

Estudio de eventos extremos enfocado a seguros y finanzas

KLEVER MEJÍA – ADRIANA UQUILLAS^{*}

Resumen

Muchos campos de la ciencia moderna y la ingeniería tienen que lidiar con eventos que son poco frecuentes pero que traen consecuencias muy significativas. La teoría de valores extremos ofrece las bases para la modelización estadística de tales extremos. El potencial de la teoría de valores extremos enfocada a problemas de finanzas ha sido reconocido recientemente. Este artículo busca introducir los fundamentos teóricos de la teoría de valores extremos además de aspectos prácticos de estimación. Principalmente presentamos dos estudios. El uno está basado en la distribución asintótica conjunta de las estadísticas de orden extremas. El otro método alternativo se refiere a modelar todos los excesos sobre un umbral usando la distribución generalizada de Pareto.

Abstract

Many fields of modern science and engineering have to deal with events which are rare but have significant consequences. Extreme value theory is considered to provide the basis for the statistical modeling of such extremes. The potential of extreme value theory applied to financial problems has only been recognized recently. This paper aims at introducing the fundamentals of extreme value theory as well as practical aspects for estimating. Mainly there have been two new approaches: One approach is based on the asymptotic joint distribution of extreme order statistics; and an alternative approach is model all the exceedances over a high threshold using a Generalized Pareto Distribution.

1. Introducción

La Teoría de Valores Extremos no es un tema nuevo pues se empezó a estudiarlo

^{*} El presente trabajo es un resumen de la tesis presentada a la Escuela Politécnica Nacional, previa a la obtención del título de ingenieros matemáticos.

desde el siglo XIX, no obstante es en la actualidad en que ha tomado gran fuerza especialmente en lo que se refiere a sus aplicaciones en finanzas y seguros. Esta teoría es la que mejor uso hace de cualquiera que sean los datos que se tengan sobre fenómenos extremos y, dado que ayuda a aclarar inquietudes respecto a lo que es la proyección de ocurrencia de eventos poco frecuentes su uso es excepcional en la estimación de riesgos financieros.

En el presente documento trataremos de explicar cómo y cuándo ocurren los eventos extremos, se expondrán métodos matemáticos apropiados para explicar aquellos eventos que ocurren con probabilidad relativamente pequeña, pero que tienen una influencia significativamente grande en el comportamiento de todo el modelo. Los métodos de extremos son de simple aplicación y funcionan bien en estudios de simulación. En ocasiones esta explicación teórica no se extiende a situaciones mucho más complicadas tales como modelos de regresión basados en distribuciones de valores extremos.

Una de las recientes aplicaciones en el campo de las finanzas es el estudio de portafolios. Considerando por ejemplo un portafolio financiero que contiene un número determinado de riesgos distintos en: bonos de gobiernos, acciones, títulos de crédito privados, etc., podría estimarse las covarianzas del portafolio para obtener la distribución total de la pérdidas o ganancias de las inversiones. Empero, los analistas de riesgos, los supervisores y reguladores pueden además estar interesados en la fijación de “requerimientos mínimos” de utilidad, o alternativamente en determinar el límite máximo para las pérdidas potenciales. En este último caso, podría obtenerse el valor en riesgo (VaR) del portafolio aplicando la técnica de valores extremos, dado un nivel de confianza. Así mismo, podría obtenerse una distribución de déficit que estima la probabilidad de exceder en un valor específico un umbral determinado, el primer resultado tiene que ver con la estimación cuantil para una función de distribución estimada. El segundo resultado tiene que ver con la estimación de la función de distribución de los excesos. La teoría recomienda para este caso usar la Distribución Generalizada de Pareto como un modelo paramétrico natural.

Se debe advertir que, existe una gran variedad de respuestas que se pueden obtener a través de esta teoría; una vez conocidos la distribución de los valores máximos (mínimos) de los datos subyacentes, se puede calcular la probabilidad de que un valor exceda el máximo(mínimo) en un cierto período de tiempo y por otro lado se podría estimar cuánto tiempo habría que esperar antes de que ocurra una cantidad de “pérdida” alta ya especificada, además saber cuál es la probabilidad para una variable, dada la máxima “pérdida” para que el siguiente período de tiempo exceda un

cierto nivel, también se podrá saber la probabilidad de que la máxima “pérdida” exceda todas las pérdidas anteriores. Si las variables en estudio describen los ingresos se puede calcular el volumen necesario de ingresos para el siguiente año, con el objeto de que cubran los futuros egresos con probabilidad lo suficientemente alta.

Existe una gran cantidad de modelos probabilísticos para describir por vía matemática eventos extremos en el caso unidimensional. El mundo verdadero sin embargo nos informa a menudo sobre tales acontecimientos con datos estadísticos: disminuciones grandes de los valores del mercado sobre cierto período del tiempo, o valores máximos y mínimos del rendimiento de un portafolio. Estos ejemplos definitivamente se refieren a preguntas sobre valores extremos de un cierto sistema subyacente de datos, de manera que, sería absurdo tratar de aplicar eficientemente en estos temas la estadística clásica..

2. Principales Tópicos de la Teoría de Valores Extremos

2.1 Fluctuaciones del Máximo

2.1.1 Introducción

Esta teoría es fundamental para muchos resultados aplicados en valores extremos. El resultado principal es el teorema de Fisher – Tippett el cual especifica la forma de la distribución límite para máximos centrados y normalizados.

Introducimos en este capítulo la noción de función de exceso media. La misma que se probará que es una herramienta útil para distinguir las f.d de sus colas derechas y juega un papel importante en la estimación de colas. La distribución Generalizada de Valores Extremos también conduce a la distribución Generalizada de Pareto.

2.1.2 Probabilidades límite para máximos

En el transcurso de esta sección X, X_1, X_2, \dots es una sucesión de v.a iid. no degeneradas con función de distribución común F . En esta sección investigamos la fluctuación de los máximos de una muestra. Notamos:

$$M_1 = X_1, \quad M_n = \max(X_1, \dots, X_n), \quad n \geq 2$$

Los resultados correspondientes a los mínimos fácilmente se los puede obtener de los máximos, se los obtiene de la siguiente manera

$$\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n), \quad n \geq 2.$$

Nota: Posteriormente analizaremos estadísticas de orden superior de la muestra X_1, \dots, X_n . ♣

La función de distribución exacta del máximo M_n la podemos escribir como:

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = F^n(x), \quad x \in \tilde{n} \text{ y } n \in \mathbb{N}$$

Los extremos se dan cerca del fin superior de la base de la distribución, aquí intuitivamente el comportamiento asintótico de M_n deberá ser descrito de la f.d.F en su cola derecha cerca del punto final derecho. Denotamos por

$$x_F = \sup\{x \in \tilde{n}: F(x) < 1\}$$

el punto final derecho de F . De donde, inmediatamente obtenemos, para todo $x < x_F$,

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

Y en el caso en que $x_F < \infty$. Tenemos para $x \geq x_F$ que

$$P(M_n \leq x) = F^n(x) = 1$$

Así $M_n \xrightarrow{p} x_F$ cuando $n \rightarrow \infty$, donde $x_F \leq \infty$. Ya que la sucesión (M_n) es no decreciente en n , esto converge c.s. y de aquí concluimos que

$$M_n \xrightarrow{c.s.} x_F \quad n \rightarrow \infty \quad (2.1)$$

Este hecho no nos da mucha información. Sin embargo, una mayor comprensión sobre la magnitud del máximo, está dada por el resultado de la convergencia débil para máximos centrados y normalizados. Este es uno de los tópicos importantes de la teoría de valores extremos clásica. Así, por ejemplo, el teorema fundamental de Fisher Tippett tiene el siguiente contenido:

Si existen constantes $c_n > 0$ y $d_n \in \tilde{n}$ tal que

$$c_n^{-1} (M_n - d_n) \xrightarrow{d} H, \quad n \rightarrow \infty \quad (2.2)$$

para alguna función de distribución H no degenerada, entonces H debe ser del tipo de una de las tres distribuciones llamadas de valores extremos estándar. Esto es similar al Teorema Central del Límite. Donde la distribución estable es solo posible para leyes de límites no degeneradas. Consecuentemente se debe considerar las probabilidades de la forma

$$P(c_n^{-1}(M_n - d_n) \leq x),$$

las cuales pueden ser escritas como

$$P(M_n \leq u_n), \tag{2.3}$$

donde $u_n = u_n(x) = u_n x + d_n$.

Para continuar con esta teoría comenzamos con un resultado elemental el cual es crucial para entender la teoría del límite débil del máximo de una muestra.

Proposición 2.1 Aproximación de Poisson

Dado $\tau \in [0, \infty]$ y la sucesión (u_n) de números reales, las siguientes relaciones son equivalentes

$$n \bar{F}(u_n) \rightarrow \tau \tag{2.4}$$

$$P(M_n \leq u_n) \rightarrow e^{-\tau} \tag{2.5}$$

Demostración

Consideramos en primer lugar que $0 \leq \tau \leq \infty$. Si (2.4) se cumple, entonces

$$P(M_n \leq u_n) = F^n(u_n) = (1 - \bar{F}(u_n))^n = \left(1 - \frac{\tau}{n} + o\left(\frac{1}{n}\right)\right)^n,$$

lo que implica (2.5). Inversamente, si (2.5) se cumple, entonces $\bar{F}(u_n) \rightarrow 0$. Consecuentemente,

$$P(M_{nk} \leq u_{nk}) = (1 - \bar{F}(u_{nk}))^{nk}$$

implicaría que $P(M_{nk} \leq u_{nk}) \rightarrow 0$.

Tomando logaritmos en (2.5) tenemos que

$$-n \ln(1 - \bar{F}(u_n)) \rightarrow t$$

Como $\lim_{x \rightarrow 0} \frac{-\ln(1-x)}{x} = 1$, esto implica que

$$n\bar{F}(u_n) = t + o(1), \text{ obteniendo entonces} \quad (2.4)$$

Si $\tau = \infty$ y (2.4) se cumple, pero no se cumple (2.5), puede haber una subsucesión (n_k) tal que

$$P(M_{n_k} \leq u_{n_k}) \rightarrow \exp\{-t'\} \text{ cuando } k \rightarrow \infty$$

para algún $t' < \infty$.

Pero entonces (2.5) implica (2.4), y $n_k \bar{F}(u_{n_k}) \rightarrow t' < \infty$ contradiciendo (2.4) con $t = \infty$. Similarmente, (2.5) implica (2.4) para $t = \infty$. ♣

Por (2.1), (M_n) converge casi seguramente al punto final derecho x_F de la f.d F, de aquí

$$P(Mn \leq x) \rightarrow \begin{cases} 0 & \text{si } x < x_F \\ 1 & \text{si } x > x_F. \end{cases}$$

El siguiente resultado extiende esta clase de comportamiento 0 – 1.

Corolario 2.2 Supongamos que $x_F < \infty$ y

$$\bar{F}(x_F^-) = F(x_F) - F(x_F^-) > 0.$$

Entonces para toda sucesión (u_n) se tiene que

$$P(M_n \leq u_n) \rightarrow r$$

con $r = 0$ o $r = 1$.

Demostración

Como $0 \leq r \leq 1$, podemos escribir $r = \exp\{-\tau\}$ con $0 \leq \tau \leq \infty$. Por la aproximación de Poisson tenemos $n\bar{F}(u_n) \rightarrow t$ cuando $n \rightarrow \infty$. Si $u_n < x_F$, para n grande tenemos $\bar{F}(u_n) \geq \bar{F}(x_{F^-}) > 0$ y de aquí $\tau = \infty$. La otra posibilidad es que $u_n \geq x_F$ para todo n suficientemente grande, con esa suposición $n\bar{F}(u_n) = 0$. Así $\tau = \infty$ o 0 , dando $\rho = 0$ o 1 . ♣

En la siguiente caracterización tenemos un resultado similar para ciertas distribuciones con punto final derecho infinito.

Teorema 2.3 Sea F una f.d con punto final derecho $x_F \leq \infty$ y sea $\tau \in [0, \infty]$. Existe una sucesión (u_n) que satisface $n\bar{F}(u_n) \rightarrow \tau$ si y solo si

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{F}(x^-)} = 1. \quad (2.6) \quad \clubsuit$$

En este caso creemos que es preferible exponer este ejemplo en vez de su demostración.

Ejemplo 2.4 Distribución de Poisson

$$P(X = x) = e^{-\lambda} \lambda^k / k!, \quad k \in \mathbb{N}_0, \lambda > 0.$$

Entonces

$$\frac{\bar{F}(k)}{\bar{F}(k-1)} = 1 - \frac{F(k) - F(k-1)}{\bar{F}(k-1)}$$

$$\begin{aligned}
 &= 1 - \frac{I^k}{k!} \left(\sum_{r=k}^{\infty} \frac{I^r}{r!} \right)^{-1} \\
 &= 1 - \left(1 + \sum_{r=k+1}^{\infty} \frac{I^{r-k}}{r!} \right)^{-1}
 \end{aligned}$$

la última suma puede ser estimada por

$$\sum_{s=1}^{\infty} \frac{I^s}{(k+1)(k+2)\dots(k+s)} \leq \sum_{s=1}^{\infty} \left(\frac{I}{k} \right)^s = \frac{I/k}{1-I/k}, \quad k > I,$$

la cual tiende a 0 cuando $k \rightarrow \infty$, así que $\bar{F}(k)/\bar{F}(k-1) \rightarrow 0$. El teorema 2.3 muestra que la distribución de límite no degenerada para el máximo no existe y, además, que el límite de la forma $P(M_n \leq u_n) \rightarrow r \in (0, 1)$ no existe, para la sucesión de constantes (u_n) .♣

Este resultado se aplica en particular a distribuciones discretas con punto final derecho infinito. Si las alturas del salto de la f.d no decaen lo suficientemente rápido, entonces la distribución límite no degenerada para el máximo no existe. Por ejemplo, si X es un valor entero y $x_F = \infty$, entonces (2.6) se transforma en $\bar{F}(n)/\bar{F}(n-1) \rightarrow 1$ cuando $n \rightarrow \infty$.

Como ya se dijo antes, el siguiente resultado es básico en la teoría de valores extremos clásica.

Teorema 2.5 Teorema de Fisher – Tippett, ley límite para el máximo

Sea (X_n) una sucesión de v.a iid. Si existe constantes normalizadas $c_n > 0$, $d_n \in \bar{\mathbf{n}}$ y alguna f.d. H no degenerada tal que

$$c_n^{-1} (M_n - d_n) \xrightarrow{d} H, \quad (2.7)$$

entonces H pertenece al tipo de una de las tres siguientes distribuciones:

$$\begin{aligned}
 \text{Fréchet:} \quad \Phi_a(x) &= \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-a}\}, & x > 0 \end{cases} \quad \text{y} \quad \mathbf{a} > 0. \\
 \text{Weibull:} \quad \Psi_a(x) &= \begin{cases} \exp\{-(-x)^a\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \text{y} \quad \mathbf{a} > 0. \\
 \text{Gumbel:} \quad \Lambda(x) &= \exp\{-e^{-x}\}, \quad x \in \tilde{\mathbf{n}}.
 \end{aligned}$$

La demostración de este teorema se la puede encontrar en Resnick S.I. (1987) Extreme Values, Regular Variation, en la proposición 0.3♣.

Sabemos además que para $t > 0$ y por la convergencia débil existe una función $\gamma(t) > 0$, $\delta(t)$ real que satisface

$$\lim_{n \rightarrow \infty} \frac{c_n}{C_{[nt]}} = \mathbf{g}(t), \quad \lim_{n \rightarrow \infty} \frac{d_n - d_{[nt]}}{C_{[nt]}} = \mathbf{d}(t), \quad t > 0,$$

y

$$H^t(x) = H(\gamma(t)x + \mathbf{d}(t)). \quad (2.8)$$

Se puede decir de (2.8) que para $s, t > 0$

$$\gamma(st) = \gamma(s)\gamma(t), \quad \delta(st) = \gamma(s)\delta(s) + \delta(t) \quad (2.9)$$

Definición 2.6 Distribución de valores extremos y variables aleatorias extremas

Las funciones de distribución Φ_α , Ψ_α y Λ como las presentadas en el teorema 2.5 son llamadas distribuciones de valores extremos estándar, las variables aleatorias correspondientes, variables aleatorias extremas estándar. Las funciones de distribuciones de tipo Φ_α , Ψ_α y Λ son distribuciones de valores extremos; las variables aleatorias correspondientes, son variables aleatorias extremas. ♣

Las distribuciones de valores extremos son precisamente las distribuciones max-estables. En particular los tres casos del teorema anterior corresponden a

$$\begin{aligned} \text{Frechét:} \quad M_n &= n^{1/\alpha} X \\ \text{Weibull:} \quad M_n &= n^{-1/\alpha} X \\ \text{Gumbel:} \quad M_n &= X + \ln n. \end{aligned}$$

2.1.3 Máximo dominio de atracción y normalización de constantes

En esta sección tratamos de responder a las siguientes preguntas:

- Cómo escoger las constantes normalizadas $c_n > 0$ y $d_n \in \tilde{n}$ tal que

$$c_n^{-1} (M_n - d_n) \xrightarrow{d} H ? \quad (2.10)$$

- Dada una distribución de valor extremo H , qué condiciones en la f.d. F implican que el máximo normalizado M_n converja débilmente a H ?

Definición 2.7 Máximo dominio de atracción

Decimos que la v.a X (la distribución de X) pertenece al Máximo Dominio de Atracción de la distribución de valores extremos H si existen constantes $c_n > 0$ y $d_n \in \tilde{n}$ tal que (2.10) se cumple. Escribimos $X \in \text{MDA}(H)$ ($F \in \text{MDA}(H)$). ♣

El siguiente resultado es una consecuencia de la proposición 2.1 y lo usaremos en las siguientes secciones.

Proposición 2.8 Caracterización del MDA(H)

La f.d. F pertenece al Máximo Dominio de Atracción de la distribución de valores extremos H con constantes normalizadas $c_n > 0$ y $d_n \in \tilde{n}$ si y solo si

$$\lim_{n \rightarrow \infty} n \overline{F}(c_n x + d_n) = \lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H(x), \quad x \in \tilde{n}.$$

Cuando $H(x) = 0$ el límite es interpretado como ∞ . ♣

Para todas las distribuciones de valores extremos se puede caracterizar su

Máximo Dominio de Atracción usando el concepto de variaciones regulares. Recordemos que la distribución de colas \bar{F} es una variación regular con índice $-\alpha$ para algún $\alpha \geq 0$. Escribimos $\bar{F} \in \hat{\mathcal{A}}_{-\alpha}$ si

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(xt)}{\bar{F}(x)} = t^{-\alpha}, \quad t > 0$$

El siguiente concepto define una relación de equivalencia en el conjunto de todas las f.d.

Definición 2.9 Inversa generalizada de una función monótona

Supongamos que h es una función no decreciente en los reales. A la inversa generalizada de h se la define como

$$h^{\leftarrow}(t) = \inf\{x \in \mathbb{N} : h(x) \geq t\}. \clubsuit$$

Definición 2.10 Función Cuantil

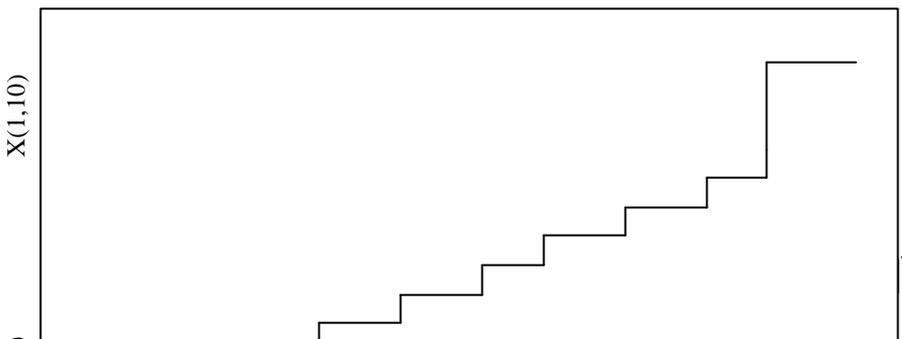
La inversa generalizada de la f.d. F

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

es llamada función cuantil de la f.d. F . La cantidad $x_t = F^{\leftarrow}(t)$ define el t -cuantil de F . \clubsuit

Gráfico No. 2.1

**Función Cuantil empírica F_n^{\leftarrow} de un conjunto de 10 v.a
exponenciales estándar**



2.1.3.1 Máximo dominio de atracción de la distribución de Fréchet

En esta sección caracterizamos del máximo dominio de atracción de Φ_a para $\alpha > 0$ que es la distribución más usada en las finanzas, especialmente para fluctuaciones grandes de precios. Por extensión de Taylor

$$1 - \Phi_a(x) = 1 - \exp\{-x^{-\alpha}\} \sim x^{-\alpha}, \quad x \rightarrow \infty,$$

De aquí que la cola de Φ_a decrece como una ley fuerte.

