición 1.3 (Propiedades). La curva de Lorenz tiene pendiente positiva en todos sus punto puesto que:

$$\left(\frac{dR}{dP} \right)_{P_0=P(r_0)} = \frac{\frac{dR(r_0)}{dr}}{\frac{dP(r_0)}{dr}} = \frac{r_0 f_r(r_0) / R_m}{f_r(r_0)} = \frac{r_0}{R_m} \geq 0$$

En el punto inicial $x_0 = 0$ la pendiente será nula. Además la curva de Lorenz es convexa puesto que la derivada segunda es siempre positiva:

$$\left(\frac{d^2 P}{dR^2} \right)_{P_0=P(r_0)} = \frac{d}{dr} \left(\frac{dR}{dP} \right) \frac{dr}{dP} = \frac{1}{R_m} \frac{dr(P_0)}{dP} = \frac{1}{R_m} \frac{1}{f_r(r_0)} \geq 0$$

| Definición 1.6. El coeficiente de Gini es una medida que se utiliza para medir la desigualdad en los ingresos. Es un número entre el 0 y el 1, donde 0 se corresponde con la perfecta igualdad (todos tienen los mismos ingresos) y el valor 1 se corresponde con la perfecta desigualdad (una persona tiene todos los ingresos y los demás ninguno). El índice de Gini es el coeficiente de Gini expresado en porcentaje. El coeficiente de Gini es la

proporción del área encerrada entre la bisectriz del cuadrado unidad y la curva de Lorenz (que denotamos a), y el área bajo la bisectriz del cuadrado unidad (denotada $a + b$), es decir, $a/(a + b)$. En la práctica se suele calcular utilizando la fórmula de Brown

$$G = \left| 1 - \sum_{k=1}^{n-1} (X_{k+1} - X_k) (Y_{k+1} + Y_k) \right|$$

De forma resumida, la Curva de Lorenz es una gráfica de concentración acumulada de la distribución de la riqueza superpuesta a la curva de la distribución de frecuencias de los individuos que la poseen, y su expresión en porcentajes es el índice de Gini.

1.2 Inferencia en muestras completas

El método de estimación que usaremos será el método de máxima verosimilitud. Vamos a ver algunas propiedades generales de estos estimadores.

El método de máxima verosimilitud es el más conocido de los métodos de estimación. El método original fue introducido por Fisher en 1922, y desde entonces ha recibido importantes contribuciones por parte de otros autores

Definición 1.7. Sea una m.a.s T_1, \dots, T_n de una variable aleatoria T con función de densidad $f(t, \underline{\theta})$ donde $\underline{\theta}$ es un vector de parámetros desconocido que toma valores en Θ . Entonces, se define la función de verosimilitud de las n variables como

$$L(t_1, \dots, t_n; \underline{\theta}) = f(t_1, \dots, t_n; \underline{\theta}) = \prod_{i=1}^n f(t_i; \underline{\theta}), \quad \forall \underline{\theta} \in \Theta. \quad (1.4)$$

Método.

El método para hallar el estimador de $\underline{\theta}$ consiste en hallar aquel valor $\hat{\underline{\theta}} = \hat{\underline{\theta}}(T_1, \dots, T_n)$ de tal forma que maximiza la función de verosimilitud, es decir, que verifica

$$L(t_1, \dots, t_n; \hat{\underline{\theta}}) = \max_{\underline{\theta} \in \Theta} L(t; \underline{\theta})$$

Este valor recibirá el nombre de **estimador máximo verosímil** o **estimador de máxima verosimilitud** (EMV) del parámetro $\underline{\theta}$.

Ahora bien, es frecuente que la función de verosimilitud en (1.7) tenga una expresión complicada, lo que dificultaría los cálculos. Esta función es positiva y sus máximos se mantienen al aplicar un logaritmo neperiano pues el \log es monótona, y por tanto conserva el orden. Por ello, es frecuente considerar el logaritmo neperiano de la función de verosimilitud. Esta nueva función recibe el nombre de **función de log-verosimilitud**, cuya expresión es

$$l(\underline{\theta}) = \log L(t; \underline{\theta}) = \sum_{i=1}^n \log f(t_i; \underline{\theta}), \quad \forall \underline{\theta} \in \Theta. \quad (1.5)$$

En muchas ocasiones, la ecuación o sistema de ecuaciones de verosimilitud se puede resolver sin muchos problemas. Sin embargo, en ocasiones hay que recurrir a métodos numéricos iterativos para poder hallar el estimador.

Caso uniparamétrico.

En el caso en el que tengamos un solo parámetro $\theta \in \Theta \subseteq \mathbb{R}$ y se verifiquen las siguientes condiciones de regularidad:

- El campo de variación de θ es un intervalo abierto del eje real.
- El campo de variación de la variable aleatoria poblacional no depende de θ .
- La fdd $f(t_i; \theta)$ es positiva y derivable respecto de θ .
- Se verifica la condición del máximo $\frac{\partial^2 l}{\partial \theta^2} |_{\theta=\hat{\theta}} < 0$.

Entonces, se cumplen las siguientes propiedades.

1. Invarianza. Si $\hat{\theta}$ es el EMV de θ , entonces $h(\hat{\theta})$ es el EMV de $h(\theta)$.
2. Consistencia. Los EMV son consistentes, es decir, $\forall \epsilon > 0$, se verifica

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| < \epsilon) = 1, \quad \forall \theta \in \Theta.$$

3. Insesgadez. En general, los EMV no serán insesgados. En cambio, suelen ser asintóticamente insesgados, es decir, se verificaría que

$$\lim_{n \rightarrow \infty} E[\hat{\theta}] = \theta$$

4. Normalidad asintótica. Al aumentar el tamaño de la muestra, la distribución de $\hat{\theta}$ puede aproximarse por una distribución normal.

Para un parámetro $\theta \in \Theta$, se verifica

$$\hat{\theta} \xrightarrow{\mathcal{L}} N\left(\theta, \sqrt{Var(\hat{\theta})}\right),$$

con $Var(\hat{\theta})$ la varianza asintótica de $\hat{\theta}$, y \mathcal{L} significa convergencia en ley.

Caso mutiparamétrico, $k \geq 2$.

En el caso en el que la función de densidad dependa de k parámetros, con $k \geq 2$, los EMV se obtienen resolviendo el sistema de ecuaciones de log-verosimilitud en $\theta_1, \dots, \theta_k$. Las ecuaciones de este sistema vendrán dadas por

$$\frac{\partial l(t_1, \dots, t_n; \theta_1, \dots, \theta_k)}{\partial \theta_j} = \sum_{i=1}^n \frac{\partial \log f(t_i; \theta_1, \dots, \theta_k)}{\partial \theta_j} = 0, \quad \forall j = 1, \dots, k.$$

En nuestro caso, será $k = 2$.

Bajo condiciones de regularidad los EMVs obtenidos tendrán buenas propiedades asintóticas, que son una generalización al caso mutiparamétrico de las expuestas anteriormente. En particular, será de gran interés la matriz de información de Fisher, pues nos permite obtener la matriz de varianza asintótica de los EMVs.

Matriz de información de Fisher.

En este documento trabajaremos con modelos de dos parámetros para T . Dada una muestra aleatoria de T , sus parámetros $\underline{\theta} = (\theta_1, \theta_2)$ se pueden estimar mediante estimadores de máxima verosimilitud. En aquellas situaciones en que se pueda suponer la normalidad asintótica de los estimadores EMV, nos interesa obtener la matriz de información de Fisher, pues a partir de ella se pueden proponer aproximaciones asintóticas para las varianzas de $\hat{\theta}_1$ y $\hat{\theta}_2$, al igual que para la covarianza $Cov(\hat{\theta}_1, \hat{\theta}_2)$.

Denotaremos por $I(\underline{\theta})$ a la matriz de información de Fisher. El elemento (i,j) de esta matriz es

$$-E \left[\frac{\partial^2 l(\underline{\theta})}{\partial \theta_i \partial \theta_j} \right], \quad i, j = 1, 2.$$

donde $l(\underline{\theta})$ es la función de log-verosimilitud de la muestra, dada por (1.5). Ahora bien, observamos en la definición de dicha matriz que depende de los valores de los parámetros, los cuales son desconocidos. Estos valores se suelen estimar por el método de máxima verosimilitud, obteniendo así la matriz de información de Fisher observada; $I(\hat{\underline{\theta}})$. El elemento (i, j) en la matriz observada viene dado por

$$-\frac{\partial^2 l(\underline{\theta})}{\partial \theta_i \partial \theta_j} \Bigg|_{\theta = \hat{\theta}}, \quad i, j = 1, 2.$$

La matriz inversa de $I(\hat{\underline{\theta}})$, $\hat{V} = I^{-1}(\hat{\underline{\theta}})$, nos da estimaciones de la varianza y covarianza de los estimadores de máxima verosimilitud de los parámetros. Concretamente, los elementos de la diagonal nos dan las varianzas y los elementos fuera de la diagonal, las covarianzas.

1.3 Inferencia en muestras censuradas

En los estudios de datos de supervivencia es muy frecuente encontrarse con muestras en las que los datos están **censurados**, es decir, la información que nos proporcionan está incompleta por algún motivo, el cual dependerá del tipo de censura. Por tanto, se necesitan mecanismos diferentes para afrontar estas muestras a los usados con las muestras completas. Hay diferentes tipos de censura, como podemos ver a continuación.

- **Censura de tipo I.** Se da cuando el investigador fija un tiempo máximo para la observación del tiempo de vida de los datos. Entonces, cuando un individuo observado no presenta un fallo dentro de ese tiempo, no aporta la información necesaria al investigador y constituye una observación censurada

- **Censura de tipo II.** Supongamos que se observa una muestra de n elementos y se fija un número r tal que $1 \leq r \leq n$. En este caso, el experimento termina cuando r individuos fallan, resultando por tanto r datos observados tales que $t_{1:n} \leq \dots \leq t_{r:n}$, y $n - r$ datos censurados, puesto que no han fallado al acabar la investigación. Por $t_{i:n}$ se ha denotado al estadístico de orden i en una muestra de tamaño n .
- **Censura aleatoria.** En este tipo, la censura no depende de ninguna decisión del investigador. Se da cuando cada individuo bajo estudio tiene un tiempo de vida T_i y un tiempo de censura C_i , siendo T_i y C_i variables aleatorias continuas e independientes entre ellas.

Para modelizar este tipo de datos, utilizaremos la siguiente notación. Supongamos una muestra aleatoria de n individuos, cuyos tiempos de vida vienen dados por las variables aleatorias T_1, \dots, T_n , respectivamente. Llamamos t_i a lo que puede ser o bien el tiempo de vida observado o bien el tiempo de censura según lo que ocurra primero, esto es, $t_i = \min\{T_i, C_i\}$, siendo C_i el tiempo de censura en la observación i . Definimos la variable δ_i , $i = 1, \dots, n$, tal que

$$\delta_i = \begin{cases} 1, & \text{si } t_i \text{ dato observado} \\ 0, & \text{si } t_i \text{ dato censurado.} \end{cases}$$

De esta manera, las muestras censuradas se escriben como (t_i, δ_i) , $i = 1, \dots, n$.

Método de máxima verosimilitud.

Para poder hacer inferencia sobre este tipo de datos, se suele utilizar el método de máxima verosimilitud, que se verá alterado por la naturaleza de estos datos. De forma general, denotaremos al vector con los parámetros del modelo a estudiar como $\underline{\theta} = (\theta_1, \theta_2)$. La función de verosimilitud cuando partimos de una muestra de datos censurados (t_i, δ_i) , $i = 1, \dots, n$, viene dada por

$$L(\underline{\theta}) = \prod_{i=1}^n f(t_i)^{\delta_i} S(t_i)^{1-\delta_i}, \quad (1.6)$$

donde f es la función de densidad de la muestra y S denota la función de supervivencia correspondiente.

Esta forma de la función de verosimilitud es común a los tres tipos de datos censurados nombrados anteriormente (ver Lawless [11]).

También puede ser de utilidad la función de log-verosimilitud, cuya expresión para este tipo de muestras es

$$l(\underline{\theta}) = \log L(\underline{\theta}) = \sum_{i=1}^n \delta_i \log f(t_i) + \sum_{i=1}^n (1 - \delta_i) \log S(t_i). \quad (1.7)$$

Matriz de información de Fisher

En el caso de muestras censuradas, sigue teniendo sentido considerar la matriz de información observada de Fisher. Al igual que en el caso de muestras completas, viene definida como la matriz donde cada elemento (i, j) viene dado por

$$-\frac{\partial^2 l(\underline{\theta})}{\partial \theta_i \partial \theta_j} \Bigg|_{\theta = \hat{\theta}}, \quad i, j = 1, 2.$$

En este caso, $l(\underline{\theta})$ representa la función de log-verosimilitud de la muestra censurada, dada por (1.7).

1.4 Distribución Exponencial

En esta sección se van a recoger algunos resultados importantes de la distribución exponencial, que nos serán útiles para abordar el estudio de modelos más complejos en capítulos posteriores.

Definición 1.8. Sea X una variable aleatoria absolutamente continua. Se dirá que X sigue una **distribución Exponencial** de parámetro $\lambda > 0$, y se denota por $X \sim E(\lambda)$ si su función de densidad se expresa como sigue:

$$f_X(x) = \lambda \exp(-\lambda x), \quad x \geq 0.$$

Por tanto, la función de distribución viene dada por

$$F(x) = \int_0^x \lambda \exp(-\lambda t) dt = [-\exp(-\lambda t)]_{t=0}^{t=x} = -\exp(-\lambda x) + 1 = 1 - \exp(-\lambda x), \quad x \geq 0$$

Este modelo tiene gran utilidad para los estudios de tiempos de vida.

