

**MODELOS DE PÉRDIDAS AGREGADAS (LDA) Y DE
LA TEORÍA DEL VALOR EXTREMO PARA
CUANTIFICAR EL RIESGO OPERATIVO
TEORÍA Y APLICACIONES**

GUILLERMO LEÓN ARIAS PINEDA



**ESCUELA DE CIENCIAS Y HUMANIDADES
DEPARTAMENTO DE CIENCIAS BÁSICAS
MAESTRÍA EN MATEMÁTICAS APLICADAS
MEDELLÍN
2010**

**MODELOS DE PÉRDIDAS AGREGADAS (LDA) Y DE
LA TEORÍA DEL VALOR EXTREMO PARA
CUANTIFICAR EL RIESGO OPERATIVO
TEORÍA Y APLICACIONES**

GUILLERMO LEÓN ARIAS PINEDA

Trabajo presentado como requisito parcial para optar al Título de
Magíster en Matemática Aplicada

Director:
JUAN GUILLERMO MURILLO GÓMEZ
Magister en Ingeniería Administrativa
Universidad Nacional de Colombia

**ESCUELA DE CIENCIAS Y HUMANIDADES
DEPARTAMENTO DE CIENCIAS BÁSICAS
MAESTRÍA EN MATEMÁTICAS APLICADAS
MEDELLÍN
2010**

UNIVERSIDAD EAFIT
ESCUELA DE CIENCIAS BÁSICAS Y HUMANAS

Los jurados abajo firmantes certifican que han leído y que aprueban la tesis titulada **“MODELOS DE PÉRDIDAS AGREGADAS (LDA) Y DE LA TEORÍA DEL VALOR EXTREMO PARA CUANTIFICAR EL RIESGO OPERATIVO, TEORÍA Y APLICACIONES”**, presentada por **GUILLERMO LEÓN ARIAS PINEDA**.

Fecha: Junio de 2010

Director: _____

Juan Guillermo Murillo Gómez

Jurado: _____

Jurado: _____

AGRADECIMIENTOS

Agradezco a todas aquellas personas que me han colaborado con la realización de esta tesis especialmente al Asesor Magister Juan Guillermo Murillo Gómez por su contribución acertada sobre el desarrollo del proceso. También debo agradecer a mi esposa Lucia Celina Torres Torres por la paciencia que ha tenido de tolerar tanto tiempo de ausencia para poder dedicarme a sacar adelante este proyecto.

Muchas gracias.

RESUMEN

El aspecto más relevante desarrollado en el presente trabajo es implementar dos métodos para medir el riesgo operativo: El Valor en Riesgo (VaR) y uno de los modelos Picos sobre un umbral (POT) de la Teoría Básica del Valor Extremo (EVT)¹.

Primero se hace una introducción a lo que es el riesgo y riesgo operativo y un poco de contexto histórico de los dos métodos. Segundo se presentan las diferentes distribuciones de probabilidad para los eventos de riesgo (frecuencias y severidad) más usadas, se presenta la forma de estimar los parámetros y como es el generador de procesos para simularlas. Posteriormente se aclaran dos modelos para tratar valores extremos de la Teoría del Valor Extremo el Máximo por bloques y el Picos sobre un umbral (POT) con sus respectivas distribuciones más usadas: la Distribución Generalizada de Pareto y la Distribución de Valor Extremo Generalizada (en el Anexo A, B, y C se presenta también un resumen de las distribuciones: Familia Beta Transformada Familia Gamma Transformada y Distribuciones para grandes pérdidas). Se hace explícita la forma de calcular el VaR y Pérdidas Esperadas y la pérdidas inesperadas que nos permite reservar un capital para proteger una empresa del Riesgo Operativo.

Una vez abordada la teoría se pasa a un tercer capítulo donde se hace el análisis de una línea de negocios la Banca minorista y de un evento de riesgo Fraude Externo de los datos de la frecuencia que se obtuvieron de la superintendencia Financiera y de la Severidad de una institución financiera.