Para $F \in \text{MDA}(\Phi_a)$ podemos elegir c_n por medio de la función cuantil, más precisamente por

$$\begin{aligned} c_n &= F^{\leftarrow}(1 - n^{-1}) = \inf \{x \in R : F(x) \geq 1 - n^{-1}\} \\ &= \inf \{x \in R : (1/\bar{F})(x) \geq n\} \\ &= (1/\bar{F})^{\leftarrow}(n) \end{aligned} \quad (2.11)$$

y d_n puede ser elegida como 0 aunque no necesariamente. Cabe notar que las f.d. de F conforman el $\text{MDA}(\Phi_a)$.

Teorema 2.11 Máximo dominio de atracción de Φ_a

La f.d. F pertenece al máximo dominio de atracción de Φ_a , $\alpha > 0$, si y solo si $\bar{F}(x) = x^{-\alpha} L(x)$ para alguna función suavemente variante L .

Si $F \in \text{MDA}(\Phi_a)$, entonces

$$c_n^{-1} M_n \xrightarrow{d} \Phi_a,$$

donde las constante normalizadas c_n pueden ser elegidas como en (2.11).

Este resultado muestra que el punto final derecho $\bar{x} = \infty$ para toda $F \in \text{MDA}(\Phi_a)$, y $c_n = n^{1/\alpha} L_1(n)$ para una función suavemente variante L_1 .

Demostración

Sea $\bar{F} \in \mathfrak{R}_{-a}$ para $\alpha > 0$. Con una adecuada elección de c_n y variación regular,

$$\bar{F}(c_n) \sim n^{-1}, \quad n \rightarrow \infty,$$

de aquí $\bar{F}(c_n) \rightarrow 0$ cuando $c_n \rightarrow \infty$. Para $x > 0$,

$$n\bar{F}(c_n x) \sim \frac{\bar{F}(c_n x)}{\bar{F}(c_n)} \rightarrow x^{-a}, \quad n \rightarrow \infty$$

Para $x < 0$, se ve claramente que $\bar{F}(c_n x) \leq F^n(0) \rightarrow 0$, ya que la variación regular requiere $F(0) < 1$. Por la proposición 2.8, $F \in \text{MDA}(\Phi_a)$.

Inversamente, suponemos que $\lim_{n \leftarrow \infty} F^n(c_n x + d_n) = \Phi_a(x)$ para todo $x > 0$ y constantes apropiadas $c_n > 0, d_n \in \tilde{\mathfrak{n}}$. Esto nos lleva a

$$\lim_{n \rightarrow \infty} F^n(c_{[ns]}x + d_{[ns]}) = \Phi_a^{1/s}(x) = \Phi_a(s^{1/a}x), \quad s > 0, \quad x > 0.$$

Por el teorema de convergencia a tipos (anexo A.7)

$$\frac{c_{[ns]}}{c_n} \rightarrow s^{1/a} \quad \text{y} \quad (d_{[ns]} - d_n)/c_n \rightarrow 0.$$

(c_n) es una sucesión que varía regularmente con índice α en particular $c_n \rightarrow \infty$.

Suponemos primero que $d_n = 0$, entonces $n\bar{F}(c_n x) \rightarrow x^{-a}$ así que $\bar{F} \in \mathfrak{R}_{-a}$ debido a la proposición del anexo A.4. El caso de $d_n \neq 0$ es más complicado, en efecto se tiene que mostrar que $d_n/c_n \rightarrow 0$. Si esto último se cumple se puede repetir el argumento de arriba reemplazando d_n por 0. ♣

Por otra parte $F \in MDA(\Phi_a) \Leftrightarrow \bar{F} \in R_{-a}$ indica hasta dónde puede trasladarse sin perder la condición de pertenencia a $MDA(\Phi_a)$. Estas clases de distribuciones contienen “distribuciones de colas muy pesadas”.

Corolario 2.12 Condición de Von Mises

Sea F una f.d. absolutamente continua con densidad f que satisface

$$\lim_{x \rightarrow \infty} \frac{xf'(x)}{\bar{F}(x)} = a > 0, \quad (2.12_f)$$

entonces $F \in MDA(\Phi_a)$. ♣

Proposición 2.13 Propiedad de encierro del $MDA(\Phi_a)$

Sean F y G dos f.d. y asumamos que $F \in MDA(\Phi_a)$ con constantes normalizadas $c_n > 0$, es decir

$$\lim_{n \rightarrow \infty} F^n(c_n x) = \Phi_a(x), \quad x > 0.$$

Entonces

$$\lim_{n \rightarrow \infty} G^n(c_n x) = \Phi_a(cx), \quad x > 0,$$

para algún $c > 0$ si y solo si F y G son colas equivalentes con

$$\lim_{x \rightarrow \infty} \bar{F}(x) / \bar{G}(x) = c^a.$$

Demostración

Condición suficiente

Supongamos que $\bar{F}(x) = q\bar{G}(x)$ cuando $x \rightarrow \infty$ para algún $q > 0$. Por la proposición 2.8 la relación límite $\lim_{n \rightarrow \infty} F^n(c_n x) = \Phi_a(x)$, $x > 0$ es equivalente a

$$\lim_{n \rightarrow \infty} n\bar{F}(c_n x) = x^{-a}$$

para todo $x > 0$. Para ese x , $c_n x \rightarrow \infty$ cuando $n \rightarrow \infty$ y aquí, por colas equivalentes,

$$n\bar{G}(c_n x) \sim nq^{-1}\bar{F}(c_n x) \rightarrow q^{-1}x^{-a},$$

e.d. nuevamente por la proposición 2.8

$$\lim_{n \rightarrow \infty} G^n(c_n x) = \exp\left\{-\left(q^{1/a}x\right)^{-a}\right\} = \Phi_a\left(q^{1/a}x\right),$$

ahora fijamos $c = q^{1/a}$ y así tendremos

$$\lim_{n \rightarrow \infty} G^n(c_n x) = \exp\left\{-\left(q^{1/a}x\right)^{-a}\right\} = \Phi_a(cx)$$

El otro sentido de la demostración se la puede encontrar en Resnick, S.I. (1986) Valores Extremos, Variación Regular y Procesos Puntuales, proposición 2.19. ♣

De los teoremas anteriores podemos resumir lo siguiente:

El $\text{MDA}(\Phi_a)$ consiste de f.d. que satisface la condición de Von Mises y sus f.d. de colas equivalentes.

2.1.3.2 Máximo dominio de atracción de la distribución de Weibull

Aquí caracterizaremos el máximo dominio de atracción de Ψ_a para $\alpha > 0$.

Un hecho importante, aunque de ninguna manera obvio es que todas las funciones de distribución F en $\text{MDA}(\Psi_a)$ tiene punto final derecho x_F finito. Y como ya habíamos indicado que Ψ_a y Φ_a están estrechamente relacionadas, en efecto

$$\Psi_a(-x^{-1}) = \Phi_a(x), \quad x > 0.$$

De aquí podemos esperar que todo $\text{MDA}(\Psi_a)$ y $\text{MDA}(\Phi_a)$ estén estrechamente relacionadas. Se confirma esto en el siguiente teorema.

Teorema 2.14 Máximo dominio de atracción de Ψ_a

La f.d. F pertenece al máximo dominio de atracción de Ψ_a , $\alpha > 0$, si y solo si $x_F < \infty$ y $\bar{F}(x_F - x^{-1}) = x^{-\alpha}L(x)$ para alguna función suavemente variante L .

Si $F \in \text{MDA}(\Psi_a)$, entonces

$$c_n^{-1}(M_n - x_F) \xrightarrow{d} \Psi_a,$$

donde la constante normalizada c_n puede ser elegida como $c_n = x_F - F^{\leftarrow}(1 - n^{-1})$ y $d_n = x_F$.

Su demostración se puede encontrar en Resnick S.I. (1987) Extreme Values, Regular Variation ♣

Debido a que las $F \in \text{MDA}(\Psi_a)$ presentan el punto final derecho finito en algunas ocasiones no es una buena opción aplicar a modelos de finanzas o modelos

de seguros, ya que estos modelos no deben tener limitaciones, aunque en la práctica todo tiene límite.

Anteriormente hablamos de la condición de Von Mises para Φ_a , ahora la extendemos a Ψ_a .

Corolario 2.15 Condición de Von Mises

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

$$\lim_{x \uparrow x_F} \frac{(x_F - x)f(x)}{\overline{F}(x)} = a > 0, \quad (2.12_w)$$

entonces $F \in \text{MDA}(\Psi_a)$.♣

Aplicando la transformación $F^*(x) = F(x_F - x^{-1})$, $x > 0$, la proposición 2.13 puede ser reformulada como sigue.

Proposición 2.16 Propiedad de encierro del $\text{MDA}(\Psi_a)$

Sean F y G dos f.d. con punto final derecho $x_F = x_G < \infty$ y asumimos que $F \in \text{MDA}(\Psi_a)$ con constante normalizada c_n e.d.

$$\lim_{n \rightarrow \infty} F^n(c_n x + x_F) = \Psi_a(x), \quad x < 0.$$

entonces

$$\lim_{n \rightarrow \infty} G^n(c_n x + x_F) = \Psi_a(cx), \quad x < 0,$$

para algún $c > 0$ si y solo si F y G son colas equivalente con

$$\lim_{x \uparrow x_F} \bar{F}(x) / \bar{G}(x) = c^{-a} \cdot \clubsuit$$

Notemos que el teorema de representación para funciones suavemente variantes implica que toda $F \in \text{MDA}(\Psi_a)$ es cola equivalente de una f.d. absolutamente continua que satisface (2.12_w).

2.1.4 La distribución generalizada de valores extremos y la distribución generalizada de Pareto

Anteriormente habíamos presentado que la distribución de valores extremos estándar, provee solo las leyes límites no degeneradas para máximos transformados de v.a iid. La representación uni - paramétrica de los tres casos estándar en una familia de distribuciones resultará ser útil. Ellos pueden ser representados introduciendo un parámetro ξ tal que

$\xi = \alpha^{-1} > 0$ corresponde a la distribución de Fréchet Φ_a

$\xi = 0$ corresponde a la distribución de Gumbel Λ

$\xi = -\alpha^{-1} < 0$ corresponde a la distribución de Weibull Ψ_a

Definición 2.17 Representación de Jenkinson – Von Mises de las distribuciones de valores extremos: La distribución generaliza de valores extremos (GVE).

Definimos la f.d. H_x por

$$H_x(x) = \begin{cases} \exp\{- (1 + \alpha x)^{-1/\alpha}\} & \text{si } \alpha \neq 0, \\ \exp\{- \exp\{-x\}\} & \text{si } \alpha = 0, \end{cases}$$

donde $1 + \alpha x > 0$.

Aquí la base de H_x corresponde a

$$X > -\alpha^{-1} \quad \text{para } \alpha > 0,$$

$$\begin{aligned} X &< -\mathbf{x}^{-1} && \text{para } \xi < 0, \\ X &\in \tilde{\mathbf{n}} && \text{para } \xi = 0. \end{aligned}$$

$H_{\mathbf{x}}$ es llamada la distribución generalizada de valores extremos estándar (GEV).♣

Consideramos la distribución H_0 como el límite de $H_{\mathbf{x}}$ cuando $\xi \rightarrow 0$. Con esta interpretación

$$H_{\mathbf{x}}(x) = \exp\left\{- (1 + \mathbf{x}x)^{-1/\mathbf{x}}\right\}, \quad 1 + \mathbf{x}x > 0,$$

sirve como una representación para todo $\xi \in \tilde{\mathbf{n}}$.

El siguiente teorema es uno de los resultados básicos de la teoría de valores extremos. De una manera analítica, da la información esencial, proporcionada por el máximo dominio de atracción. Primero recordamos la noción de la función cuantil F^{\leftarrow} de F y definimos

$$U(t) = F^{\leftarrow}(1 - t^{-1}), \quad t > 0.$$

Teorema 2.18 Caracterización del MDA($H_{\mathbf{x}}$)

Para $\xi \in \tilde{\mathbf{n}}$ los siguientes literales son equivalentes:

- a) $F \in \text{MDA}(H_{\mathbf{x}})$.
- b) Existe una función positiva, medible $a(\cdot)$ tal que para $1 + \xi x > 0$,

$$\lim_{u \uparrow x_F} \frac{\overline{F}(u + xa(u))}{\overline{F}(u)} = \begin{cases} (1 + \mathbf{x}x)^{-1/\mathbf{x}} & \text{si } \mathbf{x} \neq 0, \\ e^{-x} & \text{si } \mathbf{x} = 0, \end{cases} \quad (2.13)$$

- c) Para $x, y > 0, y \neq 1$

$$\lim_{s \rightarrow \infty} \frac{U(sx) - U(s)}{U(sy) - U(s)} = \begin{cases} \frac{x^x - 1}{y^x - 1} & \text{si } x \neq 0, \\ \frac{\ln x}{\ln y} & \text{si } x = 0. \end{cases} \quad (2.14)$$

La demostración de este teorema se la puede encontrar en Haan, L(1984) En Variación Regular y Sus Aplicaciones ♣

Nota: la condición (2.13) tiene una interesante interpretación probabilística. En efecto, sea X una v.a con f.d. $F \in \text{MDA}(H_x)$, entonces a (2.13) se la reformula como

$$\lim_{u \uparrow x_F} P\left(\frac{X - u}{a(u)} > x \mid X > u\right) = \begin{cases} (1 + x\mathbf{x})^{-1/x} & \text{si } x \neq 0, \\ e^{-x} & \text{si } x = 0. \end{cases} \quad (2.15)$$

de aquí (2.15) da una distribución aproximada de excesos sobre un umbral u . El factor escalar apropiado es $a(u)$. Esta interpretación será crucial en muchas aplicaciones. ♣

En esta nota usamos la noción de exceso. La siguiente definición habla más de esto.

Definición 2.19 Función de Distribución de Exceso, función de exceso media

Sea X una v.a con f.d. F y punto final derecho x_F . Para $u < x_F$, u fijo,

$$F_u(x) = P(X - u \leq x \mid X > u), \quad x \geq 0, \quad (2.16)$$

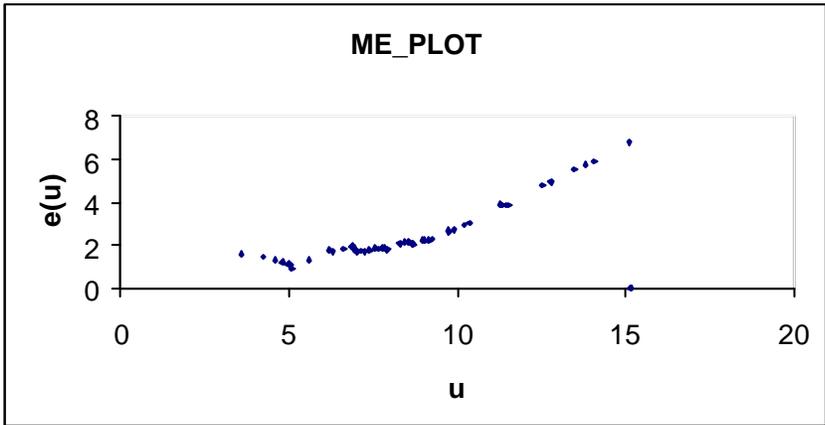
es la función de *distribución de exceso* de la v.a X sobre el umbral u . La función

$$e(u) = E(X - u \mid X > u)$$

es llamada *función de exceso media* de X . ♣

Gráfico No. 2.2

Función de exceso media



Este gráfico representa la función de exceso medio para distintos valores del umbral u

El exceso sobre el umbral juega un papel importante en muchos campos.

Tabla No. 2.20

Función de exceso media para algunas distribuciones estándar

Pareto	$\frac{k + u}{a - 1}, \quad a > 1$
Burr	$\frac{u}{at - 1}(1 + o(1)), \quad at > 1$
Loggamma	$\frac{u}{a - 1}(1 + o(1)), \quad a > 1$
Lognormal	$\frac{s^2 u}{\ln u - m}(1 + o(1))$
Benktander-tipo-I	$\frac{u}{a + 2b \ln u}$
Benktander-tipo-II	$\frac{u^{1-b}}{a}$

Weibull	$\frac{u^{1-r}}{ct} (1 + o(1))$
Exponencial	1^{-1}
Gamma	$b^{-1} \left(1 + \frac{a-1}{ub} + o\left(\frac{1}{u}\right) \right)$
Normal Truncada	$u^{-1} (1 + o(1))$

Ejemplo 2.21 Cálculo de la función de exceso media

Usando la definición de $e(u)$ y la integración por partes, las siguientes fórmulas son fáciles de obtener. Ellas son útiles para calcular la función de exceso media en casos especiales. Supongamos por facilidad de representación que X es un v.a con f.d. F y esperanza finita. Entonces

$$\begin{aligned}
 e(u) &= \int_u^{x_F} (x-u) dF(x) / \bar{F}(u) \\
 &= \frac{1}{\bar{F}(u)} \int_u^{x_F} \bar{F}(x) dx, \quad 0 < u < x_F.
 \end{aligned}
 \tag{2.17}$$

siempre que F sea continua,

$$\bar{F}(x) = \frac{e(0)}{e(x)} \exp \left\{ - \int_0^x \frac{1}{e(u)} du \right\}, \quad x > 0.
 \tag{2.18}$$

Se sigue inmediatamente de (2.18) que una f.d. continua es únicamente determinada por su función de exceso media. ♣

Definición 2.22 La distribución Generalizada de Pareto (DGP)

Definimos la f.d. G_x por

$$G_x(x) = \begin{cases} 1 - (1 + x\mathbf{x})^{-1/x} & \text{si } \mathbf{x} \neq 0, \\ 1 - e^{-x} & \text{si } \mathbf{x} = 0, \end{cases} \quad (2.19)$$

donde

$$\begin{aligned} x \geq 0 & \quad \text{si } \xi \geq 0, \\ 0 \leq x \leq -1/\xi & \quad \text{si } \xi < 0. \end{aligned}$$

G_x es llamada la distribución Generalizada de Pareto estándar (DGP). Se puede introducir la familia de localización y escala $G_{x;v,b}$ reemplazando el argumento de x anterior por $(x - v)/\beta$ para $v \in \tilde{\mathbf{n}}, \beta > 0$. ♣

Como en el caso de H_v , G_b puede ser también interpretado como el límite de G_x cuando $\xi \rightarrow 0$. Entonces, para economizar la notación, notaremos

$$G_{x;b}(x) = 1 - \left(1 + \mathbf{x} \frac{x}{b}\right)^{-1/x}, \quad x \in D(\mathbf{x}, \mathbf{b}), \quad (2.20)$$

donde

$$x \in D(\mathbf{x}, \mathbf{b}) = \begin{cases} [0, \infty) & \text{si } \mathbf{x} \geq 0, \\ [0, -\mathbf{b}/\mathbf{x}] & \text{si } \mathbf{x} < 0. \end{cases}$$

siempre que X tiene una DGP con parámetros ξ y β , esto es entendido como que X tiene f.d. $G_{x;b}$.

La distribución de Pareto generalizada ajustada es uno de los conceptos más útiles en la estadística de eventos extremos.