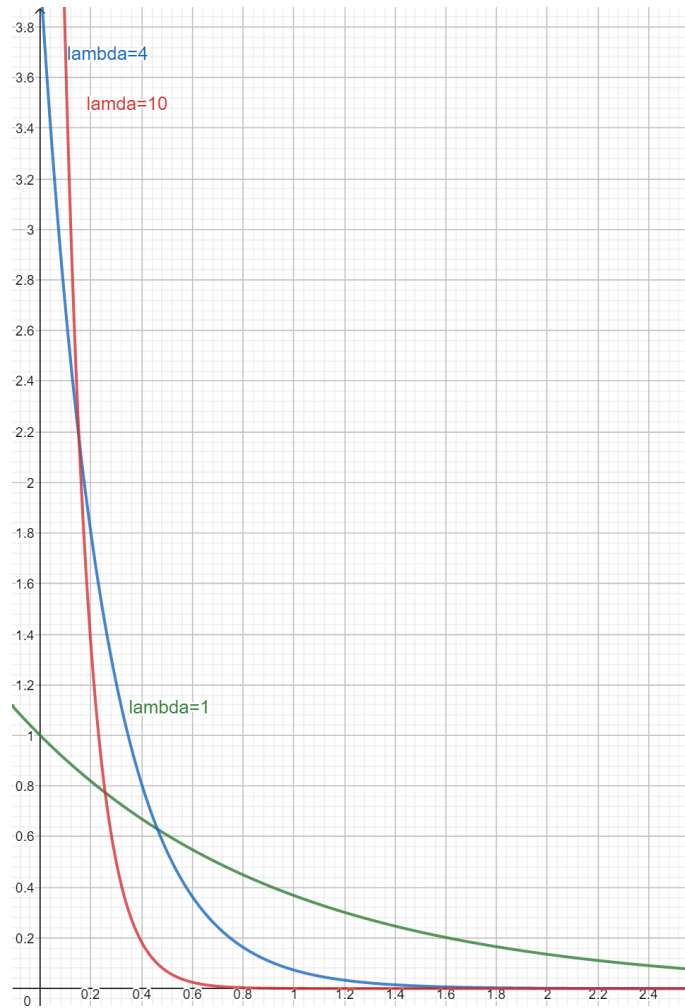


Figura 1.1: fdd de la distribución exponencial para diferentes valores de λ (azul $\lambda = 4$, rojo $\lambda = 10$ y verde $\lambda = 1$)

Proposición 1.4. Si X es una variable aleatoria tal que $X \sim E(\lambda)$, $\lambda > 0$, entonces

1. $E(X) = 1/\lambda$
2. $Var(X) = 1/\lambda^2$
3. $M_X(t) = \frac{\lambda}{\lambda - t}$, $t < \lambda$.

Demostración.

$$\begin{aligned}
 1. E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x \lambda \exp(-\lambda x) dx = [x(-\exp(-\lambda x))]_0^{\infty} \\
 &+ \int_0^{\infty} \exp(-\lambda x) dx = \left[-\frac{1}{\lambda} \exp(-\lambda x) \right]_0^{\infty} = \frac{1}{\lambda}. \\
 2. Var(X) &= E(X^2) - (E(X))^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \frac{1}{\lambda^2} = \int_0^{\infty} x^2 \lambda \exp(-\lambda x) dx - \frac{1}{\lambda^2} = \\
 &= [-x^2 \exp(-\lambda x)]_0^{\infty} + 2 \int_0^{\infty} x \exp(-\lambda x) dx - \frac{1}{\lambda^2} = \frac{2}{\lambda} \int_0^{\infty} \exp(-\lambda x) dx - \frac{1}{\lambda^2} = \\
 &= 2 \left[-\frac{\exp(-\lambda x)}{\lambda} \right]_0^{\infty} - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \\
 3. M_X(t) &= E(\exp(tX)) = \int_{-\infty}^{\infty} \exp(tX) \lambda \exp(-\lambda x) dx = \lambda \int_0^{\infty} \exp(-(\lambda - t)x) dx = \\
 &= \lambda \left[-\frac{1}{\lambda - t} \exp(-(\lambda - t)x) \right]_0^{\infty} = \lambda \frac{1}{\lambda - t} = \frac{\lambda}{\lambda - t}, \text{ para } t < \lambda.
 \end{aligned}$$

La distribución exponencial es la única distribución continua con falta de memoria. Veamos que significa esto.

| Proposición 1.5 (Pérdida de memoria). Si $X \sim E(\lambda)$ entonces se verifica

$$P(X > x + h | X > h) = P(X > x), \quad x > 0, h > 0$$

Se le "olvida" lo que ha ocurrido anteriormente

Demostración.

$$\begin{aligned}
 P(X > x + h | X > h) &= \frac{P(X > x + h, X > h)}{P(X > h)} = \frac{P(X > x + h)}{P(X > h)} = \\
 &= \frac{1 - F(x + h)}{1 - F(h)} = \frac{1 - (1 - \exp(-\lambda(x + h)))}{1 - (1 - \exp(-\lambda h))} = \frac{\exp(-\lambda x) \exp(-\lambda h)}{\exp(-\lambda h)} = \\
 &= \exp(-\lambda x) = P(X > x).
 \end{aligned}$$

2 | Una clase de familias max-estables

En este capítulo vamos a abordar el estudio general de distribuciones exponenciadas. Vamos a estudiar varios aspectos de estas distribuciones (que se van a clasificar en diferentes tipos) que se anunciarán a continuación. Este capítulo nos va a servir de puente para el estudio de un caso particular de distribución exponenciada, la distribución exponencial exponenciada, la cual veremos en el capítulo 3.

Sea F una función de distribución continua. En este capítulo vamos a estudiar varios aspectos de la clase de distribuciones max-estables $[F]^a$, a veces conocida como alternativas de Lehman. Se estudian diferentes modelos, distinguiendo entre modelos generales, modelos basados en distribuciones de vida y distribuciones extremas y modelos basados en distribuciones de ingresos. Se recogen varias propiedades probabilísticas de las distribuciones como la función de densidad, la función de hazard, la función de supervivencia, relaciones de orden, momentos y estadísticos de orden. Se proponen expresiones explícitas para las correspondientes curvas de Lorenz y los índices de Gini. Finalmente, se hace inferencia en estas distribuciones y se proponen generalizaciones de los modelos básicos. Se recogen en este capítulo aplicaciones prácticas de estas distribuciones a la modelización de rentas e ingresos.

2.1 Introducción

En el análisis de datos, es interesante partir de una distribución de probabilidad básica, y luego introducir nuevos parámetros en dicha distribución para obtener una distribución más flexible o construir modelos de regresión. Por ejemplo, la introducción de un parámetro escalar conduce a un modelo de vida acelerada y tomando po-

tencias de la función de supervivencia obtenemos modelos de hazard proporcionales. En este capítulo se estudian varios aspectos de la clase de distribuciones max-estables, en las que dadas un parámetro $a > 0$, definimos la función G_F como

$$G_F(x; a, \mu, \sigma) = [F(z)]^a, \quad x \in \mathbb{R}, \quad z = \frac{x - \mu}{\sigma}, \quad (2.1)$$

con función de densidad asociada

$$g_F(x; a, \mu, \sigma) = a \sigma^{-1} f(z) [F(z)]^{a-1}, \quad x \in \mathbb{R}, \quad (2.2)$$

donde $F(\cdot)$ es una función de distribución continua (fdD) con función de densidad asociada $f(\cdot)$, $\mu \in \mathbb{R}$ y $\sigma > 0$

La distribución (2.1) es bien conocida en el ámbito estadístico y ha sido discutida en una variedad de escenarios bajo el nombre de alternativa de Lehmann (Lehmann, 1959; Gupta et al., 1998; Arnold, 2004). Jones (2004) también ha considerado este modelo como distribuciones asociadas a los estadísticos de orden.

En primer lugar, vamos a introducir algunas notaciones. Una variable aleatoria X con función de distribución dada por (2.1) se denomina **distribución F exponenciada** y se denota por $X \sim \mathcal{E}(F; a, \mu, \sigma)$. Cuando omitimos los parámetros μ y σ nos referimos al modelo estándar con $\mu = 0$ y $\sigma = 1$. La propuesta alternativa realizada en este capítulo tiene la ventaja de estar definida en términos de la función de distribución y, por tanto, los cálculos y las simulaciones son más sencillas. La distribución propuesta tiene potenciales aplicaciones como distribución de ingresos. Si comenzamos con algunas distribuciones clásicas de ingresos (Pareto, Log-Normal...) obtenemos una nueva distribución con un parámetro de forma adicional, donde es posible obtener expresiones simples de las curvas de Lorenz (LC) y de los índices de Gini. Se pueden obtener expresiones explícitas para la distribución exponencial y la distribución Pareto III. Los resultados generales pueden utilizarse en otras familias de distribuciones de ingresos.

El contenido de este capítulo está organizado de la siguiente manera. La Sección 2 presenta el origen de la distribución y realiza un estudio de las propuestas bibliográficas existentes, distinguiendo entre modelos generales, modelos basados en tiempos de vida y extremos y modelos basados en distribuciones de ingresos. Además, se proponen nuevas clases de distribuciones. En la Sección 3 se analizan varias propiedades probabilísticas, incluyendo la función de densidad, función de hazard, función de supervivencia, cuantiles, las relaciones de orden, los momentos y los estadísticos de orden. En la Sección 4 se proporcionan expresiones explícitas de la curva de Lorenz y

los índices de Gini. En la Sección 5 discutimos algunos procedimientos de inferencia. Finalmente, en la Sección 6 se proponen algunas generalizaciones del modelo básico.

2.2 Origen de las distribuciones max-estables y algunos modelos

La distribución de $X \sim \mathcal{E}(F; n)$ es la distribución del último (máximo) estadístico de orden en una muestra aleatoria de tamaño $a = n$. Como alternativa (2.1) se puede obtener aplicando la transformación integral de la probabilidad o función cuantil de F a una variable aleatoria Beta $(a, 1)$ en lugar de a una variable aleatoria uniforme $(0,1)$. Es interesante observar también que algunas familias de tipo (2.1) surgen como soluciones de ciertas ecuaciones diferenciales. Por ejemplo, la ecuación diferencial $dF(x)/dx = F(x)(1 - F(x))g(x, F(x))$ propuesta por Burr (1942) conduce a varias soluciones de tipo (2.1). La misma situación ocurre con otros sistemas de distribución propuestos en el capítulo.

A continuación vamos a estudiar diferentes modelos. Vamos a distinguir entre modelos generales, modelos basados en tiempos de vida y extremos y modelos basados en distribuciones de ingresos.

2.2.1 Modelos generales

Si F es una distribución uniforme en el intervalo $(0,1)$ obtenemos la función de distribución potencia $G_V(x; a) = x^a$ que es un caso particular de una distribución Beta, Beta $(a, 1)$.

Nota 2.1. Recordemos que si $X \sim \text{Beta}(a, b)$ con $a, b > 0$ la función de densidad de X es

$$f_x(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1}, \quad 0 < x < 1,$$

donde $B(a, b)$ es la función beta.

Una segunda elección natural para F sería la distribución normal estándar con fdD Φ , obteniendo

$$G_{\Phi}(x; a) = [\Phi(x)]^a. \quad (2.3)$$

Este modelo fue analizado por Durrans (1992) en un contexto hidrológico. Se han dado varias generalizaciones del modelo (2.3). Por ejemplo, Balakrishnan (2002) propuso la familia de funciones de densidad ($a > 0$, $\lambda \in \mathbb{R}$)

$$g_{\beta}(x; \lambda, a) = k(\lambda, a)[\Phi(x)]^{a-1}\phi(x). \quad (2.4)$$

Cuando $\lambda = 1$ se obtiene el modelo de Durrans, y cuando $a = 2$ se obtiene la clásica distribución normal sesgada (que es una distribución de probabilidad continua que generaliza a la distribución normal permitiendo que el sesgo estadístico sea distinto de cero). Arnold y Beaver (2002) propusieron la introducción de un segundo parámetro de asimetría. Algunas propiedades de (2.3) para un entero n han sido estudiadas por Gupta y Gupta (2004). Si F es una fdD de una skew normal, obtenemos la familia

$$G_{SN}(x; \lambda, a) = [\Phi(x) - 2T(x, \lambda)]^a, \quad (2.5)$$

donde $T(\cdot, \cdot)$ es la función de Owen. Si $a = 1$ en (2.5) obtenemos la skew normal de Azzalini (Azzalini, 1985) y si $\lambda = 0$ se obtiene el modelo de Durrans. Eugene, Lee y Famoye (2002) han estudiado recientemente una generalización de (2.4).