Teorema 2.23 Propiedades de la DGP

a) Supongamos que X sigue una DGP con parámetros ξ y β . Entonces $EX < \infty$ si y

solo si $\xi < 1$. En el último caso

$$E\left(1 + \frac{\mathbf{x}}{\mathbf{b}} X\right)^{-r} = \frac{1}{1 + r\mathbf{x}}, \quad r > -1/\mathbf{x},$$

$$E\left(\ln\left(1 + \frac{\mathbf{x}}{\mathbf{b}} X\right)\right)^k = \mathbf{x}^k k!, \quad k \text{ número natural}$$

$$EX(\bar{G}_{\mathbf{x};\mathbf{b}}(x))^r = \frac{\mathbf{b}}{(r+1-\mathbf{x})(r+1)}, \quad (r+1)/|\mathbf{x}| > 0.$$

si $\xi < 1/r$ con $r \in \mathbb{N}$, entonces

$$EX^r = \frac{\mathbf{b}^r \Gamma(\mathbf{x}^{-1} - r)}{\mathbf{x}^{r+1} \Gamma(1 + \mathbf{x}^{-1})} r!.$$

b) Para todo $\xi \in \mathbb{N}$, $F \in \text{MDA}(H_{\mathbf{x}})$ si y solo si

$$\lim_{u \uparrow x_F} \sup_{0 < x < x_F - u} |F_u(x) - G_{\mathbf{x};\mathbf{b}(u)}(x)| = 0 \quad (2.21)$$

para algunas funciones positivas β .

c) Supongamos $x_i \in D(\xi, \beta)$, $i = 1, 2$, entonces

$$\frac{\bar{G}_{\mathbf{x};\mathbf{b}}(x_1 + x_2)}{\bar{G}_{\mathbf{x};\mathbf{b}}(x_1)} = \bar{G}_{\mathbf{x};\mathbf{b} + \mathbf{x}\beta}(x_2). \quad (2.22)$$

d) Supongamos que N es $\text{Poi}(\lambda)$, independiente de la sucesión iid. (X_n) con una DGP con parámetros ξ y β . Escribimos $M_n = \max(X_1, X_2, \dots, X_N)$. Entonces

$$P(M_N \leq x) = \exp\left\{-I\left(1 + \mathbf{x} \frac{x}{\mathbf{b}}\right)^{-1/\mathbf{x}}\right\} = H_{\mathbf{x};u,\mathbf{y}}(x),$$

donde $\mu = \mathbf{b}\mathbf{x}^{-1}(\mathbf{I}^{\mathbf{x}} - 1)$ y $\Psi = \beta\lambda^{\xi}$.

e) Supongamos que X sigue una DGP con parámetros $\xi < 1$ y β . Entonces para $u < x_F$,

$$e(u) = E(X - u \mid X > u) = \frac{\mathbf{b} + u\mathbf{x}}{1 - \mathbf{x}}, \quad \mathbf{b} + u\mathbf{x} > 0.$$

Demostración

a) y c) siguen por verificación directa.

b) Previamente hemos probado en el teorema 2.18 que $F \in MDA(H_z)$ si y solo si

$$\lim_{u \uparrow x_F} |F_u(x) - G_{\mathbf{x}, \mathbf{b}(u)}(x)| = 0$$

donde $\beta(u) = a(u)$. Porque la DGP es continua, la uniformidad de la convergencia se obtiene

$$\lim_{u \uparrow x_F} \sup_{0 < x < x_F - u} |F_u(x) - G_{\mathbf{x}, \mathbf{b}(u)}(x)| = 0$$

para algunas funciones positivas β .

d) inmediatamente obtenemos

$$\begin{aligned} P(M_n \leq x) &= \sum_{n=0}^{\infty} e^{-1} \frac{\mathbf{I}^n}{n!} G_{\mathbf{x}, \mathbf{b}}^n(x) \\ &= \exp \left\{ -\mathbf{I} \left(1 + \mathbf{x} \frac{x}{\mathbf{b}} \right)^{-1/\mathbf{x}} \right\} \\ &= \exp \left\{ -\mathbf{I} \bar{G}_{\mathbf{x}, \mathbf{b}}(x) \right\} \\ &= \exp \left\{ -\mathbf{I} \left(1 + \mathbf{x} \frac{x - \mathbf{x}^{-1} \mathbf{b}(\mathbf{I}^{\mathbf{x}} - 1)}{\mathbf{b}\mathbf{I}^{\mathbf{x}}} \right)^{-1/\mathbf{x}} \right\}, \quad \mathbf{x} \neq 0. \end{aligned}$$

el caso $\xi = 0$ reduce a

$$P(M_n \leq x) = \exp \left\{ -e^{-(x-b \ln I)/b} \right\}$$

e) este resultado sigue inmediatamente de

$$\begin{aligned} e(u) &= \int_u^{x_F} (x-u) dF(x) / \bar{F}(u) \\ &= \frac{1}{\bar{F}(u)} \int_u^{x_F} dF(x), \quad x > 0. \quad \clubsuit \end{aligned}$$

2.1.4.1 Ventajas y desventajas del modelo de Pareto

La distribución de Pareto es muy útil en muchas situaciones. Sus principales ventajas son las siguientes:

- La distribución está caracterizada con un solo parámetro. Esto hace más fácil obtener una intuición de su influencia.
- Matemáticamente, ésta es fácil de manejar. Todos los problemas pueden ser calculados con la ayuda de fórmulas relativamente simples.
- Se usa exactamente, y hay una buena cantidad de conocimientos en típicos valores de parámetros para ciertos perfiles.

Sin embargo, la distribución de Pareto tiene una desventaja:

- La distribución de Pareto tiene sólo un parámetro. Así, esto no es muy flexible al aproximar tan cerca como sea posible la distribución de pérdida que actualmente ocurre en el portafolio. Frecuentemente, por ejemplo, hay distribuciones de pérdidas en el mundo real en las cuales tamaños de pérdidas medianas son más probables que pérdidas grandes o pequeñas. Esto significa que las funciones de distribución de aquellas distribuciones tienen una inflexión o puntos de retorno. La distribución de Pareto no puede satisfacer sus requerimientos porque la probabilidad decrece continuamente conforme el tamaño de la pérdida crece.

2.2 Fluctuaciones de las estadísticas de orden

2.2.1 Preliminares

A continuación se extiende la teoría estudiada en la sección 2.1 al comportamiento común de muchas estadísticas de orden superior, recalcando que estos proveen información en la cola derecha de una función de distribución. Se habla de las propiedades de las estadísticas de orden de un proceso homogéneo de Poisson. Además, se realiza una pequeña síntesis de lo que son los extremos para datos dependientes y la Teoría de Valores Extremos para sucesiones estacionarias.

A continuación se define el conjunto ordenado

$$X_{n,n} \leq \dots \leq X_{1,n}$$

Donde $X_{n,n} = \min(X_1, \dots, X_n)$ y $X_{1,n} = M_n = \max(X_1, \dots, X_n)$. La variable aleatoria $X_{k,n}$ es llamada la estadística de orden superior k -ésimo.

Para $x \in \tilde{n}$ introducimos la función de distribución empírica

$$F_n(x) = \frac{1}{n} \text{card}\{i : 1 \leq i \leq n, X_i \leq x\} = \frac{1}{n} \sum_{i=1}^n I_{\{X_i \leq x\}}, \quad x \in \tilde{n}$$

donde I_A es la función indicadora del conjunto A . Ahora,

$$X_{k,n} \leq x \text{ si y sólo si } \sum I_{\{X_i > x\}} < k, \text{ lo que implica que}$$

$$P(X_{k,n} \leq x) = P(F_n(x) > 1 - \frac{k}{n}).$$

Para un conjunto X_1, \dots, X_n se nota la función de cuantil empírica por F_n^{\leftarrow} . Particularmente

$$F_n^{\leftarrow}(t) = X_{k,n} \quad \text{para } 1 - \frac{k}{n} < t \leq 1 - \frac{k-1}{n}, \text{ para } k = 1, \dots, n.$$

Nota: $F^{\leftarrow}(t) = \inf \{x \in R : F(x) \geq t\}$ $0 < t < 1$. Es la función cuantil de la función de distribución F .

Proposición 2.24 Función de Distribución de las estadísticas de orden superior

k-ésimo.

Para $k = 1, \dots, n$ $F_{k,n}$ es la función de distribución de $X_{k,n}$. Entonces

$$a) \quad F_{k,n}(x) = \sum_{r=0}^{k-1} \binom{n}{r} \bar{F}^r(x) F^{n-r}(x). \quad (\bar{F} \text{ es el complemento de } F. \quad \bar{F} = 1 - F)$$

b) Si F es continua, entonces

$$F_{k,n}(x) = \int_{-\infty}^x f_{k,n}(z) dF(z),$$

donde

$$f_{k,n}(x) = \frac{n!}{(k-1)!(n-k)!} F^{n-k}(x) \bar{F}^{k-1}(x);$$

$f_{k,n}$ es la densidad de $F_{k,n}$ con respecto a F .

Demostración

a) Para $n \in \mathbb{N}$ se define

$$B_n = \sum_{i=1}^n I_{\{X_i > x\}}.$$

Es decir que B_n es una suma de n variables iid. de Bernoulli con probabilidad de éxito

$$EI_{\{X > x\}} = P(X > x) = \bar{F}(x).$$

De esto, se sigue que para todo $x \in \tilde{\mathfrak{N}}$

$$F_{k,n}(x) = P(B_n < k)$$

$$\begin{aligned}
 &= \sum_{r=0}^{k-1} P(B_n = r) \\
 &= \sum_{r=0}^{k-1} \binom{n}{r} \bar{F}^r(x) F^{n-r}(x).
 \end{aligned}$$

b) Usando la continuidad de F, calculamos

$$\begin{aligned}
 &\frac{n!}{(k-1)!(n-k)!} \int_{-\infty}^x F^{n-k}(z) \bar{F}^{k-1}(z) dF(z) \\
 &= \frac{n!}{(k-1)!(n-k)!} \int_{\bar{F}(x)}^1 (1-t)^{n-k} t^{k-1} dt \\
 &= \sum_{r=0}^{k-1} \binom{n}{r} \bar{F}^r(x) F^{n-r}(x) = F_{k,n}(x) \mathbb{S}
 \end{aligned}$$

Teorema 2.25 Densidad conjunta de las estadísticas de orden superior k

Si F es absolutamente continua con densidad f , entonces

$$f_{X_{1,n}, \dots, X_{k,n}}(x_1, \dots, x_k) = \frac{n!}{(n-k)!} F^{n-k}(x_k) \prod_{i=1}^k f(x_i), \quad x_k < \dots < x_1.$$

Lema 2.26 Transformación Cuantil

Sean X_1, \dots, X_n iid. con función de distribución F. Además sean U_1, \dots, U_n v.a iid. uniformemente en (0, 1) y denotadas por $U_{n,n} < \dots < U_{1,n}$ las correspondientes estadísticas de orden. Entonces se obtienen los siguientes resultados:

1. $F^{\leftarrow}(U_1) = X_1$.
2. Para todo $n \in \mathbb{N}$,

$$(X_{n,n} \leq \dots \leq X_{1,n}) = (F^{\leftarrow}(U_{1,n}), \dots, F^{\leftarrow}(U_{n,n}))$$

3. La v.a $F(X_1)$ tiene una distribución uniforme en $(0,1)$ si y solo si F es una función continua.

Observación: La transformación cuantil liga la distribución uniforme a otra distribución F . Una aplicación inmediata de este resultado es la generación de números aleatorios. Así, por ejemplo, los números aleatorios exponenciales pueden ser obtenidos de números aleatorios uniformes utilizando la transformación $E_1 = -\ln(1 - U_1)$.

Teorema 2.27 Ley límite para el número de excedentes

Supongamos que (u_n) es una sucesión en \mathbb{R} tal que $n\bar{F}(u_n) \rightarrow t$ para algún $t \in [0, \infty]$ cuando $n \rightarrow \infty$. Entonces

$$\lim_{n \rightarrow \infty} P(B_n \leq k) = e^{-t} \sum_{r=0}^k \frac{t^r}{r!}, \quad k \in \mathbb{N}_0$$

Demostración

Para $t \in (0, \infty)$, es simple por el teorema de límite de Poisson como indicamos anteriormente. Para $t = 0$, tenemos

$$P(B_n \leq k) \geq P(B_n = 0) = (1 - \bar{F}(u_n))^n = \left(1 + o\left(\frac{1}{n}\right)\right)^n \rightarrow 1.$$

Para $t = \infty$ tenemos para un $\theta > 0$ arbitrario que $n\bar{F}(u_n) \geq \theta$ para n grande. Ya que la f.d. binomial es decreciente en θ , obtenemos

$$P(B_n \leq k) \leq \sum_{r=0}^k \binom{n}{r} \left(\frac{\theta}{n}\right)^r \left(1 - \frac{\theta}{n}\right)^{n-r}.$$

Así para k fija,

$$\limsup_{n \rightarrow \infty} P(B_n \leq k) \leq e^{-q} \sum_{r=0}^k \frac{q^r}{r!} \rightarrow 0, \quad q \rightarrow \infty.$$

de aquí $P(B_n \leq k) \rightarrow 0$ cuando $n \rightarrow \infty$

Para lo contrario suponemos que $\lim_{n \rightarrow \infty} P(B_n \leq k) = e^{-t} \sum_{r=0}^k \frac{t^r}{r!}$, se cumple para $k \in \mathbb{N}_0$, pero $n\bar{F}(u_n)$ no tiende a τ .

Entonces existe un $\tau' \neq \tau$ en $[0, \infty]$ y una subsucesión (n_k) tal que $n_k \bar{F}(u_{n_k}) \rightarrow \tau'$ cuando $k \rightarrow \infty$, y así B_{n_k} converge débilmente a una v.a de Poisson con parámetro τ' , contradiciendo a

$$\lim_{n \rightarrow \infty} P(B_n \leq k) = e^{-t} \sum_{r=0}^k \frac{t^r}{r!}, \quad k \in \mathbb{N}_0$$

Corolario 2.28 Distribución límite de una estadística de orden superior

Supongamos $F \in \text{MDA}(H)$ con constantes $c_n > 0$ y $d_n \in \tilde{\mathfrak{N}}$. Se define,

$$H^{(k)}(x) = H(x) \sum_{r=0}^{k-1} \frac{(-\ln H(x))^r}{r!}, \quad x \in \tilde{\mathfrak{N}}. \text{ Entonces,}$$

para cada $k \in \mathbb{N}$, se sigue que

$$\lim_{n \rightarrow \infty} P(c_n^{-1}(X_{k,n} - d_n) \leq x) = H^{(k)}(x). \tag{2.23}$$

Por otro lado, si para algún $k \in \mathbb{N}$

$$\lim_{n \rightarrow \infty} P(c_n^{-1}(X_{k,n} - d_n) \leq x) = G(x), \quad x \in \tilde{\mathfrak{N}}$$

para una función de distribución G no degenerada, entonces $G = H^{(k)}$, para alguna distribución de valor extremo H y (2.23) se cumple para todo $k \in \mathbb{N}$. ♣

Aplicación 2.29 Estadísticas de orden superior de la distribución de Gumbel

Por integración parcial,

$$H^{(k)}(x) = \frac{1}{(k-1)!} \int_{-\ln H(x)}^{\infty} e^{-t} t^{k-1} dt = \Gamma_k(-\ln H(x)), \quad x \in \tilde{\mathfrak{N}}$$

donde Γ_k denota la función Gamma incompleta. En particular, si H es la distribución de Gumbel, entonces

$$\Lambda^{(k)}(x) = \frac{1}{(k-1)!} \int_{e^{-x}}^{\infty} e^{-t} t^{k-1} dt = P\left(\sum_{i=1}^k E_i > e^{-x}\right)$$

para E_1, \dots, E_k v.a iid exponenciales estándar. Así, si $Y^{(k)}$ tiene f.d. $\Lambda^{(k)}$, entonces

$$Y^{(k)} \stackrel{d}{=} -\ln \sum_{i=1}^k E_i.$$

La distribución límite de la k -ésima estadística de orden superior fue obtenida considerando el número de excedentes del nivel u_n de X_1, \dots, X_n . ♣

Teorema 2.30 Distribución límite de la estadística de orden k superior en un conjunto aleatoriamente indexado

Supongamos que $N(t)/t \xrightarrow{p} Z$ se cumple para una v.a no negativa Z con función de distribución F_Z y que

$$\lim_{n \rightarrow \infty} n \bar{F}(c_n x + d_n) = -\ln H(x), \quad x \in \tilde{\mathfrak{N}} \text{ se satisface. Entonces,}$$

$$\begin{aligned} & \lim_{n \rightarrow \infty} P(c_n^{-1}(X_{k, N(n)} - d_n) \leq x) \\ &= \int_0^{\infty} \Gamma_k(-z \ln H(x) dF_Z(z)) \quad x \in \tilde{\mathfrak{n}} \\ &= E[\Gamma_k(-\ln H^Z(x))] \end{aligned}$$

Demostración

Escribimos $B_n = \sum_{j=1}^{N(n)} I_{\{X_j > c_n x + d_n\}}$, aplicando la probabilidad condicional se

obtiene que

$$\begin{aligned} P(c_n^{-1}(X_{k, N(n)} - d_n) \leq x | N(n)) &= P(B_n \leq k - 1 | N(n)) \\ &= \sum_{i=0}^{k-1} \binom{N(n)}{i} (F(c_n x + d_n))^{N(n)-i} (\bar{F}(c_n x + d_n))^i \\ &= (1 + \mathbf{or}(1)) \sum_{i=0}^{k-1} \frac{1}{i!} \left(\frac{N(n)}{n} (n \bar{F}(c_n x + d_n)) \right)^i \exp \left\{ \frac{N(n)}{n} (n \ln(1 - \bar{F}(c_n x + d_n))) \right\} \\ &\xrightarrow{P} \sum_{i=0}^{k-1} \frac{1}{i!} (-Z \ln H(x))^i e^{Z \ln H(x)} \\ &= H^Z(x) \sum_{i=0}^{k-1} \frac{(-\ln H^Z(x))^i}{i!}. \end{aligned}$$

Luego, tomando las esperanzas, se completa la demostración. ♣

2.3 Síntesis de la Teoría de Valores Extremos para sucesiones estacionarias

Teorema 2.31 Ley límite para el máximo de una serie estacionaria

Supongamos que $c_n^{-1}(M_n - d_n) \xrightarrow{d} G$ para alguna distribución G y constantes apropiadas $c_n > 0$, $d_n \in \tilde{\mathfrak{n}}$. Si se cumple la condición $D(c_n x + d_n)$ para todo real x , entonces G es una distribución de valor extremo.

Demostración

Recalamos en primer lugar de lo estudiado anteriormente que F es una distribución de valor extremo si y solo si F es max - estable. Luego,

$$P(M_{nk} \leq c_n x + d_n) = P^k(M_n \leq c_n x + d_n) + o(1) \rightarrow F^k(x)$$

para todo $k \geq 1$, y todo punto de continuidad x de F . Por otro lado,

$$P(M_{nk} \leq c_{nk} x + d_{nk}) \rightarrow F(x)$$

Por la convergencia del teorema del anexo A.7, existen constantes $\tilde{c}_k > 0$ y $\tilde{d}_k \in \mathfrak{N}$ tal que

$$l \lim_{n \rightarrow \infty} \frac{c_{nk}}{c_n} = \tilde{c}_k \text{ y } l \lim_{n \rightarrow \infty} \frac{d_{nk} - d_n}{c_n} = \tilde{d}_k,$$

para Y_1, \dots, Y_k v.a iid. con f.d. H ,

$$\max(Y_1, \dots, Y_k) \stackrel{d}{=} \tilde{c}_k Y_1 + \tilde{d}_k. \clubsuit$$

La condición $D(u_n)$ es una condición más débil que muchas de las formas clásicas de restricciones de dependencia.

D(u_n): Para enteros p, q y n

$$1 \leq i_1 < \dots < i_p < j_1 < \dots < j_q \leq n$$

tal que $j_1 - i_p \geq l$ se tiene que

$$\left| P\left(\max_{i \in A_1 \cup A_2} X_i \leq u_n\right) - P\left(\max_{i \in A_1} X_i \leq u_n\right) P\left(\max_{i \in A_2} X_i \leq u_n\right) \right| \leq \mathbf{a}_{n,l},$$

donde $A_1 = \{i_1, \dots, i_p\}$, $A_2 = \{j_1, \dots, j_q\}$ y $\mathbf{a}_{n,l} \rightarrow 0$ cuando $n \rightarrow \infty$ para alguna sucesión $l = l_n = o(n)$.

Ahora, se quiere encontrar condiciones suficientes para la convergencia de las probabilidades $P(M_n \leq u_n)$ para una sucesión de umbrales dada (u_n) que satisfice

$$n\bar{F}(u_n) \rightarrow t \quad (2.24)$$

para algún $t \in [0, \infty)$.

Desafortunadamente, no se puede reemplazar (\tilde{M}_n) por (M_n) sobre $D(u_n)$, todo lo que se puede derivar es

$$\liminf_{n \rightarrow \infty} P(M_n \leq u_n) \geq e^{-t}$$

A continuación se introduce una segunda condición técnica

$D'(u_n)$: La relación

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} n \sum_{j=2}^{[n/k]} P(X_1 > u_n, X_j > u_n) = 0$$

$D'(u_n)$ es una condición de anti – agrupamiento en la sucesión estacionaria (X_n) . Hay que notar que $D'(u_n)$ implica que

$$E \sum_{1 \leq i < j \leq [n/k]} I_{\{X_i > u_n, X_j > u_n\}} \leq [n/k] \sum_{j=2}^{[n/k]} E I_{\{X_i > u_n, X_j > u_n\}} \rightarrow 0$$

Teorema 2.32 Distribución límite del máximo de una serie estacionaria

Sea (X_n) una sucesión estacionaria con una función de distribución $F \in \text{MDA}(H)$ para alguna distribución de valor extremo H , existen constantes $c_n > 0, d_n \in \tilde{\mathfrak{n}}$ tal que

$$\lim_{n \rightarrow \infty} n\bar{F}(c_n x + d_n) = -\ln H(x), \quad x \in \tilde{\mathfrak{n}}$$

Asumiendo que para $x \in \tilde{\mathfrak{n}}$ las sucesiones $(u_n) = (c_n x + d_n)$ satisfacen las condiciones $D(u_n)$ y $D'(u_n)$. Entonces la igualdad anterior es equivalente a cada una de las siguientes relaciones:

$$\begin{aligned} c_n^{-1}(Mn - d_n) &\xrightarrow{d} H, \\ c_n^{-1}(\bar{M}n - d_n) &\xrightarrow{d} H. \end{aligned} \quad \mathbb{S}$$

Para finalizar con este capítulo, presentamos a continuación un lema donde se detallan las formas de chequear las condiciones $D(u_n)$ y $D'(u_n)$.