Los dos siguientes modelos son las versiones logísticas y de Cauchy

$$G_{Lo}(x; a, \mu, \sigma) = (1 + \exp[-(x - \mu)/\sigma])^{-a}, \quad (2.6)$$

$$G_{Ca}(x; a, \mu, \sigma) = \left[\frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left(\frac{x - \mu}{\sigma} \right) \right]^a. \quad (2.7)$$

La distribución (2.6), llamada distribución logística generalizada de tipo I, ha recibido una considerable atención con respecto a sus caracterizaciones, momentos y estadísticos de orden así como las estimaciones de los parámetros en el modelo general, modelo escalar y modelo de forma (ver Balakrishnan and Leung, 1988a,b). Jones y Faddy (2003) y Jones (2004) realizaron un estudio muy amplio de distribuciones más generales basadas en la distribución t de Student

2.2.2 Modelos basados en tiempos de vida y distribuciones extremas

La mayoría de los modelos clásicos para analizar los datos de vida útil solo tienen en cuenta las tasas de fallo crecientes o decrecientes. Este es el caso de la distribución exponencial o la distribución Weibull. Sin embargo, las situaciones de fallo más comunes en economía, ingeniería y ciencias biológicas son las no monótonas que implican formas unimodales o formas de curva de bañera. Este hecho motiva clases de distribuciones con una alta flexibilidad en la función de hazard que contienen distribuciones clásicas como casos particulares. En este contexto, el caso más sencillo corresponde a la distribución exponencial, que nos lleva a la clase de distribuciones

$$G_{\mathcal{E}}(x; a, \mu, \sigma) = [1 - \exp(-(x - \mu)/\sigma)]^a, \quad x \geq \mu. \quad (2.8)$$

La distribución es una re-parametrización de la distribución Gompertz-Verhulst y su origen proviene de la ciencia actuarial, donde se utiliza para comparar tablas de mortalidad humana y para representar el crecimiento de la población. En el ámbito estadístico, (2.8) se denomina distribución exponencial exponenciada, y ha sido estudiada por Gupta et al. (1998) y por Gupta y Gupta (1999, 2001). Se observa que una distribución exponencial exponenciada puede utilizarse para situaciones donde se requiere una distribución sesgada para una variable aleatoria no negativa, y también puede ser una buena alternativa a la distribución Weibull (Gupta y Kundu, 2003). Como extensión de (2.8), si F es una distribución Weibull se obtiene la familia de cuatro parámetros

$$G_W(x; a, \alpha, \mu, \sigma) = [1 - \exp(-[(x - \mu)/\sigma]^\alpha)]^a, \quad x \geq \mu, \quad \alpha > 0, \quad a > 0. \quad (2.9)$$

La distribución (2.9), llamada familia de Weibull exponenciada, ha sido estudiada por Mudholkar y Srivastava (1993) y por Mudholkar, Srivastava y Fremier (1995), en el caso $\mu = 0$. La primera ventaja de esta distribución es la flexibilidad de su función de riesgo. El espacio de los parámetros de forma α , $a > 0$ puede dividirse mediante la línea límite $\alpha = 1$ y la curva $a\alpha = 1$ en cuatro regiones sobre las que la función de hazard es, respectivamente, creciente, forma de curva de bañera, decreciente y unimodal.

Si $\alpha = 2$ en (2.9) se obtiene la distribución Burr tipo X (o exponenciada de Rayleigh), que es quizás una de las distribuciones menos conocidas que pertenece al sistema Burr. Recientemente, la distribución Burr tipo X se ha convertido en un foco de estudio, posiblemente debido a su estrecha relación con la distribución Weibull (Kundu y Raqab, 2005 y Surles y Padgett, 2001, 2005).

En un contexto de distribuciones de valores extremos, que son las distribuciones que modelan valores más extremos que los observados previamente, las versiones para los máximos son obviamente del tipo (2.1) para algunas elecciones de a , μ y σ . Nadarajah (2003, 2005) ha considerado algunas distribuciones exponenciales de valor extremo basadas en la exponenciación mínima, sin embargo estas propuestas no son de la forma (2.1)

2.2.3 Modelos basados en distribuciones de ingresos

La clase de distribuciones (2.1) ha recibido poca atención en el ámbito económico. Si F corresponde a la clásica distribución de Pareto (o distribución Pareto I), Stoppa (1990) propone la clase de distribuciones

$$G_{PI}(x; a, \alpha, \sigma) = \left[1 - \left(\frac{x}{\sigma} \right)^{-\alpha} \right]^a, \quad x \geq \sigma, \alpha > 0, a > 0. \quad (2.10)$$

En comparación con la distribución clásica de Pareto, la distribución de Stoppa es más flexible porque tiene el parámetro de potencia adicional a que permite distribuciones unimodales ($a > 1$) y cero-modales ($a \leq 1$). Recientemente, otras generalizaciones fueron propuestas por Zandonatti (2001) (ver Kleiber y Kotz, 2003).

En este contexto, proponemos dos clases de distribuciones de ingresos. La primera de ellas se basa en la distribución log-normal de tres parámetros. Entonces, tomamos como F una distribución log-normal de tres parámetros, es decir, una variable aleatoria X tal que $Z = \frac{\log(X-\theta) - \zeta}{\sigma}$ sea una distribución normal estándar (Johnson, Kotz and Balakrishnan, (1994)). Entonces, tenemos la distribución log-normal exponenciada de tres parámetros

$$G_{LN}(x; a, \theta, \zeta, \sigma) = \left[\Phi \left(\frac{\log(x - \theta) - \zeta}{\sigma} \right) \right]^a, \quad x > \theta. \quad (2.11)$$

Si elegimos dentro de la jerarquía de Pareto de Arnold (1983), obtenemos clases flexibles de distribuciones. En el caso más general, si F es la distribución Pareto IV, obtenemos la distribución de Pareto IV exponenciada, definida como

$$G_{PIV}(x; a, \alpha, \gamma, \mu, \sigma) = \left(1 - \left[1 + \left(\frac{x - \mu}{\sigma} \right)^{1/\gamma} \right]^{-\alpha} \right)^a, \quad x \geq \mu, \quad \gamma, \alpha, \sigma > 0. \quad (2.12)$$

Si $a = \gamma = 1$ en (2.12) obtenemos la distribución Pareto II; si $a = \alpha = 1$, la distribución Pareto III y si $a = 1$, la distribución Pareto IV. Por último, algunas distribuciones de ingresos son de la forma (2.1), como por ejemplo la distribución Dagum (1977)

$$F(x; a, b, \sigma) = [1 + (x/\sigma)^{-b}]^{-a}, \quad x > 0, \quad a, b > 0, \quad (2.13)$$

que es una subclase de la distribución Beta generalizada del segundo tipo (McDonald, 1984).

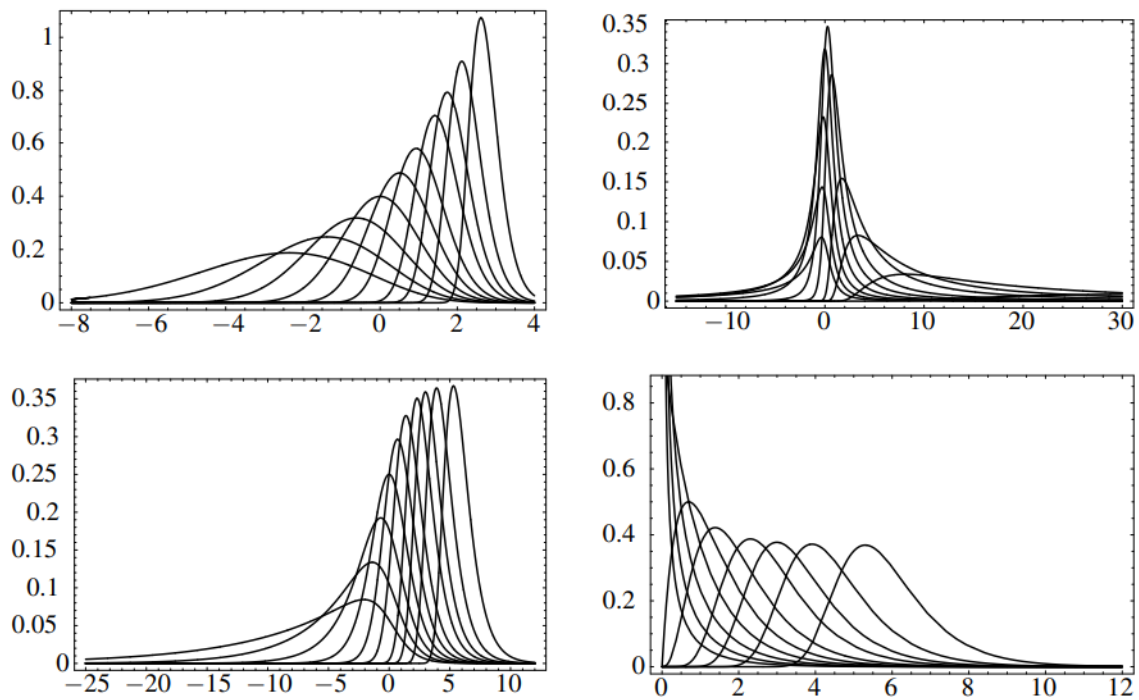


Figura 2.1: Funciones de densidad de una distribución exponenciada $\mathcal{E}(F; a)$ cuando F es una Normal, Cauchy, Logística y Exponencial (de izquierda a derecha y de arriba a abajo), para $a = 0.125, 0.25, 0.5, 1, 2, 4, 10, 20, 50$ y 200 .

La Figura 2.1 muestra algunos ejemplos de funciones de densidad con distribuciones exponenciadas.

2.2.4 Propiedades probabilísticas

En esta sección discutimos algunas propiedades probabilísticas de las anteriores familias de distribuciones

Función de supervivencia, hazard, función cuantil y relaciones de orden

Los resultados de esta sección pueden enunciarse para el modelo general $\mathcal{E}(F; a, \mu, \sigma)$, pero por simplicidad trabajaremos con el modelo normalizado. Si $X \sim \mathcal{E}(F; a)$, el modelo es la solución de la ecuación

$$(a - 1)f(x)^2 + f'(x)F(x) = 0.$$

Las funciones de supervivencia y de hazard vienen dadas por

$$S_{GF}(x; a) = 1 - G_F(x; a) = 1 - [F(x)]^a, \quad (2.14)$$

y

$$h_{GF}(x; a) = \frac{g_F(x; a)}{S_F(x; a)} = \frac{a(1 - F(x))[F(x)]^{a-1}}{1 - [F(x)]^a} h_F(x), \quad (2.15)$$

donde $h_F(x)$ denota la función de hazard de $F(\cdot)$. Los extremos de $h_{GF}(\cdot; a)$ corresponde con la solución de la ecuación

$$f'(x)F(x)(1 - [F(x)]^a) - f^2(x)(1 - a - [F(x)]^a) = 0.$$

La función cuantil viene dada por

$$Q_F(p; a, \mu, \sigma) = G_F^{-1}(p; a, \mu, \sigma) = \mu + \sigma F^{-1}(p^{1/a}), \quad 0 < p < 1, \quad (2.16)$$

y la función de densidad cuantílica $dq_F(p; a, \sigma) = (d/du)Q_F(p; a, \mu, \sigma)$ tiene la forma simple

$$q_F(p; a, \sigma) = \frac{\sigma p^{1/a-1}}{af(F^{-1}(p^{1/a}))}. \quad (2.17)$$

Con respecto a la forma de la densidad de una distribución exponenciada, el siguiente teorema proporciona condiciones suficientes

Teorema 2.1. *La densidad $g_F(x; a)$ de una distribución $\mathcal{E}(F; a)$ cumple lo siguiente:*

- Para $a < 1$, g_F es log-convexa si $f(x)$ es log-convexa y $F(x)$ es log-cóncava.
- Para $a > 1$, g_F es log-cóncava si $f(x)$ es log-cóncava y $F(x)$ es log-cóncava

Demostración. La prueba es directa, porque si f es log-cóncava, F también lo es |

Nota 2.2. Una función no negativa $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ es log-cóncava si su dominio es un conjunto convexo, y si satisface la desigualdad

$$f(\theta x + (1 - \theta)y) \geq f(x)^\theta f(y)^{1-\theta}, \quad \forall x, y \in \text{dom}(f), 0 < \theta < 1.$$

Análogamente, una función es log-convexa si satisface la desigualdad inversa

$$f(\theta x + (1 - \theta)y) \leq f(x)^\theta f(y)^{1-\theta}, \quad \forall x, y \in \text{dom}(f), 0 < \theta < 1.$$

Ejemplo:

Para la distribución exponencial exponenciada introducida en (2.8) con $\mu = 0$ y $\sigma = 1$, se tiene que

$$(d^2/dx^2) \log f(x) = 0$$

y

$$(d^2/dx^2) \log F(x) = -e^x / (1 - e^x)^2,$$

esto es, $F(x)$ es log-cóncava y en consecuencia, (2.8) es log-convexa si $a < 1$ y es log-cóncava si $a > 1$ (Gupta and Kundu, 1999).

| Teorema 2.2. Para la distribución exponenciada $\mathcal{E}(F; a)$, $a > 0$, se tiene que:

- Una función de hazard creciente si $f(x)$ es log-cóncava, $a > 1$ y $F(x)$ es log-cóncava.
- Una función de hazard decreciente si $f(x)$ es log-convexa, $a < 1$ y $F(x)$ es log-cóncava.

Demostración. La prueba se basa en la discusión de la log-convexidad de la función de densidad de Barlow and Proschan (1975). |

Usando este último resultado, la distribución exponencial exponenciada (2.8) tiene una función de hazard creciente si $a > 1$ y decreciente si $a < 1$.

Gupta et al. (1998) han estudiado alguna relación de orden entre la distribución de F y G , incluyendo el ordenamiento de la razón de verosimilitud, el ordenamiento de la tasa de fracaso, el ordenamiento de la vida residual media, el ordenamiento dispersivo y el ordenamiento estadístico. Por ejemplo, decimos que $F \geq_{LR} G$ con respecto a la ordenación de la razón de verosimilitudes si la relación $f(t)/g(t)$ es una función no decreciente de t . En nuestro caso, si $a < 1$, $F \geq_{LR} G_F$ y si $a > 1$, $F \leq_{LR} G_F$.

Momentos

Puesto que la variable aleatoria X se distribuye como $F^{-1}(Y)$ donde $Y \sim \mathcal{B}(a, 1)$ es una distribución beta clásica de parámetros $(a, 1)$, $a > 0$, la función generatriz de momentos es

$$M_X(t) = E(e^{tX}) = a \int_0^1 y^{a-1} \exp\{tF^{-1}(y)\} dy,$$

y en particular

$$E(X^r) = E\{[F^{-1}(Y)]^r\}, \quad r = 1, 2, \dots$$

Este r -ésimo momento coincide con el $(r, a-1, 1)$ -ésimo momento ponderado de probabilidad (Greenwood et al., 1979). Para la distribución exponencial exponenciada introducida en (2.8),

$$M_X(t) = aB(a, 1-t), \quad \text{si } t < 1,$$

donde $B(x, y)$ denota la función Beta,

$$E(X) = \psi(a + 1) - \psi(1),$$

y

$$\psi(x) = (d/dx) \log \Gamma(x)$$

denota la función digamma. Para la distribución de Weibull exponenciada (2.9), Mudholkar y Hutson (1996), Nassar y Eissa (2003) y Nadarajah y Gupta (2005) han proporcionado fórmulas explícitas para los momentos.