Lema 2.33 Condiciones para $D(u_n)$ y $D'(u_n)$ para una sucesión gaussiana estacionaria

Asumamos que (X_n) es Gaussiana estacionaria y sea (u_n) una sucesión de números reales.

1) Supongamos que el lado derecho de

$$\left| P(M_n \leq u_n) - \Phi^n(u_n) \right| \leq \text{const} \quad n \sum_{h=1}^n |\mathbf{g}(h)| \exp \left\{ \frac{-u_n^2}{1 + |\mathbf{g}(h)|} \right\} \text{ tiende a}$$

cero cuando $n \rightarrow \infty$ y $\sup_{h \geq 1} |\mathbf{g}(h)| < 1$. Entonces $D(u_n)$ se cumple.

2) Si adicionalmente $\limsup_{n \rightarrow \infty} n \bar{\Phi}(u_n) < \infty$. Entonces $D'(u_n)$ se cumple.

3) Si $\mathbf{g}(n) \ln n \rightarrow 0$ y $\limsup_{n \rightarrow \infty} n \bar{\Phi}(u_n) < \infty$ entonces ambas condiciones $D(u_n)$ y $D'(u_n)$ son satisfechas. ♣

La suposición $\mathbf{g}(n) \ln n \rightarrow 0$ es llamada condición de Berman y esta condición es bastante débil, especialmente en el caso gaussiano.

2.4 Estudio de los extremos por medio de procesos puntuales

La técnica de los procesos puntuales nos da una guía dentro de la estructura de las variables límite y procesos límite los cuales ocurren en la Teoría de Sumas, en la Teoría de Valores Extremos y en el análisis de Series de Tiempo. Mediante distintos métodos se pueden encontrar excelentes resultados en excedentes, límites de las estadísticas de orden superior, convergencia conjunta del máximo y mínimo, récord, etc.

Para comprender de lo que tratan las técnicas de los procesos puntuales es necesario conocer Análisis Funcional, Teoría de la Medida y ciertos argumentos de convergencia débil. En esta sección hemos tratado de disminuir al mínimo estos aspectos para que la comprensión sea mayor.

Una idea simple de un proceso puntual N es tomarlo como una distribución aleatoria de puntos X_i en el espacio. Para una configuración dada (X_i) y un conjunto A , $N(A)$ cuenta el número de $X_i \in A$. Es conveniente imaginar la distribución de N como las probabilidades

$$P(N(A_1)=k_1, \dots, N(A_m)=k_m)$$

de todas las posibles elecciones de “conjuntos buenos” A_1, \dots, A_m y todos los enteros no negativos k_1, \dots, k_m .

Los procesos puntuales más importantes son aquellos para los cuales $N(A)$ sigue una distribución de Poisson. Esto conduce a la medida aleatoria de Poisson N como una generalización del clásico proceso homogéneo de Poisson en $[0, \infty)$.

2.4.1 Notas básicas sobre los procesos puntuales

2.4.1.1 Definición

Consideremos una sucesión (X_n) de vectores aleatorios en el *espacio de estado* E , donde $A \subset E$. Naturalmente, $N(A) = N(A, \omega)$ es aleatorio para un conjunto dado A , y sobre ciertas condiciones, $N(\cdot, \omega)$ define una medida aleatoria de conteo con puntos X_n en una σ -álgebra \mathbf{e} de subconjuntos de E . El espacio de estado E , es un subconjunto de un espacio euclidiano de dimensión finita, E está equipado con la σ -álgebra \mathbf{e} de los conjuntos borelianos generados por los abiertos.

Es conveniente escribir un proceso puntual usando la medida de dirac ε_x para $x \in E$:

$$\varepsilon_x(A) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A, \end{cases} \quad A \in \mathbf{e}$$

Para una sucesión $(x_i)_{i \geq 1}$ en E ,

$$m(A) = \sum_{i=1}^{\infty} \mathbf{e}_{x_i}(A) = \sum_{i: x_i \in A} 1 = \text{card}\{i : x_i \in A\}, \quad A \in \mathbf{e},$$

define una medida contable en \mathbf{e} la cual es llamada medida puntual si $m(K) < \infty$ para todos los conjuntos compactos $K \in E$.

$M_p(E)$ es el espacio de todas las medidas puntuales en E provistas de una σ -álgebra apropiada $M_p(E)$.

Los procesos puntuales que estamos interesados pueden escribirse a menudo de la forma

$$N = \sum_{i=1}^{\infty} \mathbf{e}_{X_i}$$

para una sucesión (X_n) de vectores aleatorios d -dimensionales. Entonces, para cada $\omega \in \Omega$,

$$N(A, \mathbf{w}) = \sum_{i=1}^{\infty} \mathbf{e}_{X_i(\mathbf{w})}(A), \quad A \in \mathbf{e},$$

define una medida puntual en \mathbf{e} .

Ejemplo 2.34 Procesos puntuales de excedentes

Estos procesos están muy relacionados con la Teoría de Valores Extremos.

Sea u un número real y (X_n) una sucesión de v.a. Entonces el Proceso Puntual de Excedentes

$$N_n(\cdot) = \sum_{i=1}^n \mathbf{e}_{n^{-1} \cdot}(\cdot) I_{\{X_i > u\}}, \quad n = 1, 2, \dots, \quad (2.25)$$

con $E = (0, 1]$ cuenta el número de excedentes del umbral u de la sucesión X_1, \dots, X_n . Por ejemplo, tomando todo el intervalo $(0, 1]$. Entonces,

$$N_n(0,1] = \text{card}\{i: 0 < n^{-1}i \leq 1 \quad \text{y} \quad X_i > u\} \\ = \text{card}\{i \leq n: X_i > u\}.$$

Se puede ver inmediatamente la relación estrecha que existe con la TVE. Por ejemplo, sea $X_{k,n}$ la k -ésima estadística de orden más grande del conjunto X_1, \dots, X_n . Entonces,

$$\{N_n(0,1] = 0\} = \{\text{card}\{i \leq n: X_i > u\} = 0\} \\ = \{\text{Ningún } X_i, i \leq n, \text{ excede } u\} \\ = \{\max(X_1, \dots, X_n) \leq u\} \quad (2.26)$$

$$\{N_n(0,1] < k\} = \{\text{card}\{i \leq n: X_i > u\} < k\} \\ = \{\text{Menos de } k \text{ entre los } X_i, i \leq n, \text{ exceden } u\} \\ = \{X_{k,n} \leq u\}. \quad (2.27)$$

Notamos que el proceso puntual de excedentes puede ser escrito de la forma

$$N_n(\cdot) = \sum_{i=1}^n e_{n^{-1}i, X_i}(\cdot), \quad n = 1, 2, \dots, \quad (2.28)$$

con $E = (0,1] \cap (u, \infty)$ bi-dimensional. ♣

2.4.2 Procesos puntuales de excedentes

En esta sección se mostrará la convergencia débil de una sucesión (N_n) de los procesos puntuales de excedentes a un Proceso de Poisson Homogéneo N en el espacio de estado $E = (0,1]$.

La sucesión (X_n) es supone que es iid. o estrictamente estacionaria que satisface dos suposiciones D y D' Sección 2.3.

2.4.2.1 Caso IID

Asumamos que las X_n son v.a iid. y sea (u_n) una sucesión de umbrales reales. De la proposición 2.1 se tiene que, para cualquier $t \in [0, \infty]$, la relación

$$P(M_n \leq u_n) \rightarrow \exp\{-t\} \text{ se mantiene si y sólo si}$$

$$n\bar{F}(u_n) = E \sum_{i=1}^n I_{\{X_i > u_n\}} \rightarrow \mathbf{t} \quad (2.29)$$

Esta última condición asegura que hay en promedio de aproximadamente τ excedentes sobre el umbral u_n .

Teorema 2.35 Convergencia débil de procesos puntuales de excedentes, caso iid.

Supongamos que (X_n) es una sucesión de v.a iid. con función de distribución común F . Sea (u_n) una sucesión de umbrales tales que (2.27) se cumple para algún $\tau \in (0, \infty)$. Entonces el proceso puntual de excedentes N_n , converge débilmente en $M_p(E)$ a un proceso de Poisson Homogéneo N en $E = (0, 1]$ con intensidad τ . N es MAP($\tau \cdot | \cdot$), donde $| \cdot$ denota la medida de Lebesgue en E .

Demostración

Asumimos que el proceso límite N es un Proceso Homogéneo de Poisson en $[0, \infty)$. Aplicando el Teorema de Kallenberg (anexo A.8).

Notamos que para $A = (a, b] \subset (a, b]$ la v.a

$$\begin{aligned} N_n(A) &= \sum_{i=1}^n e_{n^{-1}i}(A) I_{\{X_i > u_n\}} \\ &= \sum_{a < n^{-1}i \leq b} I_{\{X_i > u_n\}} \\ &= \sum_{i=[na]+1}^{[nb]} I_{\{X_i > u_n\}} \end{aligned}$$

es binomial con parámetros $([nb] - [na], \bar{F}(u_n))$. Asumiendo (2.29),

$$EN_n(A) = (([nb] - [na])\bar{F}(u_n)) - (n(b-a))(n^{-1}\mathbf{t}) = \mathbf{t}(b-a) = EN(A),$$

lo que prueba que $EN_n(A) \rightarrow EN(A)$.

Así, basta probar que $P(N_n(B) = 0) \rightarrow P(N(B) = 0)$. Como $N_n(A)$ es binomial y como se cumple (2.29), tenemos que

$$\begin{aligned} P(N_n(A) = 0) &= F^{[nb]-[na]}(u_n) \\ &= \exp\{([nb] - [na]) \ln(1 - \bar{F}(u_n))\} \\ &\rightarrow \exp\{-t(b - a)\}. \end{aligned}$$

Recordando la definición del conjunto B del Teorema de Kallenberg y considerando la independencia de las X_i , concluimos del resultado anterior que

$$\begin{aligned} P(N_n(B) = 0) &= P(N_n(c_i, d_i] = 0, i = 1, \dots, k) \\ &= P\left(\max_{[nc_i] < j \leq [nd_i]} X_j \leq u_n, i = 1, \dots, k\right) \\ &= \prod_{i=1}^k P\left(\max_{[nc_i] < j \leq [nd_i]} X_j \leq u_n\right) \\ &= \prod_{i=1}^k P(N_n(c_i, d_i] = 0) \\ &\rightarrow \prod_{i=1}^k \exp\{-t(d_i - c_i)\} \end{aligned}$$

Por otro lado, por la propiedad de Poisson de N ,

$$P(N(B) = 0) = \exp\{-t|B|\} = \exp\left\{-t \sum_{i=1}^k (d_i - c_i)\right\}$$

Lo que prueba el teorema. ♣

Ejemplo 2.36 Continuación del ejemplo 2.34

Una aplicación de este último Teorema junto con las ecuaciones (2.26) y (2.27) nos lleva a que

$$P(X_{k:n} \leq u_n) = P(N_n(0,1] < k) \rightarrow P(N(0,1] < k) = e^{-t} \sum_{i=0}^{k-1} \frac{t^i}{i!}$$

Esto nos permite derivar la distribución límite de la k - ésima estadística de ordenado F en el MDA en una distribución de valor extremo. ♣

Si se tienen dos sucesiones de v.a $(X_i), (Y_i)$ iid., suponiendo que Y_i es positiva con probabilidad 1 y fijando $T_i = Y_1 + \dots + Y_i$. Entonces $N\mathcal{Q}(t) = \text{card} \{i : T_i \leq t\}$ define un proceso de renovación. Aquí consideramos el proceso puntual de excedentes correspondiente

$$\tilde{N}_n(\cdot) = \sum_{i=1}^{N'_n(n)} \mathbf{e}_{n^{-1}T_i}(\cdot) I_{\{X_i > u_n\}}, \quad (2.30)$$

en el espacio de estado $E = (0,1]$. La Ley fuerte de los grandes números implica que $n^{-1}T_{[nx]} \xrightarrow{c.s.} xEY_1 = xI^{-1}$ para $x \in (0,1]$ y entonces se puede esperar un resultado similar al del teorema 2.35.

2.4.2.2 Caso estacionario

En esta sección se hablará del problema de encontrar la distribución límite del máximo M_n y de las estadísticas de orden superior de un conjunto de una sucesión estrictamente estacionaria (X_n) . Asumimos que las condiciones $D(u_n)$ y $D\mathcal{Q}(u_n)$ se mantienen para la sucesión de umbrales (u_n) . La condición $D\mathcal{Q}(u_n)$ tiene una interpretación intuitiva en el lenguaje de los procesos puntuales: Si (u_n) es escogida para satisfacer $n\bar{F}(u_n) \rightarrow t \in (0, \infty)$ entonces existe un promedio aproximado de τ excedentes sobre u_n para X_1, \dots, X_n , y así τ/k entre $X_1, \dots, X_{[n/k]}$.

Teorema 2.37 Convergencia débil de procesos puntuales de excedentes, caso estacionario.

Supongamos que (X_n) es estrictamente estacionario y (u_n) es una sucesión de umbrales tal que (2.29), $D(u_n)$ y $D\mathcal{Q}(u_n)$ se cumplen. Sea (N_n) el proceso (2.25).

Entonces $N_n \xrightarrow{d} N$ en $M_\rho(E)$, donde N es una MAP homogénea en $E = (0,1]$ con intensidad \mathbf{t} .

Demostración

Aplicamos el Teorema de Kallenberg (anexo A.8). La prueba de que

$$EN_n(A) \rightarrow EN(A)$$

se la realiza de la misma forma que en la demostración del teorema 2.35. Entonces, basta probar que

$$P(N_n(B) = 0) \rightarrow P(N(B) = 0)$$

haciendo uso de $D(u_n)$ y $D'(u_n)$.

Por simplicidad, nos restringimos a los conjuntos $B = (c_1, d_1] \cup (c_2, d_2]$ con $0 < c_1 < d_1 < c_2 < d_2 \leq 1$. El caso general puede ser demostrado por analogía.

Tomando $(a, b] \subset (0,1]$. Usando la estacionariedad de (X_n) y el Anexo A.9 obtenemos que

$$\begin{aligned} P(N_n(a, b] = 0) &= P\left(\max_{i \in [nb] - [na]} X_i \leq u_n\right) \\ &\rightarrow \exp\{-\mathbf{t}(b-a)\} = P(N(a, b] = 0) \end{aligned} \quad (*)$$

De la condición $D(u_n)$ concluimos que

$$\begin{aligned} P(N_n(B) = 0) &= P(N_n(c_1, d_1] = 0, N_n(c_2, d_2] = 0) \\ &= P\left(\max_{c_1 < n^{-1}i \leq d_1} X_i \leq u_n, \max_{c_2 < n^{-1}i \leq d_2} X_i \leq u_n\right) \\ &= P\left(\max_{c_1 < n^{-1}i \leq d_1} X_i \leq u_n\right) P\left(\max_{c_2 < n^{-1}i \leq d_2} X_i \leq u_n\right) + o(1) \end{aligned}$$

La distancia entre dos conjuntos

$$A_1 = \{[nc_1] + 1, \dots, [nd_1]\} \text{ y } A_2 = \{[nc_2] + 1, \dots, [nd_2]\}$$

excede en $(c_2 - d_1)n > l_n = o(n)$ lo que implica que $\mathbf{a}_{n,l_n} \rightarrow 0$. Así, por (*),

$$P(N_n(B) = 0) \rightarrow \exp\{-t((d_1 - c_1) + (d_2 - c_2))\} = P(N(B) = 0),$$

Así, por el teorema de Kallenberg se prueba este teorema. ♣

Ejemplo 2.38 Probabilidades límite de las estadísticas de orden superior

Supongamos que las suposiciones del teorema anterior se mantienen. Entonces,

$$P(X_{k,n} \leq u_n) = P(N_n(0,1] < k) \rightarrow P(N(0,1] < k) = e^{-t} \sum_{i=0}^{k-1} \frac{t^i}{i!}$$

Ahora es inmediato que nosotros podamos derivar la distribución límite de una estadística de orden superior $X_{k,n}$ por la forma tradicional. Sea (\tilde{X}_n) una sucesión asociada iid. tal que X y \tilde{X} , tienen la misma distribución y se denota sus estadísticas de orden en la vía natural por $\tilde{X}_{k,n}$. ♣

El teorema del anexo A.15 muestra la semejanza entre el comportamiento asintótico de los extremos de la sucesión estacionaria (X_n) y de una sucesión asociada iid. (\tilde{X}_n) .

Lo importante ahora es generalizar estos resultados a un vector finito de estadísticas de orden. Esto es, obtener las probabilidades de la forma

$$P(X_{1,n} \leq u_n^{(1)}, \dots, X_{k,n} \leq u_n^{(k)})$$

de k sucesiones de números reales

$$u_n^{(k)} \leq \dots \leq u_n^{(1)} \tag{2.31}$$

Para esto es apropiado introducir un vector de k procesos puntuales de excedentes, uno para cada sucesión de umbrales. Como los excedentes de los niveles $u_n^{(i)}$ están muy relacionados unos con otros, ya que un excedente de $u_n^{(r)}$ es automáticamente un excedente de $u_n^{(r+1)}$ es posible por un argumento geométrico el reducir el problema de k excedentes a la convergencia débil de un proceso puntual sobre $(0,1] \times \tilde{n}$.

Antes de esto se introduce una condición análoga a $D(u_n)$ k – dimensional.

Suponemos dadas las k sucesiones de (2.31).

Condición $D_k(u_n)$: Para p, q fijos y para enteros cualquiera

$$1 \leq i_1 < \dots < i_p < j_1 < \dots < j_q \leq n$$

tal que $j_1 - i_p \geq l$ se tiene que

$$\left| P(X_{i_m} \leq u_n^{(s_m)}, m=1, \dots, p, X_{j_r} \leq u_n^{(s'_r)}, r=1, \dots, q) - P(X_{i_m} \leq u_n^{(s_m)}, m=1, \dots, p) P(X_{j_r} \leq u_n^{(s'_r)}, r=1, \dots, q) \right| \leq \alpha_{n,l},$$

para cualquier entero $1 \leq s_l, s_r \leq k$, y dado que $\alpha_{n,l} \rightarrow 0$ cuando $n \rightarrow \infty$ para alguna sucesión $l = l_n = o(n)$.

No es necesario extender la condición $D(u_n)$ pues se puede asumir que $D(u_n^{(i)})$ se mantiene separadamente para cada $i = 1, \dots, k$. Así, podemos escribir como en la Sección 2.2.

$$B_n^{(i)} = \sum_{i=1}^n I_{\{X_i > u_n^{(i)}\}}, \quad n \geq 1, i = 1, \dots, k,$$

al número de excedentes de $u_n^{(i)}$ para X_1, \dots, X_n .

Teorema 2.39 Convergencia débil conjunta del número de excedentes, caso estacionario.

Sea (X_n) una sucesión estrictamente estacionaria y supongamos que las sucesiones $(u_n^{(i)})$ satisfacen (2.31) y que $n\bar{F}(u_n^{(i)}) \rightarrow t_i \in (0, \infty)$, $i = 1, \dots, k$. Asumimos que las condiciones $D\mathcal{Q}(u_n^{(i)})$ y $D_k(u_n)$ se satisfacen.

Entonces, para $l_1, \dots, l_k \geq 0$,

$$P\left(B_n^{(1)} = \ell_1, B_n^{(2)} = \ell_1 + \ell_2, \dots, B_n^{(k)} = \ell_1 + \dots + \ell_k\right) \rightarrow \frac{t_1^{\ell_1}}{\ell_1!} \frac{(t_2 - t_1)^{\ell_2}}{\ell_2!} \dots \frac{(t_k - t_{k-1})^{\ell_k}}{\ell_k!} e^{-t_k}, \quad n \rightarrow \infty. \quad \clubsuit$$

Corolario 2.40 Ley límite conjunta de las estadísticas de orden superior, caso estacionario.

Asumamos que $F \in \text{MDA}(H)$ con constantes normalizadas $c_n > 0$ y constantes centradas $d_n \in \bar{\mathbb{N}}$. Además las condiciones $D\mathcal{Q}(u_n)$ y $D_k(u_n)$ son satisfechas por todas las sucesiones $u_n = c_n x + d_n$. Entonces, la relación límite

$$\left(c_n^{-1}(X_{i,n} - d_n)\right)_{i=1, \dots, k} \xrightarrow{d} \left(Y^{(i)}\right)_{i=1, \dots, k}, \quad k \geq 1, n \rightarrow \infty,$$

se cumple, donde $(Y^{(1)}, \dots, Y^{(k)})$ es la variable extrema k -dimensional correspondiente a la distribución de valor extremo H . \clubsuit

3. Principales herramientas estadísticas para eventos extremos en finanzas

3.1 Preliminares

El problema esencial consiste en determinar la distribución de los valores mínimos y máximos en muestras de tamaño n , obtenidas al azar de una cierta distribución subyacente dada $F(x)$. El punto importante radica en que los extremos no están fijos, pero son variables estadísticas nuevas que dependen de la distribución subyacente y del tamaño muestral.