Se pueden obtener algunas relaciones de recurrencia en el caso simétrico. Si $F(\cdot)$ tiene función de densidad f simétrica sobre 0, tenemos que

$$\int_{-\infty}^{\infty} \left\{ F(x) - \frac{1}{2} \right\}^{2m+1} f(x)x^{2n} dx = 0; \quad n, m = 0, 1, 2, \dots \quad (2.18)$$

Ahora, denotando $X_a \sim \mathcal{E}(F; a)$ y usando (2.18) obtenemos ($n, m = 0, 1, 2, \dots$)

$$E(X_{2m+2}^{2n}) = \sum_{i=1}^{2m+1} (-1)^{i+1} \binom{2m+1}{i} \frac{m+1}{(2m+2-i)2^{i-1}} E(X_{2m+2-i}^{2n}).$$

Se pueden obtener otro tipo de recurrencias entre momentos utilizando el siguiente teorema

| Teorema 2.3. *Sea $X \sim \mathcal{E}(F; a)$ una distribución exponencial con soporte en $[x_0, x_1]$ y sea $w(\cdot)$ una transformación regular. Entonces, tenemos*

$$E_G[w'(X)] = l(x_1) - l(x_0) - E_G \left[\frac{f'(X)w(X)}{f(X)} \right] - (a-1)E_G \left[\frac{f(X)w(X)}{F(X)} \right],$$

donde $l(x) = g_F(x; a)w(x)$, $|l(x_i)| < \infty$, $i = 1, 2$ y donde E_G significa que la esperanza matemática se toma con respecto a la distribución G y suponemos que todas las esperanzas matemáticas son finitas.

Estadísticos de orden

Sean X_1, \dots, X_n variables aleatorias independientes e idénticamente distribuidas según una $\mathcal{E}(F; a)$. Sea $X_{1:n} \leq \dots \leq X_{n:n}$ los estadísticos de orden. Es evidente que $X_{n:n} \sim$

$\mathcal{E}(F; an)$. La función de densidad del r -ésimo estadístico de orden $X_{r:n}$, $r = 1, \dots, n$ viene dada por

$$\begin{aligned} g_{X_{r:n};F}(x; a) &= \frac{n!}{(r-1)!(n-r)!} a f(x) F(x)^{a-1} F(x)^{a(r-1)} \{1 - F(x)^a\}^{n-r} \\ &= \sum_{j=0}^{n-r} \frac{(-1)^j n!}{(r-1)! j! (n-r-j)(r+j)!} \cdot f_{\mathcal{E}(F; a(r+j))}(x) \end{aligned}$$

En consecuencia, la función de densidad del r -ésimo estadístico de orden es una combinación lineal de las funciones de densidad de $\mathcal{E}(F; a(r+j))$, $j = 0, \dots, n-r$ variables aleatorias. Usando $g_{X_{r:n};F}$, la correspondiente fdD, función de supervivencia y los momentos se pueden obtener de forma sencilla.

Ilustración del efecto de la forma en las medidas de resumen

Para ilustrar el efecto del parámetro de forma a en las medidas de resumen de la distribución, hemos representado gráficamente como función de a unas medidas de localización, dispersión, asimetría y curtosis basadas en cuantiles

- la mediana

$$Q = Q(0.5)$$

- la mitad del rango intercuartil

$$R = (Q(0.75) - Q(0.25))/2$$

- La medida de asimetría de Bowley

$$S = (Q(0.75) - 2Q(0.5) + Q(0.25))/(Q(0.75) - Q(0.25))$$

- Medida de curtosis de Moor (1998) basada en octiles

$$T = ((Q(7/8) - Q(5/8)) - (Q(3/8) - Q(1/8)))/(Q(6/8) - Q(2/8))$$

Estas medidas son menos sensibles a los valores atípicos y existen incluso para distribuciones sin momentos. Hemos considerado estas medidas para las distribuciones

exponenciadas Normal, Cauchy, Logística y Exponencial. Para las cuatro distribuciones exponenciadas, la mediana Q es siempre una función creciente en a . El rango intercuartil medio es decreciente para la normal exponenciada y distribución logística, decreciente para la exponencial y forma de curva de bañera para el caso Cauchy alcanzando un mínimo para $a = 1$. Para las tres distribuciones simétricas (Normal, Cauchy y Logística), el coeficiente de Bowley es negativo para $a < 1$ y positivo para $a > 1$. En el caso exponencial S es una función decreciente de a . El coeficiente de Moor presenta un comportamiento diferente para las cuatro distribuciones exponenciadas: crece en el caso normal, decreciente en el caso exponencial, forma de curva de bañera para el caso Cauchy, y con un máximo para el caso logístico. La representación gráfica de estas medidas aparece en la siguiente figura.

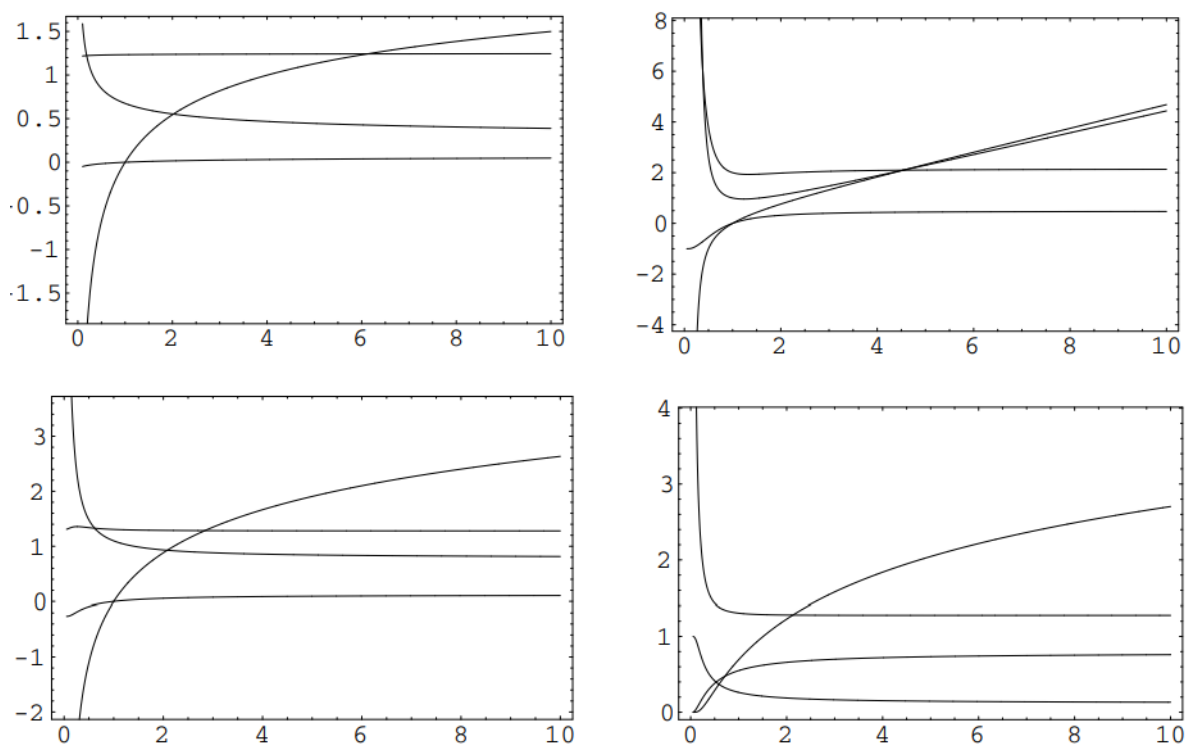


Figura 2.2: Mediana, mitad del rango intercuartil, la medida de asimetría de Bowley, medida de curtosis de Moor de una distribución exponenciada $\mathcal{E}(F; a)$ cuando F es una variable aleatoria Normal, Cauchy, Logística y Exponencial (de izquierda a derecha, de arriba a abajo) en función de a .

Dominios de atracción

Dado que la exponenciación de las funciones de distribución es estable con respecto a los dominios máximos de atracción, el máximo dominio de atracción de (2.1) coincide con el de $F(\cdot)$.

2.2.5 Medidas de desigualdad

Las herramientas básicas para el estudio de las distribuciones de ingresos son las curvas de Lorenz junto con las medidas de desigualdad asociadas, especialmente los índices de Gini. Supongamos una variable aleatoria no negativa $X \sim \mathcal{E}(F; a)$ con una esperanza finita. Usando la conocida fórmula de la curva de Lorenz obtenemos

$$L_F(p; a) = \frac{\int_0^{p^{1/a}} ay^{a-1} F^{-1}(y) dy}{\int_0^1 ay^{a-1} F^{-1}(y) dy}, \quad 0 \leq p \leq 1. \quad (2.19)$$

En relación con el índice de Gini tenemos el siguiente teorema

| Teorema 2.4. *Sea X una variable aleatoria no negativa $\mathcal{E}(F; a)$ con esperanza matemática finita μ_X . Entonces, el índice de Gini se puede obtener mediante la expresión*

$$G(L_F) = \frac{1}{\mu_X} \int_0^1 \left\{ \frac{u^a - u^{2a}}{f(F^{-1}(u))} \right\} du \quad (2.20)$$

Demostración. Para la existencia del índice de Gini, asumimos que $\mu_X < \infty$. Usando la conocida fórmula para el índice de Gini tenemos que

$$G(L_F) = \frac{1}{\mu_X} \int_0^\infty G_F(x; a) [1 - G_F(x; a)] dx = \frac{1}{\mu_X} \int_0^\infty \{ [F(x)]^a - [F(x)]^{2a} \} dx$$

y haciendo el cambio de variable $\mu = F(x)$ obtenemos lo que queremos. **|**

La distribución exponencial exponenciada

Un caso importante corresponde a la distribución exponencial exponenciada (2.8), con $\mu = 0$ y $\sigma = 1$. Usando la expresión (2.19) tenemos que la curva de Lorenz es

$$L(p; a) = \frac{-p^{1/a+1}\Psi(p^{1/a}, 1, 1+a) - p \log(1 - p^{1/a})}{\psi(a+1) - \psi(1)}, \quad (2.21)$$

donde $\Psi(z, s, a)$ denota la función trascendente de Lerch definida por $\Psi(z, s, a) = \sum_{k=0}^{\infty} \frac{z^k}{(a+k)^s}$, donde se excluye cualquier término con $k+a=0$. Si a es un entero, una expresión alternativa es

$$L(p; a) = \frac{a \sum_{j=0}^{a-1} \frac{(-1)^j}{(j+1)^2} \binom{a-1}{j} \gamma(2, -(j+1) \log(1 - p^{1/a}))}{\psi(a+1) - \psi(1)}, \quad (2.22)$$

donde $\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$ denota la función gamma incompleta. Usando (2.21) y (2.22), algunos ejemplos de curvas de Lorenz son

$$\begin{aligned} L(p; 1/2) &= \frac{-2p - (1-p) \log(1-p) + (1+p) \log(1+p)}{-2 + \log 4}, \\ L(p; 2) &= \frac{1}{3} [2\sqrt{p} + p + 2(1-p) \log(1 - \sqrt{p})], \\ L(p; 3) &= \frac{1}{11} [6p^{1/3} + 3p^{2/3} + 2p + 6(1-p) \log(1 - p^{1/3})]. \end{aligned}$$

Usando (2.20), el índice de Gini, en términos de la función psi, tiene la forma

$$G(a) = \frac{\psi(2a+1) - \psi(a+1)}{\psi(a+1) - \psi(1)}.$$

La distribución de Pareto exponenciada

Se pueden obtener resultados explícitos para la familia de Pareto exponenciada. Por ejemplo, en la distribución Pareto III normalizada exponenciada ($a, \gamma > 0$)

$$G_{PIII}(x; a, \gamma) = \left[1 - (1 + x^{1/\gamma})^{-1} \right]^a; \quad x > 0,$$

la curva de Lorenz tiene la forma ($a > \gamma$)

$$L(p; a, \gamma) = 1 - \frac{B_{1-p^{1/a}}(a - \gamma, \gamma + 1)}{B(a - \gamma, \gamma + 1)}; \quad 0 < p < 1 \quad (2.23)$$

donde $B_x(p, q) = \int_0^x u^{p-1}(1-u)^{q-1} du$ denota la función gamma incompleta. La expresión (2.23) extiende el resultado de Arnold (1983). Para la distribución exponenciada Pareto IV, si a es un entero, la curva de Lorenz viene dada por

$$L_{PIV}(p; a, \alpha, \gamma) = 1 - \frac{\sum_{j=0}^{a-1} \binom{a-1}{j} B_{(1-p^{1/a})^{1/\alpha}}(\alpha(j+1) - \gamma, \gamma + 1)}{\sum_{j=0}^{a-1} \binom{a-1}{j} B(\alpha(j+1) - \gamma, \gamma + 1)}.$$

Se pueden obtener otros resultados explícitos que cubren una amplia gama de formas funcionales. Por ejemplo si $a = 1/2$, $\alpha = 2$ y $\gamma = 1$ obtenemos la curva de Lorenz

$$L(p; 1/2, 2, 1) = \frac{2[\arcsin(p) - p]}{\pi - 2}$$

2.2.6 Inferencia

Sea x_1, \dots, x_n una muestra de tamaño n de una distribución $\mathcal{E}(F; a, \mu, \sigma)$ con $\sigma = 1/\lambda$. Si estudiamos la estimación de máxima verosimilitud, la función de log-verosimilitud viene dada por

$$\ell(a, \mu, \lambda) = n \log a + n \log \lambda + \sum_{i=1}^n \log f(z_i) + (a-1) \sum_{i=1}^n \log F(z_i) \quad (2.24)$$

donde $z_i = \lambda(x_i - \mu)$, $i = 1, 2, \dots, n$ y donde $\mu < \min\{x_1, \dots, x_n\}$ si $F(\cdot)$ es no negativa. Tomando derivadas parciales con respecto a a , μ y λ e igualando a 0 obtenemos las siguientes ecuaciones

$$\begin{aligned}\frac{\partial \ell}{\partial a} &= \frac{n}{a} + \sum_{i=1}^n \log F(z_i) = 0, \\ \frac{\partial \ell}{\partial \mu} &= -\lambda \sum_{i=1}^n \frac{f'(z_i)}{f(z_i)} - \lambda(a-1) \sum_{i=1}^n \frac{f(z_i)}{F(z_i)} = 0, \\ \frac{\partial \ell}{\partial \lambda} &= \frac{n}{\lambda} + \frac{1}{\lambda} \sum_{i=1}^n \frac{z_i f'(z_i)}{f(z_i)} + \frac{a-1}{\lambda} \sum_{i=1}^n \frac{z_i f(z_i)}{F(z_i)} = 0.\end{aligned}$$

Si sustituimos a en términos de μ y λ , estas tres ecuaciones se pueden reducir a dos, y entonces tenemos que utilizar un algoritmo iterativo o el método de Newton-Raphson para resolver simultáneamente las dos ecuaciones no lineales. Cuando $a > 1$, normalmente el máximo se alcanza lejos del límite. Cuando $a > 2$ la distribución exponencial exponenciada (2.8) satisface todas las condiciones de regularidad, como la distribuciónes gamma y Weibull (Gupta and Kundu, 1999). Cuando $a \leq 2$ los resultados asintóticos son desconocidos. Para esta distribución se han considerado también otros métodos de estimación como los estimadores de percentiles, los estimadores basados en estadísticos de orden, los estimadores de mínimos cuadrados, los estimadores de mínimos cuadrados ponderados y los estimadores basados en L -momentos (Gupta y Kundu, 2001; Raqab y Ahsanullah, 2001). Estos autores recomendaron utilizar estimadores basados en percentiles para tamaños de muestra pequeños y estimadores de máxima verosimilitud para tamaños de muestra moderados o grandes. Otras investigaciones se han realizado para el modelo exponenciado de Rayleigh (Bur tipo X) y Weibull por Surlless y Padgett (2001), Kundu y Raqab (2005) y Mudholkar, Srivastava y Freimer (1995). La estimación bayesiana para la distribución Weibull exponenciada ha sido considerada por Nassar y Eissa (2004).

Problemas de discriminación (utilizando la relación de las probabilidades maximizadas) entre la distribución exponencial exponenciada y Weibull y log-normal han sido considerados por Gupta y Kundu (2003) y Kundu, Gupta y Manglick (2005).

El método del percentil elemental

Los estimadores de máxima verosimilitud y los métodos de estimación basados en momentos pueden tener problemas, ya sea porque los rangos de las distribuciones dependen del parámetro μ (para variables aleatorias no negativas) o bien porque los momentos no existen en ciertas regiones del espacio de los parámetros (por ejemplo

para la distribución (2.10)). En esta sección describimos el método del percentil elemental (EPM) propuesto por Castillo y Hadi (1995a) y lo adaptamos al caso de las distribuciones exponenciadas. El EPM ha sido utilizado para la estimación de la distribución de valores extremos generalizada, la distribución de Pareto generalizada y otras distribuciones que tienen problemas similares a los mencionados para la distribución exponenciada (Castillo y Hadi, 1995b,c, 1997). Este método se ha utilizado también para la estimación de las curvas de Lorenz (Castillo, Hadi y Sarabia, 1998).

Sea x_1, \dots, x_n una muestra aleatoria de una distribución $\mathcal{E}(F; a, \mu, \sigma)$. El método obtiene las estimaciones en dos pasos. Las estimaciones iniciales de los parámetros se obtienen de la siguiente manera. Como tenemos que estimar tres parámetros, los estimadores elementales se basan en tres estadísticos de orden diferentes. Sea $I = \{i, j, r\}$, $i < j < r \in \{1, 2, \dots, n\}$ los índices de los tres estadísticos de orden diferentes. Igualando los cuantiles de la muestra y teóricos $\mathcal{E}(F; a, \mu, \sigma)$ en (2.16) obtenemos

$$\begin{aligned} x_{i:n} &= \mu + \sigma F^{-1} \left(p_{i:n}^{1/a} \right) \\ x_{j:n} &= \mu + \sigma F^{-1} \left(p_{j:n}^{1/a} \right) \\ x_{r:n} &= \mu + \sigma F^{-1} \left(p_{r:n}^{1/a} \right) \end{aligned} \quad (2.25)$$

donde $p_{i:n} = i/(n+1)$. Eliminando μ y σ en este sistema de tres ecuaciones obtenemos

$$\frac{x_{r:n} - x_{j:n}}{x_{r:n} - x_{i:n}} = \frac{F^{-1} \left(p_{r:n}^{1/a} \right) - F^{-1} \left(p_{j:n}^{1/a} \right)}{F^{-1} \left(p_{r:n}^{1/a} \right) - F^{-1} \left(p_{i:n}^{1/a} \right)} = \frac{1 - u_{jr}(F, a)}{1 - u_{ir}(F, a)}, \quad (2.26)$$

donde $u_{jr}(F, a) = F^{-1} \left(p_j^{1/a} \right) / F^{-1} \left(p_r^{1/a} \right)$. Se trata de una ecuación con una sola incógnita a . Resolviendo (2.26) para a usando el método de bisección, obtenemos \hat{a}_{ijr} , lo que indica que esta estimación es una función de las tres observaciones $x_{i:n}$, $x_{j:n}$ y $x_{r:n}$. La implementación de esta parte depende de cada distribución $F(\cdot)$. Sustituyendo \hat{a}_{ijr} en dos de las ecuaciones de (2.25) y resolviendo para σ y μ obtenemos

$$\hat{\sigma}_{ijr} = \frac{x_{i:n} - x_{r:n}}{F^{-1} \left(p_{i:n}^{1/\hat{a}_{ijr}} \right) - F^{-1} \left(p_{r:n}^{1/\hat{a}_{ijr}} \right)}, \quad (2.27)$$

y

$$\hat{\mu}_{ijr} = x_{i:n} - \hat{\sigma}_{ijr} F^{-1} \left(p_{i:n}^{1/\hat{a}_{ijr}} \right). \quad (2.28)$$

Finalmente, se obtienen estimaciones más eficientes y consistentes desde el punto de vista estadístico utilizando otros estadísticos de orden, como se indica a continuación. Seleccionar un número preestablecido N de subconjuntos elementales cada uno de ellos de tamaño tres, ya sea al azar o utilizando todos los subconjuntos posibles. Para cada uno de estos subconjuntos elementales, calculamos una estimación elemental de los parámetros $\theta = \{a, \mu, \sigma\}$. Denotamos estas estimaciones elementales por $\hat{\theta}_1, \dots, \hat{\theta}_N$. Estas estimaciones elementales pueden combinarse utilizando alguna función consistente adecuada para obtener una estimación final de θ . Algunos ejemplos de funciones consistentes son la mediana de la media α -recortada (TM_α), donde α denota el porcentaje de recorte. Una estimación final de $\{a, \mu, \sigma\}$ se puede obtener como

$$\begin{aligned} \hat{a}_{\text{MED}} &= \text{Mediana} (\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N), \\ \hat{\mu}_{\text{MED}} &= \text{Mediana} (\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_N), \\ \hat{\sigma}_{\text{MED}} &= \text{Mediana} (\hat{\sigma}_1, \hat{\sigma}_2, \dots, \hat{\sigma}_N), \end{aligned}$$

o de forma alternativa

$$\begin{aligned} \hat{a}_{\text{TM}_\alpha} &= \text{TM}_\alpha (\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N), \\ \hat{\mu}_{\text{TM}_\alpha} &= \text{TM}_\alpha (\hat{\mu}_1, \hat{\mu}_2, \dots, \hat{\mu}_N), \\ \hat{\sigma}_{\text{TM}_\alpha} &= \text{TM}_\alpha (\hat{\sigma}_1, \hat{\sigma}_2, \dots, \hat{\sigma}_N), \end{aligned}$$

donde $\text{Mediana}(x_1, \dots, x_N)$ es la mediana de los datos $\{x_1, \dots, x_N\}$ y $\text{TM}_\alpha(x_1, \dots, x_N)$ es la mediana obtenido después de recortar el $(\alpha/2)\%$ más grande y el $(\alpha/2)\%$ más pequeño de los estadísticos de orden de x_1, \dots, x_N . Los estimadores de la Mediana son muy consistentes pero ineficientes y los estimadores TM_α son menos consistentes

pero más eficientes que los anteriores. Cuanto más grande sea el recorte, más consistentes y menos eficientes serán los estimadores TM_α . La experiencia (Castillo et al., 2005) indica que las propiedades de los estimadores de la Mediana y TM_α son similares, pero los errores cuadráticos medios de los estimadores TM_α pueden ser ligeramente más pequeños que los de los estimadores de la Mediana.

2.2.7 Generalizaciones

Consideramos algunas extensiones del modelo básico suponiendo que el parámetro a no es constante y varía según alguna distribución de probabilidad conocida. Consideramos el modelo

$$X | a \sim \mathcal{E}(F; a), \quad (2.29)$$

$$a \sim \mathcal{G}a(\alpha, \lambda), \quad (2.30)$$

donde $\mathcal{G}a(\alpha, \lambda)$ denota una variable aleatoria Gamma clásica con fdd proporcional a $u^{\alpha-1}e^{-\lambda u}$. La composición de $\mathcal{E}(F; a)$ con una Gamma es

$$G_F(x; \alpha, \lambda) = \int_0^\infty [F(x)]^a f_{\mathcal{G}a(\alpha, \lambda)}(a) da = \frac{1}{[1 - \lambda^{-1} \log F(x)]^\alpha}. \quad (2.31)$$

Sarabia et al. (2005) consideró un modelo similar para especificar las curvas de Lorenz mixtas. Se puede obtener una segunda extensión si el parámetro a se distribuye según una Inversa Gaussiana con fdd (ver Johnson, Kotz y Balakrishnan, 1994)

$$f(x; \phi, \lambda) = \left[\frac{\lambda}{2\pi x^3} \right]^{1/2} e^\phi \exp \left\{ -\frac{1}{2} \left(\frac{\phi^2 x}{\lambda} + \frac{\lambda}{x} \right) \right\}. \quad (2.32)$$

Utilizando de nuevo el esquema (2.29)-(2.30) con la distribución Inversa Gaussiana (2.32) obtenemos

$$G_F(x; \phi, \lambda) = \exp \left\{ \phi - \sqrt{\phi^2 - 2\lambda \log F(x)} \right\}. \quad (2.33)$$

Obsérvese que (2.31) y (2.33) son de nuevo max-estables. Un tercer tipo de distribuciones se obtiene suponiendo que el parámetro a toma valores $\{r, r+1, \dots\}$ según una distribución binomial negativa con función de masa de probabilidad

$$\pi_j = \binom{j-1}{r-1} p^r (1-p)^{j-r}, \quad j = r, r+1, \dots$$

La nueva distribución viene dada por

$$G_F(x; r, p) = \sum_{j=r}^{\infty} [F(x)]^j \pi_j = \left[\frac{pF(x)}{1 - (1-p)F(x)} \right]^r. \quad (2.34)$$

Cuando $r = 1$, obtenemos una distribución estudiada por Marshall y Olkin (1997). Notemos que (2.34) puede verse como una distribución binomial negativa max-estable.

3 | Modelo Exponencial Exponenciado

En este capítulo vamos a estudiar la distribución exponencial exponenciada (EE) que es un caso particular de las distribuciones exponenciadas. Vamos a presentar este modelo como alternativa a las distribuciones Weibull y gamma. Para ello vamos a comparar las tres distribuciones y ver cuales son las ventajas y desventajas de cada una de ellas. Se van a estudiar propiedades básicas de la distribución EE, hacer inferencia sobre sus parámetros, vamos a realizar un análisis de datos de las tres distribuciones con dos conjuntos de datos para ver cuál se ajusta mejor y, finalmente, un experimento numérico en el que se dan diferentes tamaños de muestra y veremos como se comporta la distribución EE.

3.1 Definición del modelo

Definición 3.1. Sea X una variable aleatoria no negativa. Decimos que X sigue una **distribución exponencial exponenciada** de parámetros α y λ con $\alpha > 0$ y $\lambda > 0$, es decir, se denota como $X \sim EE(\alpha, \lambda)$ si y sólo si la Función de Distribución de X es

$$F(x; \alpha, \lambda) = (1 - \exp(-\lambda x))^\alpha, \quad x > 0, \alpha > 0, \lambda > 0. \quad (3.1)$$

En este caso, la función de densidad es

$$f(x; \alpha, \lambda) = F'(x; \alpha, \lambda) = \alpha \lambda (1 - \exp(-\lambda x))^{\alpha-1} \exp(-\lambda x), \quad x > 0, \alpha > 0, \lambda > 0 \quad (3.2)$$

Por tanto, su función de supervivencia $S(x; \alpha, \lambda)$ viene dada por

$$S(x; \alpha, \lambda) = 1 - F(x; \alpha, \lambda) = 1 - (1 - \exp(-\lambda x))^\alpha, \quad x > 0, \alpha > 0, \lambda > 0.$$

Añadimos ahora una gráfica con la función de densidad de una distribución exponencial exponenciada para λ constante y diferentes valores de α .