En muchos problemas prácticos de ingeniería, no se conoce la forma exacta de la distribución y tienen que plantearse supuestos con respecto a su forma. En general son tres las distribuciones estándar que siguen los valores extremos: Weibull, Fréchet y Gumbel.

En general, cuando la función subyacente de densidad de probabilidad es $f(x)$ y la función de distribución acumulativa es $F(x)$, entonces la distribución del valor mínimo en muestras de tamaño n está dada por la función de densidad de probabilidad

$$h_n(x) = nf(x)[1 - F(x)]^{n-1}$$

y la función de distribución acumulativa correspondiente es

$$H_n(x) = 1 - [1 - F(x)]^n$$

la moda de $h_n(x)$ es la solución de la ecuación

$$\frac{d}{dx} h_n(x) = 0$$

Si la experiencia muestra que los logaritmos de una serie de picos están distribuidos de manera aproximadamente normal, que es lo que sucede en general en el caso de corrientes fluviales, la probabilidad de exceder un flujo Q (evento conocido como excedencia) es

$$P(Q) = \frac{1}{T_r} = 1 - \frac{1}{s_y \sqrt{2p}} \int_{-\infty}^y e^{-(y-\bar{y})^2 / 2s_y} dy$$

donde $y = \log x$ (x variable original), y T_r es el período de retorno y se define como el promedio esperado de tiempo entre eventos extremos específicos.

3.2 Análisis exploratorio de datos

Es necesario mirar los datos antes de ocuparse de los análisis estadísticos. En la actualidad con la facilidad que nos brindan los sistemas de computación, la exploración de los datos gráficos se ha tornado más y más importante. En la sección siguiente discutiremos algunos de los métodos gráficos más usados.

3.2.1 Gráficos de Cuantil y Probabilidad

Dado un conjunto de datos a ser analizados, usualmente su análisis empieza con un histograma, uno o más gráficos de caja, un gráfico de la función de distribución empírica; en el caso multidimensional, un gráfico de dispersión en forma de matriz. En el tema del proyecto nos restringimos, sin embargo, al caso unidimensional y con una previa discusión al problema.

Posteriormente, se encuentra una función de distribución F que sea un buen modelo para los datos iid. X_1, X_2, \dots, X_n y se define el conjunto ordenado $X_{1:n} \leq \dots \leq X_{n:n}$. La teoría básica que caracteriza los gráficos de probabilidad es la transformación cuantil del lema 2.26 el cual implica que para F continua, las v.a $U_i = F(X_i)$, para $i = 1, 2, \dots, n$ son iid. uniformes en $(0, 1)$. Sobre todo

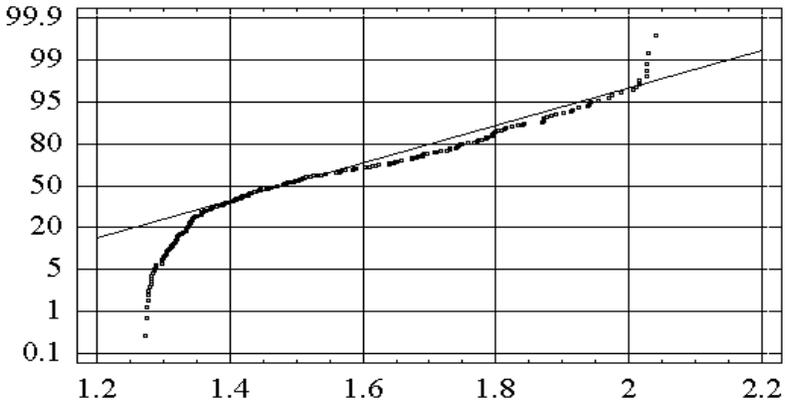
$$(F(X_{n,k}))_{k=1,\dots,n} \stackrel{d}{=} (U_{k,n})_{k=1,\dots,n},$$

de esto se sigue que

$$EF(X_{k,n}) = \frac{n-k+1}{n+1}, \quad k = 1, \dots, n.$$

También note que $F_n(X_{k,n}) = (n-k+1)/n$, donde F_n es la función de distribución empírica de F .

Gráfico No. 3.1



Este gráfico es llamado de probabilidad (PP_Plot), y se lo construye graficando

$$\left\{ \left(F(X_{k,n}), \frac{n-k+1}{n+1} \right) : k = 1, \dots, n \right\},$$

sin embargo, más común es graficar

$$\left\{ \left(X_{k,n}, F^{\leftarrow} \left(\frac{n-k+1}{n+1} \right) \right) : k = 1, \dots, n \right\} \quad (3.1)$$

que normalmente es conocido como el Gráfico Cuantil (QQ_Plot). En ambos casos la aproximación lineal del trazo es justificada por el teorema de Glivenko-Cantelli.

Existe varias variedades de (3.1) del tipo

$$\left\{ \left(X_{k,n}, F^{\leftarrow} (p_{k,n}) \right) : k = 1, \dots, n \right\} \quad (3.2)$$

donde $p_{k,n}$ es una cierta posición del trazo. Normalmente son elegidos

$$p_{k,n} = \frac{n-k+d_k}{n+g_k},$$

con (\mathbf{d}, \mathbf{g}) . En la mayoría de casos tomamos (3.1) o (3.2) con

$$p_{k,n} = \frac{n - k + 0.5}{n}.$$

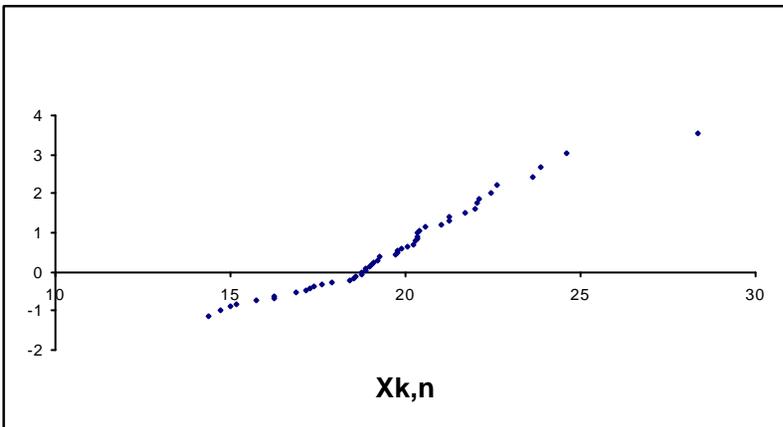
Para la distribución de Gumbel

$$\Lambda(x) = \exp\{-e^{-x}\}, \quad x \in \tilde{n},$$

el método es aplicado fácilmente y lleva a la llamada graficación doble logarítmica. Supongamos por el momento que X_1, X_2, \dots, X_n provienen de Λ . Con este fin, tomamos el conjunto ordenado y graficamos $X_{k,n}$, así $L^-(p_{k,n}) = -\ln(-\ln p_{k,n})$, donde $p_{k,n}$ es una posición en el gráfico como lo hemos discutido anteriormente. Si la distribución de Gumbel provee un buen ajuste a nuestros datos, entonces este QQ_Plot debería verse más o menos lineal.

Gráfico No. 3.2

QQ_Plot



En este gráfico se utilizó la transformación de Gumbel.

Sin embargo, los datos tienen parámetros de escala y localización donde en algunos casos m y y son la media y la desviación estándar de X . Un QQ_Plot usando estos datos debe ser aún lineal, con pendiente y e intersección m . Por ejemplo

usando regresión lineal, se puede deducir una buena estimación de los parámetros.

Hemos considerado solo familias de escala y localización. En el caso de las distribuciones de valores extremos generalizados (VEG), (definición 2.17),

$$H_{x,my} = \exp \left\{ - \left(1 + \mathbf{x} \frac{x - \mathbf{m}}{\mathbf{y}} \right)^{-1/\mathbf{x}} \right\}, \quad 1 + \mathbf{x} \frac{x - \mathbf{m}}{\mathbf{y}} \geq 0,$$

$$= \begin{cases} \Phi_a (1 + (x - \mathbf{m})/(\mathbf{a}\mathbf{x})) & \text{para } x > \mathbf{m} - \mathbf{a}\mathbf{x}, \mathbf{x} = 1/\mathbf{a} > 0 \\ \Psi_a (- (1 - (x - \mathbf{m})/(\mathbf{a}\mathbf{x}))) & \text{para } x < \mathbf{m} + \mathbf{a}\mathbf{x}, \mathbf{x} = -1/\mathbf{a} < 0 \\ \Lambda((x - \mathbf{m})/\mathbf{y}) & \text{para } x \text{ real}, \mathbf{x} = 0 \end{cases}$$

además del parámetro de localización $\mathbf{m} \in \tilde{\mathfrak{n}}$, del parámetro de escala $\mathbf{y} > 0$ interviene un parámetro de forma $\mathbf{x} \in \tilde{\mathfrak{n}}$, haciendo que la interpretación del QQ_Plot sea inmediata y más delicada. Un método preferido para pruebas gráficas si las muestras provienen de $H_{\xi_{\mu,\mathbf{y}}}$ debería ser obteniendo primero un estimador $\hat{\mathbf{x}}$ de ξ por uno de los métodos que expondremos a continuación, y consecuentemente se realiza un QQ_Plot usando $H_{x,0,1}$ donde \mathbf{m} y \mathbf{y} pueden ser estimadas o por inspección simple o por medio de regresión lineal. Estas estimaciones preliminares son frecuentemente usadas como valores iniciales en los procedimientos de iteración numérica.

3.2.2 La Función de exceso media

Otra herramienta gráfica útil, en particular para discriminación en las colas, es la función de exceso media. Ya habíamos introducido esta definición en el contexto de VEG, definición 2.19.

La cantidad $e(u)$ es frecuentemente referida a cómo el exceso medio pasa el valor del umbral u . Esta interpretación será crucial en la sección 3.5, en un contexto de seguros. $e(u)$ puede ser interpretada como el tamaño de reclamo esperado en un intervalo ilimitado, sobre la prioridad u . De aquí $e(u)$ es también llamada función de pérdida de exceso media. En un contexto médico y de seguridad, $e(u)$ es llamada función de vida residual de exceso medio. En el contexto de administración de riesgo financiero, cambiando la cola derecha a la cola izquierda $e(u)$ es conocida como déficit. Un resumen de la función de exceso media más importante podemos encontrar

en la tabla 2.20.

Ejemplo 3.1 Algunas propiedades elementales de la función de exceso media

Si X está distribuida exponencialmente con tasa λ , entonces $e(u) = \lambda/\lambda$, para todo $u > 0$. Ahora supongamos que X es una v.a no acotada a la derecha de la f.d. F . Si para todo $y \in \mathbb{R}^+$,

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x-u)}{\bar{F}(x)} = e^{-\gamma u}, \quad (3.3)$$

para algún $\gamma \in [0, \infty]$, entonces $\lim_{x \rightarrow \infty} e(x) = \gamma^{-1}$ ♣

Una prueba gráfica para el comportamiento de la cola puede ahora ser basada en la función de exceso media empírica $e_n(u)$. Supongamos que X_1, X_2, \dots, X_n son iid. con f.d. F y denotamos F_n la f.d. empírica de F y $\mathcal{D}_n(u) = \{i : i = 1, 2, \dots, n, X_i > u\}$, entonces

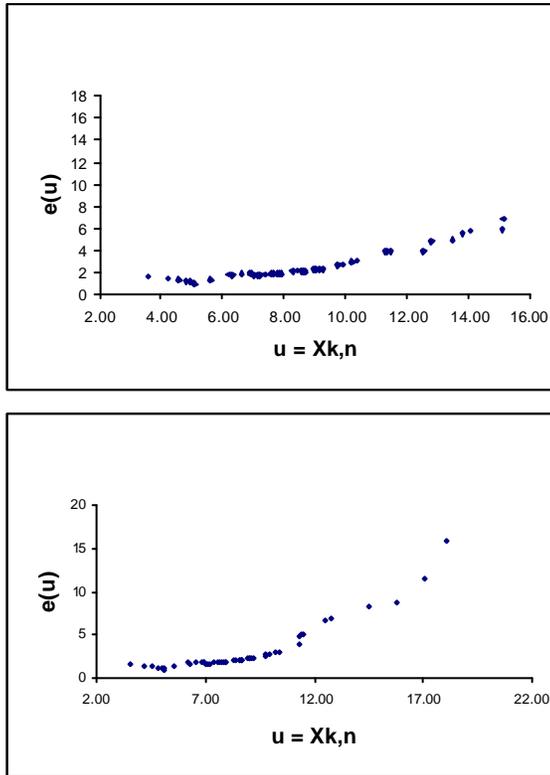
$$e_n(u) = \frac{1}{\bar{F}_n(u)} \int_u^\infty \bar{F}_n(y) dy = \frac{1}{\text{card} \mathcal{D}_n(u)} \sum_{i \in \mathcal{D}_n(u)} (X_i - u), \quad u \geq 0, \quad (3.4)$$

con la convención de que $0/0 = 0$. Un gráfico de exceso medio (ME_Plot) consiste en

$$\{(X_{k,n}, e_n(X_{k,n})) : k = 1, 2, \dots, n\}.$$

Gráfico No. 3.3

ME_Plot



En estos gráficos se puede dar cuenta que la cola de la distribución de los datos pertenecientes al ME_Plot de la derecha es más pesada que la cola de distribución de los datos pertenecientes al otro gráfico (observar las distintas inclinaciones).

Las propiedades estadísticas de $e_n(u)$ pueden ser nuevamente derivadas por uso de teoría de procesos empíricos relevantes. Para nuestro propósito el ME_Plot es usado solo como un método gráfico, especialmente para distinguir modelos de colas livianas y colas pesadas. Una precaución que es importante considerar es que los gráficos deben ser interpretados cuidadosamente, ya que debido a la dispersión de los datos disponibles para el cálculo de $e_n(u)$ para valores grandes de u , los gráficos resultantes son muy sensitivos a cambios en los datos hacia el fin del intervalo.

3.2.3 Método de Gumbel de excedentes

Hay una multitud de resultados analíticos completamente fáciles concerniente a

eventos extremos que producen información previa útil en los datos. El primer método es el de Gumbel de excedentes, que trata de analizar el número de observaciones que en el futuro excederán los récord pasados (sección 3.2.5).

Sea $X_{n,n} < \dots < X_{1,n}$ las estadísticas de orden usuales de la muestra X_1, X_2, \dots, X_n fijos en una sucesión iid. infinita (X_i) con f.d. continua F . Tomamos la k -ésima estadística de orden superior $X_{k,n}$ como un valor inicial y denotamos $S_r^n(k)$, $r \geq 1$, entonces el número de excedentes de $X_{k,n}$ en las próximas r observaciones X_{n+1}, \dots, X_{n+r} , es

$$S_r^n(k) = \sum_{i=1}^r I_{\{X_{n+i} > X_{k,n}\}}.$$

Posteriormente por facilidad de notación escribimos S por $S_r^n(k)$.

Lema 3.2 Estadísticas de orden y la f.d. hipergeométrica

La v.a S definida como antes tiene una distribución hipergeométrica, e.d.

$$P(S = j) = \frac{\binom{r+n-k-j}{n-k} \binom{j+k-1}{k-1}}{\binom{r+n}{n}}, \quad j = 0, 1, \dots, r. \quad (3.5)$$

Demostración. Condicionando, tenemos

$$P(S = j) = \int P(S = j \mid X_{k,n} = u) dF_{k,n}(u),$$

donde $F_{k,n}$ denota la f.d. de $X_{k,n}$. Ahora usando el hecho que (X_1, X_2, \dots, X_n) y $(X_{n+1}, \dots, X_{n+r})$ son independientes, que $\sum_{i=1}^r I_{\{X_i > u\}}$ tiene una distribución binomial con parámetro r y $\bar{F}(u)$, y de la proposición 2.24 que

$$dF_{k,n}(u) = \frac{n!}{(k-1)!(n-k)!} F^{n-k}(u) \bar{F}^{k-1}(u) dF(u)$$

se obtiene (3.5) ♣

Nota: Se obtiene inmediatamente de la definición de S y de los argumentos dados en la demostración anterior que $ES = rk/(n+1)$ para un número importante de excedentes del umbral aleatorio $X_{k,n}$.

Ejemplo 3.3

Supongamos $n = 270, r = 12$. Queremos calcular:

- a) la probabilidad $p_k = P(S_{12}^{270}(k) = 0)$ que no existan excedentes de $X_{k,270}, k \geq 1$, en las próximas doce observaciones.
- b) la probabilidad $p_k = P(S_{12}^{270}(k) = i)$ que existan i excedentes, $i = 1, 2, \dots, 5$ de $X_{k,270}, k \geq 1$, en las próximas doce observaciones.

Solución

- a) Para $j = 0$ la fórmula (3.5) se reduce a

$$P(S_r^n(k) = 0) = \frac{n(n-1)\dots(n-k+1)}{(r+n)(r+n-1)\dots(r+n-k+1)}.$$

en forma tabulada obtenemos para $n = 270$ y $r = 12$,

$j \setminus k$	1	2	3	4	5	6
0	0.95744681	0.9165594	0.87727828	0.83954588	0.80330649	0.76850621

Se observa que la probabilidad de no tener excedentes del valor más grande, en las próximas doce observaciones es de 0.957447.

Así, si tenemos 270 datos de puntos mensuales encaminados a designar un cierto estándar igual a la tercera observación más grande, hay cerca del 88% de posibilidad de que ese nivel no sea excedido durante el próximo año.

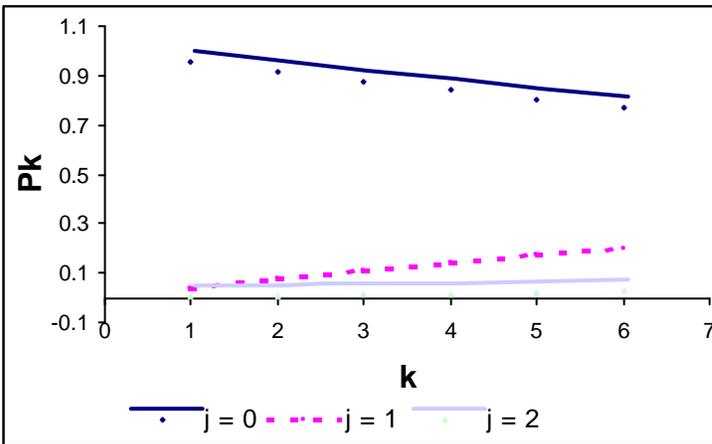
- b) Para $j = 1, \dots, 5$ usando la fórmula (3.5) se tiene

$j \setminus k$	1	2	3	4	5	6
-----------------	---	---	---	---	---	---

	0.0408874	0.0785622		0.1449575	0.1740014	
1	1	3	0.1131972	6	1	0.20047988
	0.0016062	0.0046461	0.0089580	0.0143910	0.0208045	
2	9	5	5	9	2	0.02806718
	5.7573E-	0.0002228	0.0005389	0.0010428	0.0017652	
3	05	4	9	3	3	0.0027316
	1.8639E-	9.0502E-	2.6364E-	5.9726E-	0.0001159	
4	06	06	05	05	6	0.00020262
		3.1479E-	1.0737E-	2.7901E-	6.1168E-	
5	5.383E-08	07	06	06	06	1.1919E-05

Se mira que la probabilidad de tener un excedente de la tercera observación más grande en las siguientes 12 observaciones es de 0.1131972.

Gráfico No. 3.4



Este gráfico indica las probabilidades de que no existan excedentes, que exista un excedente y que existan dos excedentes en las siguientes doce observaciones para distintos valores de k. ♣

3.2.4 El período de retorno

En esta sección estamos interesados en analizar inquietudes como: ¿Cuál es el tiempo medio esperado entre dos eventos extremos específicos?

Esta inquietud está normalmente enfocada en el siguiente sentido. Sea (X_i) una sucesión de v.as con f.d. continua F y u un umbral dado. Consideramos la sucesión $(I_{\{X_i > u\}})$ de v.a iid. de Bernoulli con probabilidad $p = \bar{F}(u)$. Consecuentemente, el tiempo de la primera ocurrencia es

$$L(u) = \min\{i \geq 1 : X_i > u\},$$

e.d. el tiempo del primer excedente del umbral u , es una v.a con f.d. geométrica.

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

Notemos que las v.a iid.

$$L_j(u) = L(u), \quad L_{n+j}(u) = \min\{i > L_n(u) : X_i > u\}, \quad n \geq 1,$$

(X_n) describe los períodos de tiempo entre los excedentes sucesivos de u . El período de retorno del evento $\{X_i > u\}$ es entonces definido como $EL(u) = p^{-1} = (\bar{F}(u))^{-1}$, que tiende al ∞ cuando $u \rightarrow \infty$.