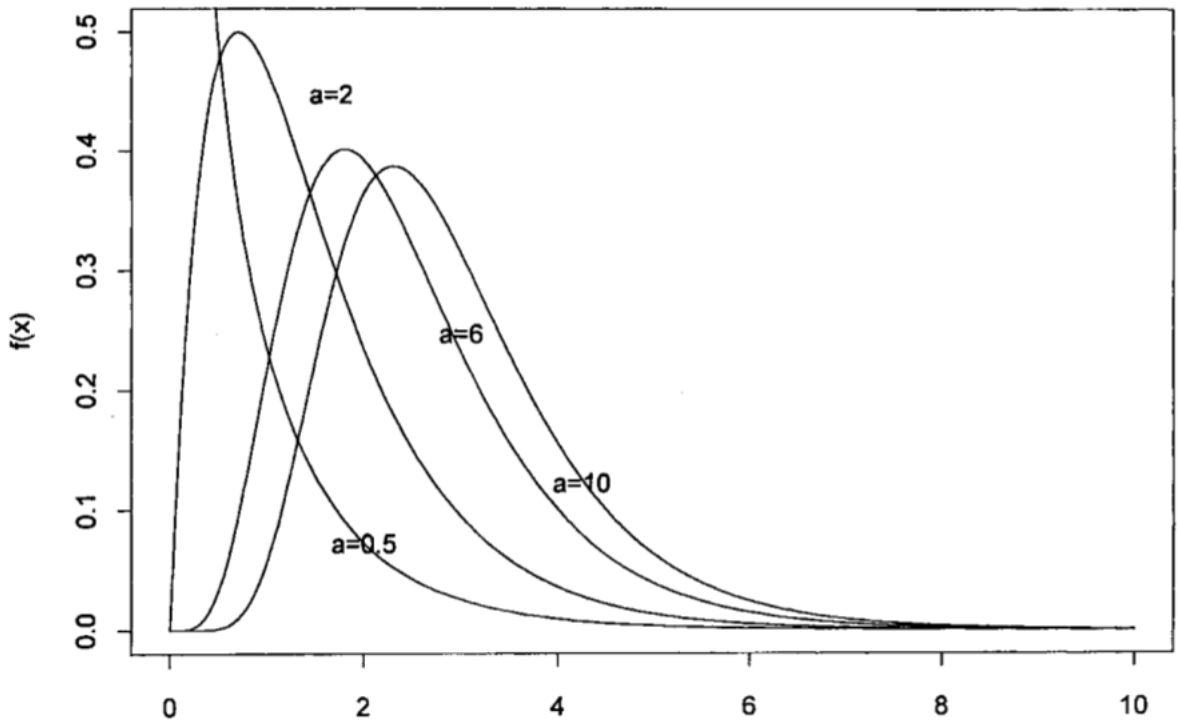


Figura 3.1: fdd de EE para diferentes valores de α

Definición 3.2. Se define la función de hazard o función de riesgo como

$$h(x; \alpha, \lambda) = \frac{f(x; \alpha, \lambda)}{S(x; \alpha, \lambda)} = \frac{\alpha \lambda (1 - \exp(-\lambda x))^{\alpha-1} \exp(-\lambda x)}{1 - (1 - \exp(-\lambda x))^\alpha}, \quad x > 0, \alpha > 0, \lambda > 0.$$

Como antes, añadimos una gráfica con la función de hazard para diferentes valores de α y λ constante ($\lambda = 1$).

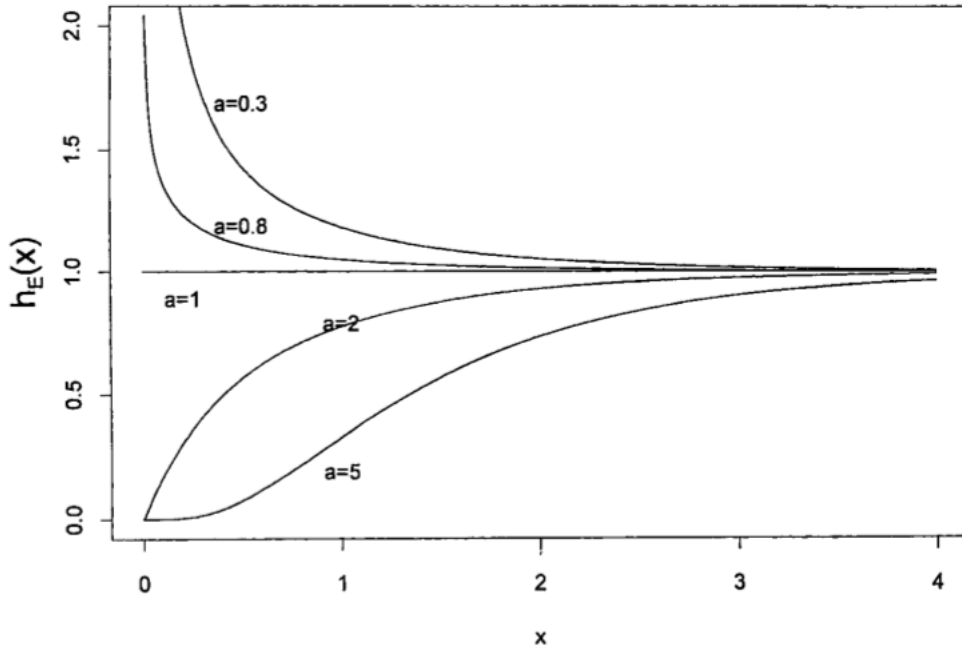


Figura 3.2: Función de hazard de EE para diferentes valores de α

Observación 3.1. Cuando $\alpha = 1$, la distribución exponencial exponenciada representa la distribución exponencial.

3.2 Algunas propiedades básicas

Proposición 3.1 (Relación con la distribución exponencial). Sean X_1, X_2, \dots, X_n variables aleatorias i.i.d con Función de Distribución común $X_i \sim \text{Exp}(\lambda)$. Entonces

$$X_{(n)} = \max_{1 \leq i \leq n} X_i$$

sigue una distribución exponencial exponenciada de parámetros n y λ , $X_{(n)} \sim EE(n, \lambda)$.

Demostración.

$$\begin{aligned} F_{X_{(n)}}(x) &= P(X_{(n)} \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = P(X_1 \leq x) \cdots P(X_n \leq x) = \\ &= (F_X(x))^n = (1 - \exp(-\lambda x))^n. \end{aligned}$$

Definición 3.3. *Un sistema en serie es una configuración en la que las componentes están conectadas sucesivamente, de tal forma que el fallo de una de ellas provoca el fallo del sistema, es decir, un sistema en serie funciona si y sólo si todas sus componentes funcionan. Definimos la probabilidad de que una componente falle en $[0, t]$ como*

$$P[T_i < t], \quad i = 1, \dots, n.$$

Por tanto, la probabilidad de que un sistema en serie funcione viene dada por

$$P[\text{ninguna falle}] = (P[T_i > t])^n, \quad t > 0.$$

Definición 3.4. *Un sistema en paralelo es una configuración en la que las componentes están colocadas de tal manera que los terminales de entrada y los terminales de salida coinciden entre sí. Un sistema en paralelo funciona cuando funciona alguna de sus componentes. Como antes, la probabilidad de que un sistema en paralelo funcione viene dada por*

$$P[\text{al menos una funcione}] = 1 - P[\text{todas fallen}] = 1 - (P[T_i < t])^n = 1 - (F(t))^n, \quad t > 0.$$

Nota 3.1. Hallamos la función de supervivencia de un sistema con n componentes e idénticamente distribuidas.

$$S_{\text{sistema}}(t) = P[\text{sistema funcione en } t]$$

- En Serie: $S_{\text{sistema}}(t) = P[T_1 > t, \dots, T_n > t] = (P[T_i > t])^n = (1 - F(t))^n.$
- En Paralelo: $S_{\text{sistema}}(t) = 1 - (F(t))^n.$

donde $F(t)$ es la Función de Distribución de las componentes.

3.2.1 Comparación del modelo EE con las distribuciones Gamma y Weibull

Definición 3.5. La función de hazard del modelo se define como:

$$h_E(x; \alpha, \lambda) = \frac{\alpha \lambda (1 - \exp(-\lambda x))^{\alpha-1} \exp(-\lambda x)}{1 - (1 - \exp(-\lambda x))^\alpha}, \quad x > 0, \alpha > 0, \lambda > 0.$$

Aquí α es el parámetro de forma y λ es el parámetro escalar. Las tres familias, llamadas Gamma, Weibull y EE, son una generalización de la familia exponencial pero de diferentes maneras. La distribución EE tiene una buena interpretación física. Supongamos que hay n componentes en un sistema paralelo y que las distribuciones de los tiempos de vida de cada componente son independientes e idénticamente distribuidas. Si la distribución del tiempo de vida de cada componente es exponencial, entonces la distribución del tiempo de vida del sistema es EE. Al contrario que la distribución Weibull, que representa un sistema en serie, la distribución EE representa un sistema en paralelo. La función de hazard de EE se muestra en la Figura 1 con $\lambda = 1$. Para cualquier λ , la función de hazard es una función no decreciente si $\alpha > 1$, y es una función no creciente si $\alpha < 1$. Para $\alpha = 1$, es constante.

Vamos a ver una comparación de la función de hazard del modelo EE y los modelos gamma y Weibull. Para ello, definamos las funciones de densidad de la distribución gamma (f_G) y la distribución Weibull (f_W):

$$f_G(x; \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\lambda x); \quad \alpha, \lambda, x > 0,$$

$$f_W(x; \alpha, \lambda) = \alpha \lambda (x\lambda)^{\alpha-1} \exp(-(\lambda x)^\alpha); \quad \alpha, \lambda, x > 0.$$

Ahora sí, veamos la comparación en la siguiente tabla:

Parámetros	Gamma	Weibull	EE
$\alpha = 1$	Constante	Constante	Constante
$\alpha > 1$	Crece de 0 a λ	Crece de 0 a ∞	Crece de 0 a λ
$\alpha < 1$	Decrece de ∞ a λ	Decrece de ∞ a 0	Decrece de ∞ a λ

Cuadro 3.1: Funciones de hazard

Por tanto, la función de hazard de la distribución EE se comporta como la función de hazard de la distribución gamma, que es muy diferente de la función de hazard de la distribución Weibull. Para la distribución Weibull, si $\alpha > 1$, la función de hazard crece desde 0 a ∞ , y si $\alpha < 1$, la función de hazard decrece desde ∞ a 0. Vamos a ver ahora que pasa si el parámetro de forma permanece constante. Ya que $f_G(x)/f_E(x)$ es una función creciente para $\alpha > 1$ y decreciente para $\alpha < 1$ podemos decir que la distribución gamma es más grande (pequeña) que la distribución EE en términos de ordenación de la razón de probabilidad si $\alpha > 1$ ($\alpha < 1$) y son iguales si $\alpha = 1$. Es especialmente interesante el caso en el que $\alpha = 1$. Cuando $\alpha = 1$ las tres distribuciones se asemejan a una distribución exponencial. Es fácil observar que no hay un ordenamiento de la razón de probabilidad entre Weibull y gamma o entre Weibull y EE. Pero se puede observar lo siguiente para todos los valores de x ;

$$h_W(x) - h_E(x) \geq 0 \quad \text{si } \alpha > 1,$$

$$h_W(x) - h_E(x) \leq 0 \quad \text{si } \alpha < 1.$$

Dado que existe una relación entre las funciones de hazard de la distribución EE y la distribución gamma, obtenemos que

$$h_W(x) \geq h_E(x) \geq h_G(x) \quad \text{si } \alpha > 1,$$

$$h_W(x) \leq h_E(x) \leq h_G(x) \quad \text{si } \alpha < 1.$$

Por tanto, tenemos una ordenación de las funciones de distribución de los tres:

$$F_W(x) \geq F_E(x) \geq F_G(x) \quad \text{si } \alpha > 1,$$

$$F_W(x) \leq F_E(x) \leq F_G(x) \quad \text{si } \alpha < 1.$$

Muchos autores señalan que, dado que la función de hazard de una distribución EE (para $\alpha > 1$) crece desde 0 a una constante finita, la distribución EE puede ser más apropiada como modelo de población cuando los individuos de la población están en un programa de mantenimiento regular. La misma observación es válida para la distribución gamma.

Por tanto, si se sabe que los datos proceden de un entorno de mantenimiento regular, puede tener más sentido ajustar los datos mediante una distribución EE o gamma que usar una distribución Weibull.

| Definición 3.6. Sea X una variable aleatoria. La función

$$M_X(t) = E(\exp(tX))$$

se conoce como la función generatriz de momentos (FGM) de X . Esta función debe existir en algún entorno del origen $\forall t \in (-t_0, t_0)$ con $t_0 > 0$.

| Teorema 3.1 (De unicidad). Si existe la función generatriz de momentos de una variable aleatoria, determina unívocamente su distribución de probabilidad.

| Teorema 3.2. Si existe la función generatriz de momentos de una variable aleatoria, $M_X(t)$, con $t \in (-t_0, t_0)$, se tiene:

1. $\exists E(X^k), \forall k \in \mathbb{N}$
2. $M_X(t) = \sum_{k=0}^{\infty} \frac{E(X^k)}{k!} t^k$
3. $E(X^k) = M_X^{(k)}(t)|_{t=0}$

| Definición 3.7. Definimos la función gamma como:

$$\Gamma(p) = \int_0^{\infty} x^{p-1} \exp(-x) dx, \quad p > 0$$

Algunas propiedades de la función gamma son:

- $\Gamma(p) = (p-1)\Gamma(p-1)$
- Si $p \in \mathbb{Z}^+$ entonces $\Gamma(p) = (p-1)!$
- $\Gamma(1/2) = \sqrt{\pi}$

| Definición 3.8. Definimos la función beta como:

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx, \quad a > 0, b > 0$$

La relación entre la función gamma y la función beta es la siguiente:

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad a > 0, b > 0$$

Proposición 3.2 (Propiedades). Si $X \sim EE(\alpha, \lambda)$ se verifica

1. Momento no central de orden k :

$$E(X^k) = \frac{\alpha \Gamma(k+1)}{\lambda^k} \sum_{i=0}^{\infty} (-1)^i c(\alpha-1, i) \frac{1}{(i+1)^{k+1}}, \quad (3.3)$$

2. Función generatriz de momentos:

$$M(t) = \alpha \int_0^1 (1-y)^{\alpha-1} y^{-\frac{t}{\lambda}} dy = \frac{\Gamma(\alpha) \Gamma(1 - \frac{t}{\lambda})}{\Gamma(\alpha - \frac{t}{\lambda} + 1)}, \quad (3.4)$$

la FGM esta definida para $0 < t < \lambda$

3. Media:

$$\frac{1}{\lambda} (\psi(\alpha+1) - \psi(1)) \quad (3.5)$$

4. Varianza:

$$\frac{1}{\lambda^2} (\psi'(1) - \psi'(\alpha+1)) \quad (3.6)$$

Demostración. 1. Supongamos que X sigue una distribución EE con parámetros α y λ , entonces

$$E(X^k) = \alpha \lambda \int_0^{\infty} x^k (1 - \exp(-\lambda x))^{\alpha-1} \exp(-\lambda x) dx.$$

Ahora bien, como $0 < \exp(-\lambda x) < 1$, para $\lambda > 0$ y $x > 0$, usando la representación en serie (finita o infinita) de $(1 - \exp(-\lambda x))^{\alpha-1}$

$$(1 - \exp(-\lambda x))^{\alpha-1} = \sum_{i=0}^{\infty} (-1)^i c(\alpha-1, i) \exp(-i\lambda x),$$

donde $c(\alpha-1, i) = \frac{(\alpha-1) \cdots (\alpha-i)}{i!}$, obtenemos que

$$E(X^k) = \frac{\alpha \Gamma(k+1)}{\lambda^k} \sum_{i=0}^{\infty} (-1)^i c(\alpha-1, i) \frac{1}{(i+1)^{k+1}}.$$

Dado que (3.3) es una serie convergente para cualquier $k > 0$, todos los momentos existen y para valores enteros de α , (3.3) puede representarse como una serie finita.