Todas las inquietudes de períodos de retorno pueden ahora ser aclaradas simplemente con las correspondientes propiedades de la distribución geométrica.

Definamos

$$r_k = P(L(u) \leq k) = p \sum_{i=1}^k (1-p)^{i-1} = 1 - (1-p)^k, \quad k \in \mathbb{N}.$$

Aquí r_k es la probabilidad de que exista al menos un excedente de u antes del tiempo k (o en las k observaciones). Esto da una relación 1 - 1 entre r_k y el período de retorno p^{-1} .

La probabilidad que haya un excedente de u antes del período de retorno es

$$P(L(u) \leq EL(u)) = P(L(u) \leq [1/p]) = 1 - (1-p)^{[1/p]},$$

Donde $[x]$ es la parte entera de x . Para un umbral alto e.d. cuando $u \rightarrow \infty$ y consecuentemente $p \rightarrow 0$, obtenemos

$$\begin{aligned} \lim_{u \uparrow \infty} P(L(u) \leq EL(u)) &= \lim_{p \downarrow 0} (1 - (1 - p)^{\lfloor 1/p \rfloor}) \\ &= 1 - e^{-1} = 0.63212. \end{aligned}$$

Esto demuestra que para un umbral alto la media de $L(u)$ es más grande que su mediana.

Ejemplo 3.4 Período de retorno, evento de t años

En un contexto de seguros, una estructura es asegurada en base de que al menos en los siguientes 30 años no exista más del 7.5% de riesgo de falla.

Qué implica esta información para el período de retorno? Usando el lenguaje anterior, la ingeniería traduce esto a

$$P(L(u) \leq 30) \leq 0.075$$

Aquí asumimos que una falla de estructura por cada año i puede ser modelada por medio del evento $\{X_i > u\}$, donde X_i es una componente crítica de estructura dependiente. Asumimos la propiedad de iid. para las X_i . De la condición anterior se tiene que $P(L(u) \leq 30) = 1 - (1 - p)^{30} = 0.075$, lo que implica inmediatamente que $p = 0.002595344$, e.d $EL(u) = 385$. En lenguaje de seguros se habla en este caso alrededor de un evento en 385 años.

La importancia de la próxima pregunta concierne la implicación del requisito de un evento en t años para un valor de umbral establecido. Por definición significa que para el umbral correspondiente u_t ,

$$t = EL(u_t) = \frac{1}{\bar{F}(u_t)},$$

de aquí

$$u_t = F^{\leftarrow}(1 - t^{-1}).$$

En este ejemplo $u_{385} = F^{\leftarrow}(0.9974)$. Esto nos lleva una vez más al problema de estimación del cuantil alto. ♣

Además nos lleva a concentrarnos en lo que se necesita para una solución de los siguientes problemas:

- Encontrar una estimación confiable para el cuantil alto de datos iid.
- Como los datos en la práctica se presentarán dependientes o no estacionarios, encontrar un procedimiento de estimación cuantil para este tipo de datos.

3.2.5 Record como una herramienta exploratoria

Supongamos que las v.a X_i son iid. con f.d. F . Un récord X_n ocurre si $X_n > M_{n-1} = \max(X_1, X_2, \dots, X_{n-1})$. Por definición tomamos X_1 como récord. En las definiciones anteriores usamos lenguaje de procesos puntuales con el fin de describir récord y tiempos récord L_n . Los últimos son los tiempos aleatorios en los cuales el proceso (M_n) salta. Se define el proceso de conteo de récord como

$$N_1 = 1, \quad N_n = 1 + \sum_{k=2}^n I_{\{X_k > M_{k-1}\}}, \quad n \geq 2.$$

Lema 3.5 Momentos de N_n

Supongamos (X_i) son iid. con f.d. F continua y (N_n) definida como antes. Entonces

$$EN_n = \sum_{k=1}^n \frac{1}{k} \quad \text{y} \quad \text{var}(N_n) = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k^2} \right)$$

Demostración

De la definición de N_n obtenemos

$$\begin{aligned}
 EN_n &= 1 + \sum_{k=2}^n P(X_k > M_{k-1}) \\
 &= 1 + \sum_{k=2}^n \int_{-\infty}^{\infty} P(X_k > u) dP(M_{k-1} \leq u).
 \end{aligned}$$

ahora usamos $P(M_{k-1} \leq u) = F^{k-1}(u)$. De donde,

$$\begin{aligned}
 EN_n &= 1 + \sum_{k=2}^n \int_{-\infty}^{\infty} P(X_k > u) dF^{k-1}(u) \\
 &= 1 + \sum_{k=2}^n (k-1) \int_0^1 P(X_k > u) F^{k-2}(u) dF(u) \\
 &= 1 + \sum_{k=2}^n (k-1) \int_0^1 (1-F(u)) F^{k-2}(u) dF(u).
 \end{aligned}$$

Haciendo el cambio de variable se tiene que,

$$EN_n = 1 + \sum_{k=2}^n (k-1) \int_0^1 (1-v) v^{k-2} dv$$

de donde

$$\begin{aligned}
 EN_n &= 1 + \sum_{k=2}^n (k-1) \left(\frac{1}{k-1} - \frac{1}{k} \right) \\
 &= 1 + \sum_{k=2}^n \frac{1}{k},
 \end{aligned}$$

y por lo tanto $EN_n = \sum_{k=1}^n \frac{1}{k}$.

De la misma manera se puede demostrar la varianza de N_n . ♣

Notemos que la EV_n y la $\text{var}(N_n)$ son ambos de orden $\ln n$ cuando $n \rightarrow \infty$. Más precisamente, $EN_n - \ln n \rightarrow g$ donde $g = 0.5772\dots$ que es la constante de Euler.

Como una consecuencia: el número de récord de datos iid. crece muy lentamente.

A continuación presentamos una tabla de algunos resultados de récord que tendríamos que esperar para distintas n observaciones.

Tabla No. 3.6

$n = 10^k, k =$	EN_n	$\ln n$	$\ln n + g$	$\sqrt{\text{var}(N_n)}$
1	2.9	2.3	2.9	1.2
2	5.2	4.6	5.2	1.9
3	7.5	7.0	7.5	2.4
4	9.8	9.2	9.8	2.8
5	12.1	11.5	12.1	3.2
6	14.4	13.8	14.4	3.6
7	16.7	16.1	16.7	3.9
8	19.0	18.4	19.0	4.2
9	21.3	20.7	21.3	4.4

En esta tabla se muestra el número esperado de récord EN_n en una sucesión (X_n) iid., junto con la aproximación asintótica $\ln n$, $\ln n + \gamma$, y la desviación estándar $\sqrt{\text{var}(N_n)}$, basados en el lema anterior.

3.3 Estimación de parámetros para la distribución de valores extremos generalizada

Retomamos la definición de distribución de valores extremos generalizada

$$H_{x,my} = \exp\left\{-\left(1 + x \frac{x-m}{y}\right)^{-1/x}\right\}, \quad 1 + x \frac{x-m}{y} \geq 0. \quad (3.6)$$

En el caso de que $\xi = 0$ corresponde a la distribución de Gumbel

$$H_{x,my} = \exp\left\{-e^{-(x-m)/y}\right\} \quad x \in \bar{n} \quad (3.7)$$

El parámetro $\mathbf{q} = (\mathbf{x} \ \mathbf{m} \ \mathbf{y}) \in \tilde{\mathbf{n}} \times \tilde{\mathbf{n}} \times \tilde{\mathbf{n}}_+$ consiste en un parámetro de forma \mathbf{x} un parámetro de localización \mathbf{m} y un parámetro de escala \mathbf{y} . La metodología de la estadística estándar de la teoría de estimación de parámetros es válida si nuestros datos consisten de una muestra

$$X_1, X_2, \dots, X_n \text{ iid. de } H_{\mathbf{q}} \quad (3.8)$$

Mencionamos aquí la suposición de que la v.a X_i tiene una distribución de valor extremo exacta $H_{\mathbf{q}}$ lo cual no es tal vez lo más realista. En la próxima sección retornamos a la suposición más robusta de que los X_i tienen aproximadamente una distribución $H_{\mathbf{q}}$. La “próximalidad” será interpretada como “perteneciente al MDA”.

Para la fijación del máximo anual:

Los datos son válidos cuando los X_i pueden ser interpretados como máximo en períodos disjuntos de tiempo de longitud s , estos períodos, muchas veces, son de un año, este período de tiempo es elegido para intentar compensar la estacionalidad intra año. En lo que los datos originales pueden mirarse como

$$\begin{aligned} X^{(1)} &= (X_1^{(1)}, \dots, X_s^{(1)}) \\ X^{(2)} &= (X_1^{(2)}, \dots, X_s^{(2)}) \\ &\vdots \\ X^{(n)} &= (X_1^{(n)}, \dots, X_s^{(n)}). \end{aligned}$$

Donde los vectores $(X^{(i)})$ se supone son iid. Pero en cada vector $X^{(i)}$ los componentes pueden ser dependientes. La longitud del tiempo s es elegida de tal manera que las condiciones de arriba son probablemente satisfechas. La muestra básica iid. de $H_{\mathbf{q}}$ en la cual la inferencia estadística está siendo ejecutada entonces consiste de

$$X_i = \max\{X_1^{(i)}, \dots, X_s^{(i)}\}, \quad i = 1, \dots, n. \quad (3.9)$$

Referente a estudios realizados y a que frecuentemente s es un año, la inferencia estadística para $H_{\mathbf{q}}$ basada en los datos de la forma (3.9) se refiere a cómo ajustar el máximo anual.

Luego expondremos algunas de las principales técnicas para la estimación de \mathbf{q}

en modelos como los de (3.8).

3.3.1 Estimación de máxima verosimilitud

El modelo (3.8) corresponde al caso paramétrico estándar de inferencia estadística y aquí, en principio, puede ser resuelto por el método de máxima verosimilitud. Supongamos que H_q tiene densidad h_q . Entonces la función de verosimilitud basada en el dato $\mathbf{X} = (X_1, X_2, \dots, X_n)$ está dada por

$$L(\mathbf{q} ; \mathbf{X}) = \prod_{i=1}^n h_q(X_i) I_{\{1+x(X_i-m)/y > 0\}}.$$

Denotamos por $l(\mathbf{q} ; \mathbf{X}) = \ln L(\mathbf{q} ; \mathbf{X})$ la función del log-verosimilitud. El estimador de máxima verosimilitud (EMV) de \mathbf{q} entonces es $\hat{\mathbf{q}}_n = \hat{\mathbf{q}}_n(X_1, \dots, X_n)$ que maximiza $l(\mathbf{q} ; \mathbf{X})$ sobre un espacio de parámetro apropiado \mathbf{q} . En el caso de $H_{0, \mathbf{m}y}$ nos da

$$l((0, \mathbf{m}y); \mathbf{X}) = -n \ln y - \sum_{i=1}^n \exp\left\{-\frac{X_i - \mathbf{m}}{y}\right\} - \sum_{i=1}^n \frac{X_i - n}{y}.$$

Diferenciando esta última función respecto a $\mathbf{m}y$ a y nos da el siguiente sistema en el caso de Gumbel.

$$\begin{aligned} 0 &= n - \sum_{i=1}^n \exp\left\{-\frac{X_i - \mathbf{m}}{y}\right\}, \\ 0 &= n + \sum_{i=1}^n \frac{X_i - \mathbf{m}}{y} \left(\exp\left\{-\frac{X_i - \mathbf{m}}{y}\right\} - 1 \right) \end{aligned}$$

No existe solución explícita en esta ecuación. La situación para H_x cuando $x \neq 0$ es aún más complicada. El cálculo numérico del (EMV) $\hat{\mathbf{q}}_n$ para H_q no tiene problemas serios en principio, decimos en principio porque en el llamado *caso regular* la estimación de máxima verosimilitud ofrece una técnica que produce estimadores eficientes, consistentes y estimadores asintóticamente normales.

Por otro lado, nuestros datos han consistido de n observaciones iid. de máximos

los cuales hemos supuesto siguen exactamente una VEG H_q . Con una definición apropiada los períodos de tiempo básicos, los asignamos independientes en el modelo. Supongamos ahora que, solo tengamos disponible las k observaciones más grandes de cada período de tiempo. En la notación de (3.9) esto nos lleva a los datos

$$X_{k,s}^{(i)} \leq \dots \leq X_{1,s}^{(i)} = X_i, \quad i = 1, \dots, n.$$

La teoría de máxima verosimilitud basada en estas n por k observaciones debe usar la densidad conjunta de los vectores independientes $(X_{k,s}^{(i)}, \dots, X_{1,s}^{(i)}), i = 1, \dots, n$. Solo pocas veces en casos prácticos podríamos suponer que para cada i , los vectores se derivan de datos iid. Si ese fue el caso entonces el estimador de máxima verosimilitud debe basarse solo en la densidad conjunta de k estadísticas de orden superior de un VEG como lo discutido en el teorema 2.26.

$$\frac{s!}{(s-k)!} H_q^{s-k}(x_k) \prod_{i=1}^n h_q(x_i), \quad x_k < \dots < x_1,$$

donde, dependiendo de q , el valor de x satisface las restricciones de dominios importantes. Los errores estándar de EMV para m y y pueden ser reducidos considerablemente si $k = 2$ e.d. tomamos las dos observaciones más grandes.

Una afirmación final concerniente a la metodología de máxima verosimilitud se resume de la siguiente manera:

La gran ventaja de los procedimientos de máxima verosimilitud es que ellos pueden ser generalizados, con muy pocos cambios en la metodología básica, a modelos mucho más complicados en los cuales otros efectos pueden estar presentes. Además los estimadores que se obtienen eficientes, consistentes y mantienen la condición de normalidad asintótica, por lo tanto la construcción de los intervalos de confianza para estos estimadores se la realiza de la forma clásica.

3.3.2 Estimación de Colas y Cuantil, un primer paso

Regresamos a la estructura básica de (3.8) y (3.9) e.d. que tenemos una muestra iid. X_1, \dots, X_n de H_θ . En esta situación un estimador cuantil puede ser previamente obtenido. En efecto, por los métodos discutidos en las secciones previas, obtenemos una estimación \hat{q} de q . Dado algún $p \in (0, 1)$, el p -cuantil x_p es definido por medio de

$x_p = H_q^{\leftarrow}(p)$ (Definición 2.10). Un estimador natural para x_p , basado en X_1, \dots, X_n , viene dado por

$$\hat{x}_p = H_{\hat{q}}^{\leftarrow}(p)$$

por la definición de H_{θ} esto conduce a

$$\hat{x}_p = \hat{m} - \frac{\hat{y}}{\hat{x}} \left(1 - (-\ln p)^{-\hat{x}} \right)$$

El correspondiente estimador de cola para $\overline{H}_q(x)$, para x en un dominio apropiado, corresponde a

$$\hat{H}_{\hat{q}}(x) = 1 - \exp \left\{ - \left(1 + \hat{x} \frac{x - \hat{m}}{\hat{y}} \right)^{-1/\hat{x}} \right\},$$

donde $\hat{q} = (\hat{x}, \hat{m}, \hat{y})$ es estimado por EMV o por una estimación de momentos ponderados de probabilidad.

Nota: Si la distribución de valores extremos es conocida e.d. que sea de Fréchet, Weibull o Gumbel, los métodos anteriores pueden ser adaptados a una f.d. específica bajo ciertas consideraciones. Esto podría simplificar el problema de estimación si $\xi \geq 0$, es decir, en el caso de la distribución de Fréchet o Gumbel, pero no en el caso de la distribución de Weibull.

3.4 Estimación bajo condiciones de máximo dominio de atracción

3.4.1 Introducción

Relajando la condición (3.8), suponemos en esta sección que para $\xi \in \mathfrak{N}$,

$$X_1, \dots, X_n \text{ pertenecen a } F \in \text{MDA}(H_x) \text{ y son iid.} \quad (3.10)$$

Por la proposición 2.8, $F \in \text{MDA}(H_x)$ es equivalente a

$$\lim_{n \rightarrow \infty} n\bar{F}(c_n x + d_n) = -\ln H_{\mathbf{x}}(x) \quad (3.11)$$

para sucesiones normalizadas apropiadas (c_n) y (d_n) , y x perteneciente a un dominio apropiado dependiendo del signo de \mathbf{x}

Existen diferencias fundamentales entre (3.8) y (3.10). Consideremos por propósito ilustrativo solo el caso de Fréchet que es de nuestro interés con $\mathbf{x} = 1/\alpha > 0$. Ahora (3.8) indica que nuestra muestra X_1, \dots, X_n sigue exactamente la distribución de Fréchet, e.d

$$\bar{F}(x) = 1 - \exp\{-x^{-\alpha}\}, \quad x > 0.$$

Por otra parte, por el teorema 2.11 la suposición (3.10) se reduce en el caso de Fréchet a

$$\bar{F}(x) = x^{-\alpha} L(x), \quad x > 0,$$

para algunas funciones suavemente variante L . Claramente, en este caso la estimación de la cola $\bar{F}(x)$ es mucho más compleja debido al carácter no paramétrico de L . Además (3.8) tiene muchas suposiciones paramétricas, y por otro lado (3.10) es esencialmente semiparamétrica: es decir que hay una parte paramétrica α y una parte no paramétrica L .

Por estas diferencias, (3.10) es considerada más generalmente como inferencia para distribuciones de colas pesadas como en forma opuesta a la inferencia para la VEG en (3.8).

Una consecuencia de (3.11) es que para u grande $u = c_n x + d_n$,

$$n\bar{F}(u) \approx \left(1 + \mathbf{x} \frac{u - d_n}{c_n} \right)^{-1/\mathbf{x}},$$

así que un estimador de colas podría tomar la forma

$$(\bar{F}(u))^\wedge = \frac{1}{n} \left(1 + \hat{\mathbf{x}} \frac{u - \hat{d}_n}{c_n} \right)^{-1/\hat{\mathbf{x}}}, \quad (3.12)$$

para estimadores apropiados $\hat{\mathbf{x}}$, \hat{c}_n y \hat{d}_n . Como (3.10) es esencialmente una propiedad de cola, la estimación de ξ puede ser basada en k estadísticas de orden superior $X_{k,n} \leq \dots \leq X_{1,n}$.

Las siguientes condiciones matemáticas son usualmente impuestas:

- $k(n) \rightarrow \infty$ usa un número suficientemente grande de estadísticas de orden,

pero

- $\frac{n}{k(n)} \rightarrow \infty$ como estamos interesados en una propiedad de cola, debemos concentrarnos solo en estadísticas de orden superior. (3.13)

De (3.12) deberíamos en principio estar en la posición de estimar el cuantil $x_p = F^{\leftarrow}(p)$, para un p fijo en $(0, 1)$, como sigue

$$x_p = \hat{d}_n + \frac{\hat{c}_n}{\hat{\mathbf{x}}} \left((n(1-p))^{-\hat{\mathbf{x}}} - 1 \right) \quad (3.14)$$

Típicamente, estaremos interesados en estimar cuantiles altos fuera de la muestra X_1, \dots, X_n . Esto significa que $p = p_n$ es escogido de tal forma que $p > 1 - 1/n$, de aquí la f.d. empírica satisface $\bar{F}_n(p) = 0$ y no produce ninguna información acerca de tales cuantiles. Para obtener un buen estimador de ξ , c_n y d_n en (3.14) se necesita un artificio en la sucesión. Suponemos por conveniencia la notación que $n/k \in \mathbb{N}$. Una aproximación estándar ahora consiste en pasar a una subsucesión (n/k) con $k = k(n)$ satisfaciendo (3.13). El cuantil x_p entonces es estimado por

$$\hat{x}_p = \hat{d}_{n/k} + \frac{\hat{c}_{n/k}}{\hat{\mathbf{x}}} \left(\left(\frac{n}{k} (1 - p_n) \right)^{-\hat{\mathbf{x}}} - 1 \right) \quad (3.15)$$

La razón tras esta construcción es que necesitamos estimar en dos niveles. Primero tenemos que encontrar una estimación confiable para ξ . Segundo,

necesitamos estimar constantes normalizadas c_n y d_n las cuales ellas mismas son definidas por medio de cuantiles de F .

En el caso de Fréchet sabemos que $c_n = F^{\leftarrow}(1 - n^{-1})$ (teorema 2.11). De aquí estimar c_n es equivalente al problema de estimar x_p en nuestro rango de datos. Nos movemos siempre del valor crítico $1 - n^{-1}$ al más seguro $1 - (n/k)^{-1}$ acercándonos a lado de la subsucesión (n/k) . La estimación de $c_{n/k}$ nos lleva a estimar el cuantil en el rango de nuestros datos. Argumentos semejantes se emplean para $d_{n/k}$, y realmente para el caso de la distribución de Gumbel y Weibull.