2. La función generatriz de momentos, $M_X(t)$, se puede escribir como

$$M(t) = E(\exp(tX)) = \alpha \lambda \int_0^{\infty} (1 - \exp(-\lambda x))^{\alpha-1} \exp((t - \lambda)x) dx$$

Haciendo el cambio de variables $y = \exp(-\lambda x)$, se reduce a

$$M(t) = \alpha \int_0^1 (1 - y)^{\alpha-1} y^{-\frac{t}{\lambda}} dy = B\left(1 - \frac{t}{\lambda}, \alpha\right) = \frac{\Gamma(\alpha) \Gamma\left(1 - \frac{t}{\lambda}\right)}{\Gamma\left(\alpha - \frac{t}{\lambda} + 1\right)}.$$

Ahora bien, como la función beta está definida para valores positivos, tenemos que $1 - \frac{t}{\lambda} > 0 \Leftrightarrow 1 > \frac{t}{\lambda} \Rightarrow t < \lambda$

3 y 4. Diferenciando $\ln(M(t))$ y evaluando en $t = 0$, conseguimos la media y la varianza de X como

$$E(X) = \frac{1}{\lambda} (\psi(\alpha + 1) - \psi(1)) \quad y \quad var(X) = \frac{1}{\lambda^2} (\psi'(1) - \psi'(\alpha + 1))$$

|

Nota 3.2. $\psi(\cdot)$ es la función digamma y $\psi'(\cdot)$ su derivada. Los momentos centrales superiores pueden obtenerse en términos de las funciones poligamma.

3.3 Inferencia en la distribución EE

En esta sección vamos a discutir los estimadores de máxima verosimilitud (EMV) de una distribución EE de dos parámetros y sus propiedades asintóticas. Sea x_1, \dots, x_n una muestra de EE, entonces, usando (3.2), la función de verosimilitud puede escribirse como:

$$L(\alpha, \lambda) = \prod_{i=1}^n f(x_i; \alpha, \lambda) = \prod_{i=1}^n \alpha \lambda (1 - \exp(-\lambda x_i))^{\alpha-1} \exp(-\lambda x_i). \quad (3.7)$$

Aplicando logaritmo a (3.7), obtenemos la función de log-verosimilitud

$$l(\alpha, \lambda) = n \log(\alpha) + n \log(\lambda) + (\alpha - 1) \sum_{i=1}^n \log(1 - \exp(-\lambda x_i)) - \lambda \sum_{i=1}^n x_i. \quad (3.8)$$

Por tanto, para obtener los EMV de α y λ podemos maximizar (3.8) directamente respecto de α y λ ó podemos calcular las derivadas parciales e igualarlas a 0 como sigue:

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \log(1 - \exp(-\lambda x_i)) = 0, \quad (3.9)$$

$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} + (\alpha - 1) \sum_{i=1}^n \frac{x_i \exp(-\lambda x_i)}{1 - \exp(-\lambda x_i)} - \sum_{i=1}^n x_i. \quad (3.10)$$

A partir de (3.9), obtenemos el EMV de α en función de λ , designado $\hat{\alpha}(\lambda)$, como

$$\hat{\alpha}(\lambda) = -\frac{n}{\sum_{i=1}^n \log(1 - \exp(-\lambda x_i))}. \quad (3.11)$$

Por tanto, si el parámetro escalar es conocido, el EMV del parámetro de forma, $\hat{\alpha}$, se obtiene directamente de (3.11). Sin embargo, si ambos parámetros son desconocidos, en primer lugar podemos obtener el estimador del parámetro escalar maximizando

$$g(\lambda) = l(\hat{\alpha}(\lambda), \lambda) = C - n \log\left(-\sum_{i=1}^n \log(1 - \exp(-\lambda x_i))\right) \quad (3.12)$$

$$+ n \log(\lambda) - \sum_{i=1}^n \log(1 - \exp(-\lambda x_i)) - \lambda \sum_{i=1}^n x_i, \quad (3.13)$$

con respecto a λ . Aquí C es una constante que no depende de λ . Una vez obtenido $\hat{\lambda}$, $\hat{\alpha}$ puede ser obtenido de (3.11) como $\hat{\alpha}(\hat{\lambda})$. Por lo tanto, se reduce el problema bidimensional a un problema unidimensional que es relativamente más fácil de resolver.

En esta situación utilizamos los resultados de la normalidad asintótica para obtener el intervalo de confianza asintótico. Podemos exponer los resultados de la siguiente manera:

$$\sqrt{n} (\hat{\theta} - \theta) \rightarrow N_2(0, I^{-1}(\theta)) \quad (3.14)$$

donde $I(\theta)$ es la matriz de Información de Fisher, es decir,

$$I(\theta) = -\frac{1}{n} \begin{bmatrix} E\left(\frac{\partial^2 l}{\partial \alpha^2}\right) & E\left(\frac{\partial^2 l}{\partial \alpha \partial \lambda}\right) \\ E\left(\frac{\partial^2 l}{\partial \lambda \partial \alpha}\right) & E\left(\frac{\partial^2 l}{\partial \lambda^2}\right) \end{bmatrix},$$

con $\theta = (\alpha, \lambda)$. Ya que $\alpha > 0$, la distribución EE satisface todas las condiciones de regularidad (ver BAIN, 1976), por tanto (3.14) se cumple. Ahora, vamos a proporcionar los elementos de la matriz de Información de Fisher, que puede ser útil en la práctica. Para $\alpha > 2$

$$E\left(\frac{\partial^2 l}{\partial \alpha^2}\right) = -\frac{n}{\alpha^2},$$

$$E\left(\frac{\partial^2 l}{\partial \alpha \partial \lambda}\right) = \frac{n}{\lambda} \left[\frac{\alpha}{\alpha - 1} (\psi(\alpha) - \psi(1)) - \psi(\alpha + 1) - \psi(1) \right],$$

$$E\left(\frac{\partial^2 l}{\partial \lambda^2}\right) = -\frac{n}{\lambda^2} \left[1 + \frac{\alpha(\alpha - 1)}{\alpha - 2} (\psi'(1) - \psi'(\alpha - 1) + (\psi(\alpha - 1) - \psi(1))^2) \right],$$

$$-\frac{n\alpha}{\lambda^2} [\psi'(1) - \psi(\alpha) + (\psi(\alpha) - \psi(1))^2],$$

y para $0 < \alpha \leq 2$,

$$E\left(\frac{\partial^2 l}{\partial \alpha^2}\right) = -\frac{n}{\alpha^2},$$

$$E\left(\frac{\partial^2 l}{\partial \alpha \partial \lambda}\right) = \frac{n\alpha}{\lambda} \int_0^\infty x \exp(-2x) (1 - \exp(-x))^{\alpha-2} dx < \infty,$$

$$E\left(\frac{\partial^2 l}{\partial \lambda^2}\right) = -\frac{n}{\lambda^2} - \frac{n\alpha(\alpha - 1)}{\lambda^2} \int_0^\infty x \exp(-2x) (1 - \exp(-x))^{\alpha-2} dx < \infty.$$

Ya que θ es desconocido en (3.14), $I^{-1}(\theta)$ es estimado por $I^{-1}(\hat{\theta})$ con $\hat{\theta} = (\hat{\alpha}, \hat{\lambda})$, y este resultado se puede utilizar para obtener los intervalos de confianza asintóticos de α y λ . En presencia de censuras de Tipo 1 o Tipo 2 los resultados pueden modificarse adecuadamente. Cabe mencionar que para datos censurados de Tipo 1, la matriz de Información de Fisher puede obtenerse de la misma manera que en el caso de una muestra completa. Sin embargo, para datos censurados de Tipo 2, no es posible obtener dicha matriz de una forma cerrada.

3.4 Análisis de datos

En esta sección utilizamos dos conjuntos de datos no censurados y ajustamos los tres modelos: Gamma, Weibull y Exponencial Exponenciado. Vamos a utilizar el test chi-cuadrado de bondad de ajuste. El test chi-cuadrado de bondad de ajuste se utiliza para discriminar si una colección de datos o muestra se ajusta a una distribución teórica de una determinada población. La fórmula de cálculo del estadístico chi-cuadrado utilizado en el test chi-cuadrado de bondad de ajuste corresponde a

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - e_i)^2}{e_i} \sim \chi_{k-p-1}^2,$$

donde k es el número de clases (en nuestro caso van a ser $k=5$ intervalos), p es el número de parámetros que se estiman en el modelo (en nuestro caso $p=2$), O_i son los valores observados y e_i el valor esperado del intervalo i -ésimo

Conjunto de Datos 1: El primer conjunto de datos es el siguiente; (Lawless, 1986 página 228). Los datos que se dan aquí surgieron en las pruebas de resistencia de los rodamientos rígidos de bolas. Los datos consisten en el número de millones de revoluciones antes del fallo para cada uno de los 23 rodamientos de bolas en el test de vida y son:

17.88, 28.92, 33.00, 41.52, 42.12, 45.60, 48.80, 51.84, 51.96, 54.12, 55.56, 67.80, 68.64, 68.64, 68.88, 84.12, 93.12, 98.64, 105.12, 105.84, 127.92, 128.04 y 173.40.

Hemos ajustado Gamma, Weibull y EE a este conjunto de datos. Tenemos los estimadores de máxima verosimilitud, la log-verosimilitud (LL) (calculada a partir de los EMVs), los valores observados y esperados y los estadísticos χ^2 . Los resultados son los siguientes.

Para la distribución Gamma

$$\hat{\lambda} = 0.0556, \quad \hat{\alpha} = 4.0196, \quad LL = -113.0274, \quad \chi^2 = 1.040.$$

Para la distribución Weibull

$$\hat{\lambda} = 0.0122, \quad \hat{\alpha} = 2.1050, \quad LL = -113.6887, \quad \chi^2 = 1.791.$$

Para la distribución EE

$$\hat{\lambda} = 0.0314, \quad \hat{\alpha} = 5.2589, \quad LL = -112.9763, \quad \chi^2 = 0.783.$$

Las frecuencias observadas y esperadas son las siguientes:

Intervalos	Observado	EE	Weibull	Gamma
0-35	3	2.94	3.01	3.54
35-55	7	5.70	5.31	4.54
55-80	5	6.55	6.52	6.05
80-100	3	3.39	3.62	3.86
100-	5	4.43	4.54	5.02

Cuadro 3.2: Frecuencias observadas y esperadas Datos 1

Se observa que este conjunto de datos se ajusta mejor con una distribución EE en términos de verosimilitud y en términos de chi-cuadrado puesto que

$$p - \text{valor} = P[\chi_{k-p-1}^2 \geq \chi_{\text{obs}}^2]$$

y al ser más bajo el χ_{obs}^2 de la distribución EE, el p-valor será mayor.

Conjunto de datos 2: (Linhart and Zucchini (1986, page 69). Los siguientes datos se corresponden con los tiempos de fallo del sistema de aire acondicionado de un avión:

23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95.

De nuevo, hemos ajustado los datos con las tres distribuciones. Los resultados son los siguientes:

Para la distribución Gamma

$$\hat{\lambda} = 0.0136, \quad \hat{\alpha} = 0.8134, \quad LL = -152.2312, \quad \chi^2 = 3.302.$$

Para la distribución Weibull

$$\hat{\lambda} = 0.0183, \quad \hat{\alpha} = 0.8554, \quad LL = -152.007, \quad \chi^2 = 3.056.$$

Para la distribución EE

$$\hat{\lambda} = 0.0145, \quad \hat{\alpha} = 0.8130, \quad LL = -152.264, \quad \chi^2 = 3.383.$$

Las frecuencias observadas y esperadas son las siguientes;

Intervalos	Observado	EE	Weibull	Gamma
0-15	11	7.97	8.45	8.07
15-30	5	4.91	5.06	4.94
30-60	3	6.44	6.33	6.43
60-100	6	4.84	4.55	4.80
100-	5	5.84	5.62	5.77

Cuadro 3.3: Frecuencias observadas y esperadas Datos 2

Por el mismo motivo que antes, este conjunto de datos se ajusta mejor con una distribución Weibull. Por tanto, no está garantizado que la distribución EE se comporte mejor que una Weibull o una Gamma, pero al menos podemos decir que en ciertas circunstancias la distribución EE podría funcionar mejor que una Weibull o una Gamma.