Podemos por lo tanto esperar que la construcción de (3.15) conduzca a un buen estimador para x_p . La discusión anterior es solo una heurística, un análisis estadístico detallado muestra que esta aproximación puede ser utilizada.

En el contexto de estadísticos de eventos extremos esto también puede ser de interés para estimar la siguiente cantidad la cual está cercanamente relacionada con el cuantil x_p :

$$x_{p,r} = F^{\leftarrow}(p^{1/r}), \quad r \in \mathbb{N}.$$

Notemos que $x_p = x_{p,1}$. La interpretación de $x_{p,r}$ es obvia de

$$p = F^{\leftarrow}(x_{p,r}) = P(\max(X_{n+1}, \dots, X_{n+r}) \leq x_{p,r}),$$

así $x_{p,r}$ es el nivel el cual, con una probabilidad dada p , no será excedido en las próximas r observaciones X_{n+1}, \dots, X_{n+r} . Como un estimador entonces obtenemos de (3.15)

$$\hat{x}_{p,r} = \hat{d}_{n/k} + \frac{\hat{c}_{n/k}}{\hat{x}} \left(\left(\frac{n}{k} (1 - p^{1/r}) \right)^{-\hat{x}} - 1 \right)$$

De la heurística anterior obtenemos una secuencia a seguir para el resto de esta sección.

- Encontrar estimadores apropiados para el parámetro de forma ξ de VEG.
- Encontrar estimadores apropiados para las constantes normalizadas c_n y d_n .
- Demostrar que los estimadores propuestos arriba dan como resultado una aproximación razonable a la distribución de cola.

d) Determinar las propiedades estadísticas de esos estimadores.

3.4.2 Estimación del parámetro de forma x

En esta sección estudiamos estimadores diferentes para los parámetros de forma ξ pertenecientes a $F \in \text{MDA}(H_\xi)$. También damos algunas de sus propiedades estadísticas.

3.4.2.1 Estimador de Pickands para $x \hat{I} \tilde{n}$

La idea básica tras este estimador consiste en encontrar una condición equivalente a $F \in \text{MDA}(H_\xi)$ el cual encierra el parámetro ξ en un camino fácil. La clave del estimador de Pickands y sus generalizaciones están en el teorema 2.18, donde se muestra que para $F \in \text{MDA}(H_\xi)$, $U(t) = F^{\leftarrow}(1-t^{-1})$ satisface

$$\lim_{t \rightarrow \infty} \frac{U(2t) - U(t)}{U(t) - U(t/2)} = 2^x.$$

Además, la siguiente propiedad de uniformidad se cumple: siempre que $\lim_{t \rightarrow \infty} c(t) = 2$ para una función positiva c ,

$$\lim_{t \rightarrow \infty} \frac{U(c(t)t) - U(t)}{U(t) - U(t/c(t))} = 2^x. \quad (3.16)$$

La idea básica ahora consiste en construir un estimador empírico usando (3.16). Para ese efecto, sea

$$V_{n,n} \leq \dots \leq V_{1,n}$$

la estadística de orden de la muestra iid. V_1, \dots, V_n con f.d. común de Pareto $F_V(x) = 1 - x^{-1}$, $x \geq 1$. Esto se da en la misma forma que como para la transformación cuantil (lema 2.26), que

$$\left(X_{k,n} \right)_{k=1,\dots,n} \stackrel{d}{=} \left(U(V_{k,n}) \right)_{k=1,\dots,n},$$

donde X_1, \dots, X_n son iid. con f.d. F . Nótese que $V_{k,n}$ es el $(1-k/n)$ -cuantil empírico de F_V . Haciendo uso de la transformación cuantil, no es difícil ver que

$$\frac{k}{n} V_{k,n}^p \rightarrow 1, \quad n \rightarrow \infty,$$

donde $k = k(n) \rightarrow \infty$ y $k/n \rightarrow 0$. En particular

$$V_{k,n}^p \rightarrow \infty \quad \text{y} \quad \frac{V_{2k,n}^p}{V_{k,n}^p} \rightarrow \frac{1}{2}, \quad n \rightarrow \infty.$$

Combinando esto con (3.16), y usando el argumento de la subsucesión, se obtiene

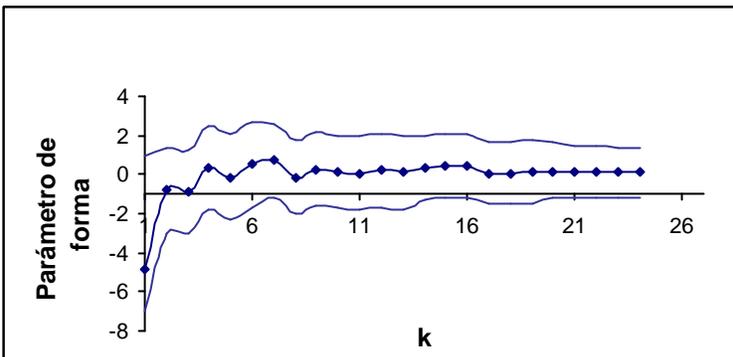
$$\frac{U(V_{k,n}) - U(V_{2k,n})}{U(V_{2k,n}) - U(V_{4k,n})} \rightarrow 2^x, \quad n \rightarrow \infty.$$

De esta manera definimos ahora el estimador de Pickands

$$\hat{\mathbf{x}}_{k,n}^{(p)} = \frac{1}{\ln 2} \ln \frac{X_{k,n} - X_{2k,n}}{X_{2k,n} - X_{4k,n}} \quad (3.17)$$

Gráfico No. 3.5

Pickands_Plot



Se presenta el Pickands_Plot con una banda asintótica del 95%. Aquí se

observa claramente que existe un equilibrio entre la varianza y el sesgo a partir de $k = 17$.

Teorema 3.7 Propiedades del estimador de Pickands

Supongamos que (X_n) es una sucesión iid. con f.d. $F \in \text{MDA}(H_\xi)$, $\xi \in \bar{\mathbb{N}}$. Sea $\hat{\mathbf{X}}^{(p)} = \hat{\mathbf{X}}_{k,n}^{(p)}$ el estimador de Pickands de (3.17).

- a) (Consistencia débil) Si $k \rightarrow \infty, k/n \rightarrow 0$ para $n \rightarrow \infty$, entonces

$$\hat{\mathbf{X}}^{(p)} \xrightarrow{P} \mathbf{X}, \quad n \rightarrow \infty$$

- b) (Consistencia fuerte) Si $k/n \rightarrow 0, k/\ln \ln n \rightarrow \infty$, para $n \rightarrow \infty$, entonces

$$\hat{\mathbf{X}}^{(p)} \xrightarrow{c.s.} \mathbf{X}, \quad n \rightarrow \infty \clubsuit$$

3.4.2.2 Estimador de Hill Para $\mathbf{x} = \mathbf{a}^{-l} > 0$

Para muchas aplicaciones el conocimiento del índice α de variación regular es de mayor importancia; así, si $\alpha < 2$ entonces $EX_1^2 = \infty$. Este caso es frecuentemente observado en la modelización de datos de seguros.

Por otro lado, estudios empíricos en las colas de los log-retornos diarios en finanzas han indicado que frecuentemente se encuentran valores de α entre 3 y 4. Esto implica que mientras que las covarianzas de los datos están bien definidas, la construcción de intervalos de confianza para el conjunto de autocovarianzas y auto correlaciones en base a la teoría asintótica puede ser cuestionable debido a que se pide la condición de finitud del momento de orden cuatro.

El estimador de Hill de α toma la siguiente forma:

$$\hat{\mathbf{a}}^{(H)} = \hat{\mathbf{a}}_{k,n}^{(H)} = \left(\frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1},$$

donde $k = k(n) \rightarrow \infty$ en forma apropiada, como en el caso del estimador de

Pickands, una sucesión creciente de estadísticas de orden superior es usada. Uno de los hechos interesantes concernientes a esta última ecuación es que se pueden derivar varias versiones asintóticamente equivalentes mediante distintos métodos. Más adelante discutiremos tales versiones.

Respecto al acercamiento al EMV, asumimos por el momento que X es una v.a con f.d. F tal que para $\alpha > 0$.

$$P(X > x) = \bar{F}(x) = x^{-\alpha}, \quad x \geq 1$$

Luego, esto inmediatamente muestra que $Y = \ln X$ tiene como f.d.

$$P(Y > y) = e^{-\alpha y}, \quad y \geq 0,$$

Y sigue una distribución $\text{Exp}(\alpha)$ y así el EMV de α está dado por

$$\hat{\alpha}_n = \bar{Y}_n^{-1} = \left(\frac{1}{n} \sum_{j=1}^n \ln X_j \right)^{-1} = \left(\frac{1}{n} \sum_{j=1}^n \ln X_{j,n} \right)^{-1}$$

Una generalización trivial concierne a

$$\bar{F}(x) = Cx^{-\alpha}, \quad x \geq u > 0, \quad \text{con } u \text{ conocido.} \quad (3.18)$$

Para $C = u^\alpha$, entonces inmediatamente se obtiene como EMV de α

$$\hat{\alpha}_n = \left(\frac{1}{n} \sum_{j=1}^n \ln \left(\frac{X_{j,n}}{u} \right) \right)^{-1} = \left(\frac{1}{n} \sum_{j=1}^n \ln X_{j,n} - \ln u \right)^{-1} \quad (3.19)$$

Generalmente se asume que \bar{F} se comporta como una distribución de Pareto sobre un cierto umbral conocido u .

Para $k =$ número de excedentes, el EMV de α y C en (3.18) se reduce a la maximización de la densidad conjunta de $(X_{k,n}, \dots, X_{1,n})$. Del teorema 2.25 se deduce que

$$f_{X_{k,n}, \dots, X_{1,n}}(x_k, \dots, x_1) = \frac{n!}{(n-k)!} (1 - Cx_k^{-a})^{n-k} C^k a^k \prod_{i=1}^k x_i^{-(a+1)},$$

De un cálculo directo se obtiene que el estimador de Hill tiene la misma forma que los EMV condicionales en el modelo subyacente (3.19) pero ahora se tiene reemplazado el valor determinístico u por el umbral aleatorio $X_{k,n}$. Además se obtiene

inmediatamente un estimador de la cola $(\bar{F}(x))^{\wedge} = \frac{k}{n} \left(\frac{x}{X_{k,n}} \right)^{-\hat{a}_{k,n}^{(H)}}$ y para el p -

cuantil

$$\hat{x}_p = \left(\frac{n}{k} (1-p) \right)^{-1/\hat{a}_{k,n}^{(H)}} X_{k,n}.$$

Además,

$$\hat{C}_{k,n} = \frac{k}{n} X_{k,n}^{\hat{a}_{k,n}^{(H)}}$$

Es importante hablar del acercamiento a la función de exceso media. Supongamos que X es una va con f.d. $F \in \text{MDA}(\Phi_\alpha), \alpha > 0$, y por conveniencia asumimos que $X > 1$ casi seguramente. Se puede escribir ahora,

$$E(\ln X - \ln t \mid \ln X > \ln t) \rightarrow \frac{1}{a}, \quad t \rightarrow \infty.$$

Denotando $u = \ln t$ y $e^*(u)$ la función de exceso media de $\ln X$ obtenemos que

$$e^*(u) \rightarrow \frac{1}{a} \quad u \rightarrow \infty$$

En sí, el estimador de Hill puede ser interpretado como la función de exceso media empírica de $\ln X$ calculada en el umbral $u = \ln X_{k,n}$.

Teorema 3.8 Propiedades del estimador de Hill

Supongamos que (X_n) es estrictamente estacionario con distribución marginal F que satisface para algún $\alpha > 0$ y $L \in \mathfrak{R}_0$, que

$$\bar{F}(x) = P(X > x) = x^{-\alpha} L(x), \quad x > 0.$$

Sea $\hat{\mathbf{a}}^{(H)} = \hat{\mathbf{a}}_{k,n}^{(H)}$ el estimador de Hill.

a) *Consistencia débil.* Asumamos que una de las siguientes condiciones es satisfecha.

- (X_n) es iid.
- (X_n) es débilmente dependiente (corta memoria)
- (X_n) es un proceso lineal

Si $k \rightarrow \infty, k/n \rightarrow 0$ cuando $n \rightarrow \infty$, entonces

$$\hat{\mathbf{a}}^{(H)} \xrightarrow{p} \mathbf{a}$$

b) *Consistencia fuerte.* Si $k / \ln \ln n \rightarrow \infty, k/n \rightarrow 0$ cuando $n \rightarrow \infty$ y (X_n) es una sucesión iid., entonces

$$\hat{\mathbf{a}}^{(H)} \xrightarrow{c.s.} \mathbf{a}$$

c) *Normalidad asintótica.* Si se impone la suposición de variación regular de segundo orden en \bar{F} y además (X_n) es una sucesión de v.a iid, entonces

$$\sqrt{k}(\hat{\mathbf{a}}^{(H)} - \mathbf{a}) \xrightarrow{d} N(0, \mathbf{a}^2) \clubsuit$$

3.4.2.3 Estimador de Deckers – Einmahl – de Haan para $\mathbf{x} \hat{\mathbf{I}} \tilde{\mathbf{n}}$

Una desventaja del estimador de Hill es que éste ha sido diseñado para $F \in \text{MDA}(H_\xi)$, $\xi > 0$, y sabemos que esta clase de modelos satisfacen muchas aplicaciones en las finanzas y en los seguros. En este método se realiza una extensión del estimador de Hill para que este cubra toda la clase H_ξ , $\xi \in \tilde{\mathbf{n}}$. Aquí, se presenta la siguiente propuesta:

$$\hat{\mathbf{x}} = 1 + H_n^{(1)} + \frac{1}{2} \left(\frac{(H_n^{(1)})^2}{H_n^{(2)}} - 1 \right)^{-1},$$

donde

$$H_n^{(1)} = \frac{1}{k} \sum_{j=1}^k (\ln X_{j,n} - \ln X_{k+1,n})$$

es el estimador recíproco de Hill y

$$H_n^{(2)} = \frac{1}{k} \sum_{j=1}^k (\ln X_{j,n} - \ln X_{k+1,n})^2.$$

Como $H_n^{(1)}$ y $H_n^{(2)}$ pueden ser interpretados como momentos empíricos, $\hat{\mathbf{x}}$ es entendido también como un estimador de momento de ξ

Hasta el momento varios estimadores han sido presentados, pero en realidad no tenemos todos los acercamientos relevantes para la estimación del parámetro de forma ξ . Sin embargo, como ya se ha dicho, la metodología de los procesos puntuales se usa cuando se discuten extremos, y esta teoría nos ayuda a encontrar procedimientos alternativos de estimación.

Teorema 3.9 El equilibrio entre el sesgo y la varianza para el estimador de Hill

Supongamos que dado que $F(x) = x^{-a} L(x)$, $a > 0$ con datos iid. se cumple la propiedad de segundo orden, es decir que

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(tx)/\bar{F}(x) - t^{-a}}{a(x)} = t^{-a} \frac{t^r - 1}{r}, \quad t > 0, \text{ existe.} \quad (3.20)$$

Donde $a(x)$ es una función medible donde su signo no varía. El lado derecho de la ecuación es interpretado como $t^{-a} \ln t$ si $r = 0$. La constante $r \leq 0$ es el parámetro de segundo orden que rige la tasa de convergencia de $\bar{F}(tx)/\bar{F}(x)$ a t^{-a} .

Y además $k = k(n) \rightarrow \infty, k/n \rightarrow 0$ cuando $n \rightarrow \infty$. Si

$$\lim_{n \rightarrow \infty} \sqrt{k} A\left(\frac{n}{k}\right) = \lambda \in \tilde{\mathfrak{n}}$$

entonces cuando $n \rightarrow \infty$

$$\sqrt{k}(\hat{\mathbf{a}}^{(H)} - \mathbf{a}) \xrightarrow{d} N\left(\frac{\mathbf{a}^3 \mathbf{l}}{\mathbf{r} - \mathbf{a}}, \mathbf{a}^2\right)$$

donde $A(x) = \mathbf{a}^{-2} a(F^{\leftarrow}(1 - x^{-1}))$. ♣

Observación. El escoger el estimador del parámetro de forma ξ a usarse no tiene una respuesta realmente clara. Todo depende de los posibles valores de ξ y de cómo se han presentado, en las propiedades de la f.d. F. Para $\xi = \alpha^{-1} > 0$ y f.d. que satisface (3.20) se probaron resultados similares al último teorema enunciado para el estimador de Pickands y el estimador de Deckers – Einmahl – de Haan (DEdH).

En el caso que $\mathbf{r} = 0$, el estimador de Hill tiene el mínimo error medio cuadrático. La eficiencia asintótica para estos estimadores depende críticamente de la interacción entre \mathbf{r} y α . Para $\xi > -2$ el estimador DEdH tiene varianza menor al de Pickands.

3.4.3 Estimación Cuantil y de Cola

Supongamos, como antes, que tenemos una muestra X_1, \dots, X_n de v.a iid. con f.d. $F \in \text{MDA}(H_\xi)$ para algún $\xi \in \tilde{\mathfrak{n}}$. Sea $0 < p < 1$ y x_p el p -cuantil correspondiente. Lo que se desea es encontrar estimadores para la cola $\bar{F}(x)$ con valores de x grandes. Se discutirá algunos de ellos con más detalle y exponiendo sus propiedades y limitaciones más importantes.

Para empezar es necesario destacar que la estimación fuera del rango de los datos puede ser hecha sólo si se imponen suposiciones extras del modelo. $X_{k,n}$ aparece como un estimador natural del $(1 - (k-1)/n)$ -cuantil. El rango $[X_{n,n}, X_{1,n}]$ nos permite construir la estimación dentro del conjunto de datos del $(1 - n^{-1})$ -cuantil. A pesar de que en una aplicación práctica, p es fijo, desde el punto de vista matemático la diferencia entre cuantiles altos dentro y fuera del conjunto puede ser descrita como sigue:

- cuantiles altos dentro del conjunto: $p = p_n \uparrow 1, n(1 - p_n) \rightarrow c, c \in (1, \infty]$,
- cuantiles altos fuera del conjunto: $p = p_n \uparrow 1, n(1 - p_n) \rightarrow c, 0 \leq c < 1$.

Teorema 3.10 Estimación de cuantiles altos I

Supongamos una muestra X_1, \dots, X_n de v.a iid. con f.d. $F \in \text{MDA}(H_\xi)$ para algún $\xi \in \bar{\mathfrak{n}}$, y que F tiene una densidad positiva f . Asumimos que la densidad U' está en $\mathfrak{R}_{\xi,1}$. Escribiendo $p = p_n$ y $k = k(n) = [n(1-p_n)]$, donde $[x]$ es la parte entera de x . Si se cumplen las condiciones

$$p_n \rightarrow 1 \quad \text{y} \quad n(1 - p_n) \rightarrow \infty$$

entonces,

$$\sqrt{2k} \frac{X_{k,n} - x_p}{X_{k,n} - X_{2k,n}} \xrightarrow{d} N\left(0, 2^{2x+1} \mathbf{x}^2 / (2^x - 1)^2\right) \clubsuit$$

Nota: La condición $U' \in \hat{\mathbf{A}}_{x,1}$ puede ser reformulada en términos de F . Por ejemplo, para $\xi > 0$, la condición se convierte en $f \in \hat{\mathbf{A}}_{1,1/x}$.

Teorema 3.11 Estimación de cuantiles altos II

Supongamos una muestra X_1, \dots, X_n de v.a iid. con f.d. $F \in \text{MDA}(H_\xi)$ para algún $\xi \in \bar{\mathfrak{n}}$, y que $\lim_{n \rightarrow \infty} n(1 - p) = c$ para algún $c > 0$.

Sea \hat{x}_p definido por

$$\hat{x}_p = \frac{(k / (n(1 - p)))^{\hat{x}} - 1}{1 - 2^{-\hat{x}}} (X_{k,n} - X_{2k,n}) + X_{k,n}$$

siendo $\hat{\mathbf{X}}$ el estimador de Pickands. Entonces, para todo $k > c, k$ fijo,

$$\frac{\hat{x}_p - x_p}{X_{k,n} - X_{2k,n}} \xrightarrow{d} Y,$$

donde

$$Y = \frac{(k/c)^x - 2^{-x}}{1 - 2^{-x}} + \frac{1 - (Q_k/c)^x}{\exp\{xH_k\} - 1}$$

Las v.a H_k y Q_k son independientes, Q_k sigue una distribución Gamma con parámetro $2k+1$ y $H_k = \sum_{j=k+1}^{2k} \frac{E_j}{j}$ para v.a iid. exponenciales estándar E_1, E_2, \dots ♦

3.5 Construcción de excesos sobre un umbral

3.5.1 Construcción de la distribución generalizada de Pareto (DGP)

La metodología introducida hasta ahora fue obtenida ya sea de la suposición de que los datos vienen de una Distribución Generalizada de Valores Extremos (VEG) o de funciones de distribución que pertenecen a su máximo dominio de atracción. Ahora, se exhibe una alternativa basada en los excedentes de umbrales altos.