3.5 Experimentos numéricos y discusiones

En esta sección vamos a realizar algunos experimentos numéricos para ver como los EMVs y sus resultados asintóticos funcionan para una muestra finita. Para la generación de datos vamos a usar el método de la transformación inversa. El método de la transformación inversa se utiliza para generar valores pseudoaleatorios en las distribuciones en las ue se puede hallar, de forma explícita, la inversa de la función de distribución.

| Definición 3.9. *La transformación integral de la probabilidad establece que si X es una variable aleatoria continua cualquiera con función de distribución F_X , entonces $Y = F_X(X)$ tiene una distribución uniforme en $[0,1]$. The inverse probability integral transform no es más que la inversa de lo anterior: concretamente, si Y tiene una distribución uniforme en $[0,1]$ y si X tiene una función de distribución $F_X(X)$, entonces la variable aleatoria $F_X^{-1}(Y)$ tiene la misma distribución que X .*

Descripción del método

El problema que resuelve este método es el siguiente:

- Sea X una variable aleatoria cuya distribución puede ser descrita por la función de distribución F_X .
- Queremos generar valores de X que se distribuyen según esta distribución

El método actúa como sigue:

1. Genera un número cualquiera u de la distribución uniforme estándar en el intervalo $[0,1]$, es decir, $U \sim U[0, 1]$.
2. Encuentra la inversa de la función de distribución, es decir, F_X^{-1} .
3. Calcula $x = F_X^{-1}(u)$. La variable aleatoria calculada X tiene como función de distribución F_X .

Dicho de otro modo, dada una variable aleatoria continua uniforme U en $[0,1]$ y una función de distribución invertible F_X , la variable aleatoria $X = F_X^{-1}(U)$ tiene función de distribución F_X (o X se distribuye según F_X).

En el caso de la distribución exponencial exponenciada, el método funciona como sigue:

$$\begin{aligned}
 F(x) &= (1 - \exp(-\lambda x))^\alpha, \\
 (1 - \exp(-\lambda x))^\alpha &= u \Rightarrow 1 - \exp(-\lambda x) = u^{1/\alpha} \Rightarrow \\
 \Rightarrow 1 - u^{1/\alpha} &= \exp(-\lambda x) \Rightarrow \ln(1 - u^{1/\alpha}) = -\lambda x \Rightarrow \\
 \Rightarrow x &= -\frac{\ln(1 - u^{1/\alpha})}{\lambda}
 \end{aligned}$$

Ahora si, vamos a comentar un poco que es lo que vamos a hacer. Vamos a considerar los siguientes tamaños de muestra (SS), $n = 10, 15$ (pequeño), $40, 50$ (mediano), y 100 (grande). Para cada tamaño de muestra, generamos esa cantidad de valores por el método que hemos mostrado anteriormente y calculamos los EMVs de α y λ de la siguiente forma:

1. A partir de (3.9), el EMV de α en función de λ , denotado como $\hat{\alpha}(\lambda)$, se obtiene de (3.11).
2. Sin embargo, como λ es desconocido, en primer lugar tenemos que obtener el estimador del parámetro escalar, es decir, $\hat{\lambda}$, maximizando (3.12) con respecto a λ .
3. Una vez obtenido $\hat{\lambda}$, $\hat{\alpha}$ puede ser obtenido de (3.11) como $\hat{\alpha}(\hat{\lambda})$.

Repetimos este proceso 1000 veces, es decir, generamos mil muestras del tamaño deseado mediante el método anterior y calculamos los EMVs de α y λ . Una vez hecho esto, se dan la media de las estimaciones(AE), la raíz cuadrada de los errores cuadráticos medios (SMSE) y las probabilidades de cobertura (CP). Vamos a considerar diferentes modelos para ver como se comportan estos datos.

Modelo 1: $\alpha = 2, \lambda = 0.1$, Modelo 4: $\alpha = 2, \lambda = 0.2$,

Modelo 2: $\alpha = 1, \lambda = 0.1$, Modelo 5: $\alpha = 1, \lambda = 0.2$,

Modelo 3: $\alpha = 0.5, \lambda = 0.1$, Modelo 6: $\alpha = 0.5, \lambda = 0.2$.

SS	Par	$\alpha = 2.0, \lambda = 0.1$			$\alpha = 1.0, \lambda = 0.1$			$\alpha = 0.5, \lambda = 0.1$		
		AE	SMSE	CP	AE	SMSE	CP	AE	SMSE	CP
10	α	2.720	1.703	0.97	1.261	0.667	0.97	0.599	0.277	0.97
	λ	0.112	0.033	0.98	0.113	0.037	0.98	0.113	0.037	0.98
15	α	2.470	1.181	0.95	1.171	0.470	0.96	0.565	0.120	0.96
	λ	0.109	0.029	0.97	0.109	0.031	0.97	0.110	0.034	0.97
40	α	2.160	0.579	0.96	1.061	0.238	0.96	0.524	0.103	0.96
	λ	0.104	0.019	0.95	0.104	0.021	0.96	0.105	0.025	0.96
50	α	2.130	0.501	0.96	1.051	0.210	0.96	0.520	0.092	0.96
	λ	0.103	0.016	0.95	0.104	0.019	0.96	0.105	0.023	0.96
100	α	2.053	0.303	0.95	1.024	0.130	0.96	0.509	0.058	0.96
	λ	0.101	0.011	0.95	0.102	0.013	0.95	0.102	0.016	0.95

SS	Par	$\alpha = 2.0, \lambda = 0.2$			$\alpha = 1.0, \lambda = 0.2$			$\alpha = 0.5, \lambda = 0.2$		
		AE	SMSE	CP	AE	SMSE	CP	AE	SMSE	CP
10	α	2.896	2.004	0.97	1.315	0.750	0.97	0.616	0.300	0.97
	λ	0.233	0.080	0.98	0.236	0.087	0.98	0.241	0.096	0.99
15	α	2.558	1.375	0.95	1.201	0.520	0.96	0.575	0.211	0.96
	λ	0.221	0.066	0.97	0.225	0.074	0.97	0.230	0.084	0.97
40	α	2.167	0.604	0.95	1.065	0.248	0.96	0.526	0.107	0.96
	λ	0.208	0.039	0.95	0.210	0.046	0.95	0.213	0.056	0.96
50	α	2.131	0.504	0.96	1.052	0.213	0.96	0.521	0.093	0.96
	λ	0.206	0.034	0.95	0.208	0.040	0.95	0.212	0.050	0.96
100	α	2.053	0.303	0.95	1.021	0.130	0.96	0.509	0.058	0.96
	λ	0.203	0.022	0.95	0.203	0.026	0.95	0.205	0.033	0.96

Los experimentos numéricos indican un buen comportamiento de los estimadores. Se observa que para todos los valores paramétricos considerados, los errores cuadráticos medios así como los sesgos disminuyen a medida que aumenta el tamaño de la muestra. Se verifica las propiedades de consistencia de los estimadores de máxima verosimilitud mencionadas anteriormente en este trabajo. Para un λ fijo, a medida que aumenta α , los errores cuadráticos medios y los sesgos de $\hat{\alpha}$ aumentan mientras que los correspondientes errores cuadráticos medios y los sesgos de $\hat{\lambda}$ disminuyen para todos los tamaños de muestra. Por tanto, la estimación de α es mejor a medida que α disminuye mientras que la estimación de λ es más precisa a medida que α aumenta. Por otro lado, para un α fijo, a medida que λ aumenta, los errores cuadráticos medios y los sesgos de $\hat{\alpha}$ y $\hat{\lambda}$ aumentan. Notemos que para tamaños grandes de muestra, $\hat{\lambda}/\lambda$ se mantiene constante para todo valor de α . Esto no es muy sorprendente porque λ es el parámetro escalar y también se deduce de la normalidad asintótica del estimador de máxima verosimilitud. Es interesante que para tamaños de muestra moderados o grandes, se observa que para un α fijo, el EMV de α y su correspondiente error cuadrático medio permanece constante para diferentes valores de λ . Es claro que los EMVs de α y de λ están positivamente sesgados aunque los sesgos tienden a 0 a medida que aumenta el tamaño de la muestra. Es también interesante observar que el intervalo de confianza asintótico mantiene las probabilidades de cobertura nominal incluso para tamaños de muestra pequeños. Por tanto, los EMVs y sus resultados asintóticos pueden ser usados para la estimación y para la construcción de intervalos de confianza

incluso para tamaños de muestra pequeños.

3.6 Conclusiones

En este capítulo se ha estudiado la familia de distribuciones EE. Se observa que la distribución EE de dos parámetros es bastante similar en su naturaleza a otras distribuciones de dos parámetros utilizadas en fiabilidad como la distribución Weibull y la distribución gamma. Hemos visto que la mayoría de las propiedades de la distribución EE son bastante similares a las de una distribución gamma pero computacionalmente se comporta de manera similar a la distribución Weibull. Por tanto, esta distribución puede utilizarse como alternativa a la distribución Weibull o a la distribución gamma y se espera que en algunas situaciones pueda funcionar mejor (en términos de ajuste) que estas distribuciones aunque no se puede garantizar.

Hemos presentado dos conjuntos de datos reales en los que hemos observado que en el primero la distribución EE tiene mejor ajuste que la Weibull o gamma, pero que en el segundo Weibull tiene mejor ajustes que las otras dos distribuciones. Además, es bien sabido que la distribución gamma tiene ciertas ventajas en comparación con la distribución Weibull en cuanto a la convergencia más rápida de los EMVs. Se espera que la distribución EE también goce de estas propiedades. Los experimentos numéricos confirman que para la distribución EE los resultados asintóticos pueden utilizarse incluso para tamaños de muestra pequeños para diferentes α y λ .

Bibliografía

- [Arnold, B. C. (1983)] Pareto distributions, International Co-operative Publishing House, Fairland, MD.
- [Azzalini, A. (1985)] A Class of distributions which includes the normal ones, Scandinavian Journal of Statistics, 12, 171–178.
- [Bain, L. J. (1976)] Statistical analysis of reliability and life testing model. Marceland Dekker Inc., New York.
- [Balakrishnan, N. and Leung, M. Y. (1988a)] Order statistics from the type I generalized logistic distribution, Communications in Statistics, Simulation and Computation, 17, 25–50.
- [Barlow, R. E. and Prochan, F. (1975)] Statistical theory of reliability and life testing: probability models, Holt, Reinhardt and Winston, New York.
- [Castillo, E. and Hadi, A. S. (1995a)] A method for estimating parameters and quantiles of distributions of continuous random variables, Computational Statistics and Data Analysis, 20, 421–439.
- [Castillo, E. and Hadi, A. S. (1995b)] Modeling lifetime data with applications to fatigue models, Journal of the American Statistical Association, 90, 1041–1054.
- [Castillo, E. and Hadi, A. S. (1997)] Fitting the generalized Pareto distribution to data, Journal of the American Statistical Association, 92, 1609–1620.
- [Castillo, E., Hadi, A. S., and Sarabia, J. M. (1998)] A method for estimating Lorenz curves, Communications in Statistics, Theory and Methods, 27, 2037–2063.
- [Dagum, C. (1977)] A new model of personal income distribution: specification and estimation, Economie Appliquee ´ , 30, 413–437.

- [Durrans, S. R. (1992)] Distributions of fractional order statistics in hydrology, *Water Resources Research*, 28, 1649–1655.
- [Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J.R. (1979)] Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form, *Water Resources Research*, 15, 1049–1054.
- [Greenwood, J. A., Landwehr, J. M., Matalas, N. C., and Wallis, J.R. (1979)] Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form, *Water Resources Research*, 15, 1049–1054.
- [Gupta, R. C., Gupta, P. L., and Gupta, R. D. (1998)] Modeling failure time data by Lehmann alternatives, *Communications in Statistics, Theory and Methods*, 27, 887–904.
- [Gupta, R. D. and Kundu, D. (1999)] Generalized exponential distributions, *Australian and New Zealand Journal of Statistics*, 41, 173–188.
- [Gupta, R. D. and Kundu, D. (2001)] Exponentiated exponential family: and alternative to gamma and Weibull distributions, *Biometrical Journal*, 43, 117–130.
- [Gupta, R. D. and Kundu, D. (2003)] Discriminating between Weibull and generalized exponential distributions, *Computational Statistics and Data Analysis*, 43, 179–196.
- [Jackson, O. A. Y. (1963)] Fitting a gamma or log normal distribution to fiber diameter measurements of wooltops. *Applied Statistics* 1, 161–166.
- [Jones, M. C. (2004)] Families of distributions arising from distributions of order statistics (with discussion), *Test*, 13, 1–43.
- [Kotz, S., Balakrishnan, N., and Johnson, N.L. (2000)] Continuous multivariate distributions, vol. 1: Models and applications, John Wiley and Sons, New York.
- [McDonald, J. B. (1984)] Some generalized functions for the size distribution of income, *Econometrica*, 52, 647–663.
- [Moors, J. J. A. (1998)] A quantile alternative for kurtosis, *The Statistician*, 37, 25–32.

- [Mudholkar, G. S. and Hutson, A. D. (1996)] The exponentiated Weibull family: some properties and a flood data application, *Communications in Statistics, Theory and Methods*, 25, 3059– 3083. *Communications in Statistics, Theory and Methods*, 34, 253–256.
- [Nassar, M. M. and Eissa, F. H. (2003)] On the exponentiated Weibull distribution, *Communications in Statistics, Theory and Methods*, 32, 1317–1336.
- [Raqab, M. Z. and Ahsanullah, M. (2001)] Estimation of the location and scale parameters of generalized exponential distribution based on order statistics, *Journal of Statistical Computation and Simulation*, 69, 109–124.
- [Sarabia, José & Castillo, Enrique. (2005)] About a class of max-stable families with applications to income distributions. *Metron - International Journal of Statistics*. LXIII. 505-527.
- [Stoppa, G. (1990)] A new model for income size distributions, In: *Income and wealth distribution*,
- [Team, R. C. (2013)] *R: A language and environment for statistical computing*.
- [Weibull, W. (1939)] *A statistical theory of the strength of material*. Ingeniors Vetenskaps Akademiens, Stockholm 151.