Supongamos que X, X_1, \dots, X_n son iid. con f.d. $F \in \text{MDA}(H_\xi)$ para algún $\xi \in \tilde{n}$.

Primero, se escoge un umbral alto u y se denota por

$$N_u = \text{card}\{i : i = 1, \dots, n, X_i > u\}$$

al número de excedentes de u derivados de X_1, \dots, X_n . Los excesos correspondientes se denotan por Y_1, \dots, Y_{N_u} .

Gráfico No. 3.6

Excedentes



En la figura se observan los datos X_1, \dots, X_{13} y los correspondientes excesos.

Y_1, \dots, Y_{N_u} sobre el umbral u .

La función de distribución de exceso de X está dada por

$$F_u(y) = P(X - u \leq y \mid X > u) = P(Y \leq y \mid X > u), \quad y \geq 0$$

La última relación puede ser escrita también como:

$$\bar{F}(u + y) = \bar{F}(u)F_u(y). \quad (3.21)$$

Récordando la definición de la DGP: $G_{\xi, \beta}$ con parámetros $\xi \in \mathbb{R}$ y $\beta > 0$ tiene una distribución de cola

$$\bar{G}_{x, \beta}(x) = \begin{cases} \left(1 + \frac{x}{\beta}\right)^{-1/\xi} & \text{si } x \neq 0, \\ e^{-x/\beta} & \text{si } x = 0, \end{cases}$$

donde

$$\begin{aligned} x &\geq 0 && \text{si } \xi \geq 0, \\ 0 \leq x \leq -\beta/\xi &&& \text{si } \xi < 0. \end{aligned}$$

Del teorema 2.23 se obtiene que para $\bar{F}_u(y)$,

$$\lim_{u \uparrow x_F} \sup_{0 < x < x_F - u} \left| \bar{F}_u(x) - \bar{G}_{\mathbf{x}; \mathbf{b}(u)}(x) \right| = 0$$

donde $\beta > 0$. De esto, se sigue que para u grande

$$\bar{F}_u(y) \approx \bar{G}_{\mathbf{x}; \mathbf{b}(u)}(y) \quad (3.22)$$

De donde se obtiene un estimador de la forma:

$$(\bar{F}_u(y))^\wedge = \bar{G}_{\hat{\mathbf{x}}; \hat{\mathbf{b}}}(y) \quad (3.23)$$

para $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{N_u}$ y $\hat{\mathbf{b}} = \hat{\mathbf{b}}_{N_u}$.

Un estimador natural de $\bar{F}(u)$ está dado por la f.d. empírica

$$(\bar{F}(u))^\wedge = \bar{F}_n(u) = \frac{1}{n} \sum_{i=1}^n I_{\{X_i > u\}} = \frac{N_u}{n}$$

En el caso que $\xi \geq 0$, un estimador del cuantil x_p resulta inmediatamente:

$$\hat{x}_p = u + \frac{\hat{\mathbf{b}}}{\hat{\mathbf{x}}} \left(\left(\frac{n}{N_u} (1-p) \right)^{-\hat{\mathbf{x}}} - 1 \right)$$

Además, para $\hat{\mathbf{x}} < 0$ un estimador del punto final derecho x_F de F está dado por

$$\hat{x}_F = u - \frac{\hat{\mathbf{b}}}{\hat{\mathbf{x}}}$$

Las propiedades estadísticas de los estimadores resultantes dependen crucialmente de las propiedades que se derivan de las distribuciones de los procesos puntuales de los excedentes (N_u) .

La opción del tamaño de la cola puede afectar nuestras estimaciones con su efecto sobre nuestra estimación del índice de la cola, pero la teoría nos da poca guía sobre cuán grande debe ser la cola. El escoger el tamaño de la cola (valor de u) es complicado por una compensación entre la varianza y el sesgo. Si aumentamos el tamaño de la cola, el umbral de la cola se mueve hacia el centro de la distribución y conseguimos más observaciones de la cola. Esto aumenta la precisión y reduce la varianza de nuestro estimador de cola, pero también aumenta su sesgo poniendo relativamente más peso en el centro antes que en las observaciones extremas.

Si disminuimos el tamaño de la cola, nos movemos hacia fuera a lo largo de la cola; esto disminuye al sesgo pero incrementa la varianza porque tenemos menos observaciones con las cuales trabajar.

Quizás la mejor manera de manejar la opción del tamaño de la cola es estimar el índice de cola para una gama de tamaños de cola, y después elegir un tamaño de la cola donde el diagrama del estimador de la cola contra el tamaño de la cola (esperanzadamente) llega a ser más o menos horizontal. Este acercamiento tiene la atracción que intenta extraer la máxima información posible de nuestros datos, no obstante, de una manera informal, y es recomendado por los investigadores de Zurich (Embrechts, McNeil, etc.). Una alternativa, sugerida por Danielsson y de Vries, es la de estimar un tamaño de cola óptimo que minimice el error medio cuadrático de la función de pérdida y así se produzca una compensación óptima entre la varianza y el sesgo. Sin embargo, este acercamiento tiene sus limitaciones, entre ellas está que requiere de muchos datos.

Un método que es inmediatamente usado en la práctica está basado en la linealidad de la función de exceso media $e(u)$ para la DGP.

Sabemos, que si la v.a X tiene una f.d. $G_{\xi,\beta}$,

$$e(u) = E(X - u \mid X > u) = \frac{\beta + \xi u}{1 - \xi}$$

donde $\xi < 1$

$$\begin{array}{ll} u \geq 0 & \text{si } \xi \geq 0, \\ 0 \leq u \leq -\beta/\xi & \text{si } \xi < 0. \end{array}$$

Por lo tanto $e(u)$ es lineal. Lo que se sugiere es un acercamiento gráfico para escoger u : escoger $u > 0$ tal que $e_n(u)$ es aproximadamente lineal para $x \geq u$.

Embrechts, McNeil, etc. recomiendan el uso de gráficos para reforzar el juicio y el

sentido común y comparar las estimaciones resultantes a través de una variedad de valores de u .

3.5.1.1 Estimación de máxima verosimilitud

Cabe recordar el hecho de que nuestros datos originales $X = (X_1, \dots, X_n)$ son iid. con función de distribución F . Asumimos que F es la DGP con parámetros ξ y β .

Las ecuaciones de verosimilitud pueden ser derivadas y resueltas numéricamente obteniendo $\hat{\mathbf{x}}_n$ y $\hat{\mathbf{b}}_n$.

Las propiedades usuales del EMV como consistencia y eficiencia asintótica se mantienen.

Debido a la ecuación (3.22) es más real asumir una DGP para los excedentes. Las ecuaciones de verosimilitud condicional resultantes pueden ser resueltas por medio de una reparametrización $(\xi, \beta) \rightarrow (\xi, \tau)$, donde $\tau = -\xi/\beta$. De esto, obtenemos los siguientes resultados:

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}(\mathbf{t}) = N_u^{-1} \sum_{i=1}^{N_u} \ln(1 - tY_i),$$

donde τ satisface

$$h(\mathbf{t}) = \frac{1}{\mathbf{t}} + \frac{1}{N} \left(\frac{1}{\hat{\mathbf{x}}(\mathbf{t})} + 1 \right) \sum_{i=1}^N \frac{Y_i}{1 - tY_i} = 0.$$

La función $h(\tau)$, definida para $\tau \in (-\infty, \max(Y_1, \dots, Y_{N_u}))$, es continua en cero.

3.5.2 Principios básicos para la adecuación de la DGP

En aplicaciones prácticas, se deben resolver dos asuntos importantes:

1. Construir la f.d. condicional $F_u(x)$ para el umbral u apropiado.
2. Construir la f.d. incondicional $F(x)$ nuevamente para un umbral u apropiado.

En la práctica se debe tener cuidado sobre el rango preciso disponible de datos

y/o el intervalo sobre el cual queremos adecuarnos.

Después de estas estimaciones, usando (3.23) se puede dibujar una estimación de $F_u(x)$. Para estimar la cola de la f.d. incondicional F la cual nos da información sobre la frecuencia con la cual un nivel u es excedido, se utiliza

$$\left(\bar{F}(u+y)\right)^\wedge = \frac{N_u}{n} \left(1 + \hat{\mathbf{x}} \frac{y}{\hat{\mathbf{b}}}\right)^{-1/\hat{\mathbf{x}}}$$

Un cálculo directo nos permite expresar $\hat{F}(z)$ como una DGP de tres parámetros, esta f.d. está diseñada solo para una construcción adecuada de los datos sobre el umbral u :

$$\hat{F}(z) = 1 - \left(1 + \hat{\mathbf{x}} \frac{z - u - \hat{v}}{\hat{\mathbf{b}}'}\right)^{-1/\hat{\mathbf{x}}}, \quad z \geq u,$$

donde

$$\hat{v} = \frac{\hat{\mathbf{b}}}{\hat{\mathbf{x}}} \left(\left(\frac{N_u}{n} \right)^{\hat{\mathbf{x}}} - 1 \right) \quad y \quad \hat{\mathbf{b}}' = \hat{\mathbf{b}} \left(\frac{N_u}{n} \right)^{\hat{\mathbf{x}}}.$$

Bajo u , donde los datos son típicamente abundantes, varias técnicas estándar pueden ser usadas. Construyendo tanto la DGP sobre u y la f.d. empírica bajo u , se puede obtener una buena construcción global.

Finalmente, usando la construcción de $F(z)$, se pueden dar estimaciones para los p – cuantiles, $p \geq F(u)$.

La confiabilidad estadística de estas estimaciones en general, llega a ser muy difícil de juzgar. Aunque se pueden estimar intervalos de confianza para estos estimadores, tales construcciones son fuertemente basadas en suposiciones matemáticas que no son verificables en la práctica. Es por esto que estos modelos basados en estimaciones deben ser la base para una discusión detallada con los responsables “de los datos”. La TEV ofrece una plataforma ideal para la simulación

de experimentos y énfasis de escenarios.

3.6 Tópicos especiales

3.6.1 El índice extremo

3.6.1.1 Definición y propiedades elementales

Anteriormente presentamos un buen material de extremos. La mayoría de veces nos restringimos al caso de observaciones iid. Sin embargo, en la realidad los extremos tienden a ocurrir en grupos causados por dependencia local en los datos. Las observaciones siempre están entrelazados con otras y por tanto no son independientes, algunos pueden ser observaciones financieras como tasas de cambio, ganancia que genera la variación de precios, etc. Si un valor grande en una serie de tiempo ocurre, podemos ver generalmente un grupo de valores grandes en cortos períodos de tiempo.

El índice extremo es una cantidad que, en una manera intuitiva, permite caracterizar la relación entre la estructura de dependencia de los datos y su comportamiento de extremos. Notamos M_n como el máximo de la muestra X_1, \dots, X_n , (\tilde{X}_n) esta asociada a una sucesión iid. y (\tilde{M}_n) denota la sucesión correspondiente de máximos.

3.6.1.2 Interpretación y estimación del índice extremo

Empezamos con algo de un ejemplo simple mostrando la relevancia de la noción del índice extremo.

Ejemplo 3.12 Supongamos que se quiere construir un canal en el río Machángara para canalizar las aguas que fluyen en el transcurso de la ciudad de Quito con una probabilidad del 95% en los próximos 10 años. Supongamos que ha sido establecido que los percentiles 98.97% y 99.488% de la altura de los caudales anuales son de 2m y 3m respectivamente. Si los máximos anuales se cree que son iid., entonces el canal debe ser de 3m de altura ($0.99488^{10} \approx 0.95$). Pero si los máximos anuales son estacionarios con índice extremo $q = 0.5$, entonces la altura de 2m es suficiente ($0.9897^5 \approx 0.95$). ♣

Este ejemplo nos lleva a calcular previamente la estimación del índice q , debe ser un resultado central en el estadístico de valores extremos para datos dependientes. La estimación de q será basada en un número de diferentes interpretaciones probabilísticas del índice extremo. Empezamos entonces a la construcción de diferentes estimadores.

Empezamos con una primera Aproximación de la Estimación de q :

El Método de Bloques

Partiendo de la definición del índice extremo, tenemos

$$P(M_n \leq u_n) \approx P^q(\tilde{M}_n \leq u_n) = F^{nq}(u_n),$$

siempre que $n \bar{F}(u_n) \rightarrow \tau > 0$. De aquí

$$\lim_{n \rightarrow \infty} \frac{\ln P(M_n \leq u_n)}{n \ln F(u_n)} = q. \quad (3.24)$$

Esta relación de límite simple sugiere la construcción de estimadores simples de q . Ya que no conocemos $F(u_n)$ ni $P(M_n \leq u_n)$, esas cantidades tiene que ser reemplazadas por estimadores. Un candidato obvio para estimar la cola $\bar{F}(u_n)$ es su versión empírica

$$\frac{N}{n} = \frac{1}{n} \sum_{i=1}^n I_{\{X_i > u_n\}}.$$

Esta elección es motivada por el teorema de Glivenko_Cantelli para sucesiones ergódicas estacionaria (X_n) . Para encontrar un estimador empírico para $P(M_n \leq u_n)$ no es sencillo. Recordemos que la condición $D(u_n)$ implica

$$P(M_n \leq u_n) \approx P^k(M_{[nk]} \leq u_n). \quad (3.25)$$

Con k lentamente creciente $k = k(n)$, la aproximación (3.25) forma la base para el método de bloques. Para el bien los de argumentos suponemos $n = rk$, para r entero

$r = r(n) \rightarrow \infty$ y $k = k(n) \rightarrow \infty$. Por otra parte $r = [n/k]$ donde $[x]$ es la parte entera de x . Esto divide a la muestra en k bloques de tamaño r :

$$X_1, \dots, X_r; \dots; X_{(k-1)r+1}, \dots, X_{kr} \quad (3.26)$$

Para cada bloque calculamos el máximo

$$M_r^{(i)} = \max\{X_{(i-1)r+1}, \dots, X_{ir}\}, \quad i = 1, 2, \dots, k.$$

relacionando (3.25) entonces sugiere la aproximación

$$\begin{aligned} P(M_n \leq u_n) &= P\left(\max_{1 \leq i \leq k} M_r^{(i)} \leq u_n\right) \approx P^k(M_r \leq u_n) \\ &\approx \left(\frac{1}{k} \sum_{i=1}^k I_{\{M_r^{(i)} \leq u_n\}}\right)^k = \left(1 - \frac{K}{k}\right)^k. \end{aligned}$$

Una combinación de este argumento heurístico con (3.24) lleva a la siguiente estimación de q :

$$\hat{q}_n^{(1)} = \frac{k \ln(1 - K/k)}{n \ln(1 - N/n)} = \frac{1 \ln(1 - K/k)}{r \ln(1 - N/n)}. \quad (3.27)$$

Aquí N es el número de excedentes del umbral u_n por X_1, \dots, X_n , y K es el número de bloques con uno o más excedentes. Un argumento de expansión de Taylor produce un segundo estimador

$$\hat{q}_n^{(2)} = \frac{K}{N} = \frac{1}{r} \frac{K/k}{N/n} \approx \hat{q}_n^{(1)}.$$

El método de bloques ocurre por agrupación en los datos. Si el evento

$$\{M_r^{(i)} = u_n\} = \bigcup_{j=1}^r \{X_{(i-1)r+j} > u_n\}$$

dice que un grupo ocurrió en el i -ésimo bloque. Esos eventos caracterizan el comportamiento extremo de (X_n) si suponemos que el tamaño de $r(n)$ de los bloques crecen lentamente con n . Esto nos da alguna sensación para la estructura de

dependencia de los datos en la sucesión (X_n) . En ese sentido, el índice extremo es una medida de la tendencia del grupo de excedentes del umbral alto en una sucesión estacionaria.

Regresando a lo anterior podemos decir que no es fácil obtener la estimación de q . Esto parece estar en las mismas condiciones del sección 3.1, donde probamos estimar el índice ξ de una distribución de valor extremo.

Conclusiones

Como es de conocimiento, los resultados de la mayoría de estudios se obtienen a través de la aplicación de la Estadística Clásica, la misma que se ocupa del estudio de valores centrales. En este caso con el uso del Teorema Central del Límite se construye generalmente una distribución empírica normal para así hacer uso de su información. En los casos en que es necesario el estudio de acaecimientos extremos la estadística tradicional tiene dos problemas intrínsecos: Puesto que la mayoría de las observaciones yacen cerca del centro de cualquier distribución empírica, los acercamientos paramétricos tradicionales nos conducen a construir las curvas que acomodan las observaciones centrales, antes que acomodar las observaciones de la cola que son más importantes para estos propósitos. Esto conduce a una mala estimación. Por otro lado, estos acercamientos imponen generalmente distribuciones ante los datos que no tienen ningún sentido cuando están utilizados para la valoración de la cola. La Teoría de Valores Extremos nos dice cómo la distribución de extremos debería ser, por lo menos asintóticamente, esta teoría está libre de estos problemas y está diseñada específicamente para la evaluación de la cola.

El uso de la Teoría de Valores Extremos nos conduce a una estimación más eficiente de los parámetros, porque nos indica en cierto modo la forma que deberían tener las colas. Además la información que brinda va más allá de la totalidad de nuestros datos muestrales. La TVE da las mejores estimaciones de acontecimientos extremos y representa el mejor acercamiento para medir la incertidumbre inherente en el problema.

La Teoría de Valores Extremos ofrece una solución muy natural a un problema práctico extremadamente difícil, que es el de cómo estimar los cuantiles extremos donde tenemos, por definición, escasos datos que ocurren con poca frecuencia. La contribución de la TVE seguirá siendo básica y útil pues nos ayuda a dibujar curvas suaves a través de las colas de funciones empíricas. Cualquier información que se

pueda extraer es inevitablemente muy tentativa ya que estamos trabajando únicamente en la cola.

Los datos que se usan para la modelización matemática usando Teoría de Valores Extremos deben cumplir con todas las exigencias teóricas, de no ser así se debe realizar un tratamiento especial a los datos, además deben cumplir con ciertas características básicas como por ejemplo colas pesadas, independencia, etc.

La estimación de los parámetros requieren no solo de buenas técnicas de estimación sino también de buenos métodos numéricos, recomendamos usar el método de Newton Raphsón, además se debe tener una buena interpretación analítica de los gráficos de los parámetros.

Bibliografía

- Birkhäuser, Basel, S joint with J. Hüslér und R.-D. Reiß., (1994); *Extremes and Rare Events*. DMV, Seminar Bd. 23.
- Ec. Geneve, (2000), *Introduction to Extreme Value Theory*. E.E.U.U.
- Embrechts, P., C. Klüppelberg, and T. Mikosch, (1997), *Modelling extremal events for insurance and finance*. Berlin: Spring Verlag.
- Falk, M. and Reiss R.-D., (1989), *Efficient Estimation of the Canonical Dependence Function*.
- Gilbert W. Bassett Jr., Hsiu Lang Chen, (1999), *Quantile Style: Return-Based Attribution Using Regression Quantiles*, U.S.A.
- Gomes, M. Ivette, (1749), *Generalized Jackknife Moment estimators of the tail index*, Lisboa, Portugal, Faculty of Science of Lisbon.
- Gunther Matthys _ & Jan Beirlant, (2000), *Adaptive Threshold Selection in Tail Index Estimation*, University Center for Statistics and Catholic University Leuven.
- Kennedy John. B, (), *Estadística Para Ciencias e Ingeniería*.
- Kevin Dowd, Ph.D., (1999), *The Extreme Value Approach to VaR*.
- Manfred Gilli, Evis Këllezi, (1997), *Extreme Value Theory for Tail – Related Risk Measures*, Switzerland: University of Geneva.

Markus Schmutz, Richard R. Doerr, (1998), *The Pareto Model in Property Reinsurance Formulas and applications whit reference to other loss distributions functions*. Zurich: Products Management Department.

McNeil, A. J., (1998), "Calculating quantile risk measures for financial return series using extreme value theory." Mimeo. ETH Zentrum, Zürich.

Mendenhall, Scheaffer, Wackerly, (1968), *Estadística Matemática con Aplicaciones*. Nueva York: Grupo Editorial Iberoamérica.

MD* TECH, *Fitting GP Distributions to the upper tail*. www.mdtech.de

———, *Parameters estimators for GP Model*. www.mdtech.de

M.D. Pandey, J.K. Vrigling, (1998), *The Use of L- Momentes in the POT Approach for Estimating Extreme Quantiles of Wind Velocity*, Canadá, Ontario: University of Waterloo.

Nolan, John P., (1999), Tail Behavior, Modes and other characteristics of Stable Distributions. *E.E.U.U*: American University.

Ser-Huang Poon, Ichael Rockinger y Jonathan Tawn, (2001), *New Extreme-Value Dependence Measures and Finance Applications*, Huang Poon University.

Sidney Resnick, Catalin Starica, (1996), *Smoothing the Moment Estimator of the Extreme Value Parameter*, NY: School of ORIE, Cornell University.