Dentro del análisis de los datos primero se aplican pruebas de bondad de ajuste a los eventos de riesgo (frecuencia y severidad), se estiman los parámetros de los

¹La Teoría del Valor Extremo, EVT por sus siglas en inglés es una teoría que ha sido aplicada a la medición del riesgo. Análogo al teorema central del Límite provee resultados que se pueden aplicar en el límite, es decir en forma asintótica. La teoría permite analizar de una mejor manera una medida tan universal como el VaR.

modelos VaR y EVT, se obtiene la distribución de pérdida agregada, se calculan las pérdidas esperadas, las pérdidas inesperadas y el capital regulatorio para el Riesgo Operacional.

Se establece una comparación de la eficiencia de los dos métodos Valor en riesgo y Teoría de Valor Extremo en la medición del Riesgo Operativo.

ÍNDICE GENERAL

INTRODUCCIÓN	21
CAPÍTULO 1 RIESGO Y RIESGO OPERATIVO	25
1.1 Contextualización del riesgo y el riesgo operativo	25
1.2 Valor en Riesgo	25
1.3 Calculo del VaR de forma paramétrica	26
1.4 Desviación estándar	27
CAPÍTULO 2 TEORÍA Y MODELOS DE VALORES EXTREMOS	33
2.1 Teoría Básica de Valor Extremo	33
2.1.1 Introducción	33
2.1.2 Formulación de la relación límite para el máximo y el mínimo .	35
2.1.3 Formulaciones alternativas de la relación límite	37
2.1.4 Distribución Generalizada de Valores Extremos	46
2.1.5 Dominios de Atracción y Convergencia Débil	50
2.2 Métodos de umbral y Distribución Generalizada de Pareto	61
2.2.1 Introducción	61
2.2.2 Ajuste de excesos sobre un umbral	62
2.3 Estimación del índice de valor extremo y pruebas	69
2.3.1 Introducción	69
2.3.2 El estimador de Hill $\xi > 0$	72

CAPÍTULO 3 MÉTODO DE DISTRIBUCIÓN DE PÉRDIDA AGREGADA (LDA)	87
3.1 Introducción	87
3.2 Distribuciones de frecuencias	90
3.3 Distribución de severidad	91
3.4 Valor en riesgo de mercado y operacional	92
CAPÍTULO 4 APLICACIÓN	97
4.1 Introducción	97
4.2 Análisis exploratorio de los datos de frecuencia y severidad	99
4.3 Pruebas de bondad de ajuste para los datos	102
4.3.1 Pruebas de Bondad de ajuste para los datos de frecuencia . . .	102
4.3.2 Pruebas de bondad de ajuste para los datos de severidad . . .	102
4.4 Distribución de pérdida agregada	103
4.4.1 Aplicando el Método de pérdidas agregadas (LDA)	105
4.4.2 Aplicando la Teoría Básica del Valor Extremo	107
CONCLUSIONES	117
ANEXO A FUNCIÓN DE DISTRIBUCIÓN ACUMULADA PARA LA PÉRDIDA POR FRAUDE EXTERNO CON SEVERIDAD NORMAL	119
ANEXO B GRÁFICO Q-Q PARA LA PÉRDIDA PARA FRAUDE EXTERNO CON SEVERIDAD NORMAL	121
ANEXO C VaR PARA DIFERENTES PERCENTILES Y VALOR DEL ESTADÍSTICO DE PRUEBA SMIRNOV-KOLMOGOROV	123
ANEXO D CÓDIGO DE MATLAB SIMULACIÓN DE MONTECARLO (BINOMIAL NEGATIVA-NORMAL)	125
BIBLIOGRAFÍA	129

ÍNDICE DE TABLAS

CAPÍTULO 1 RIESGO Y RIESGO OPERATIVO	25
1.1 Matriz de pérdidas	30
CAPÍTULO 4 APLICACIÓN	97
4.1 Eventos de riesgo durante los últimos cinco años	99
4.2 Pérdidas económicas por eventos de riesgo en los últimos cinco años	99
4.3 Tabla de frecuencia de la frecuencia del fraude externo	100
4.4 Medidas Frecuencia de la Frecuencia del Fraude Externo	101
4.5 Frecuencia de la Severidad del Fraude Externo	101
4.6 Medidas de resumen de la Severidad del Fraude Externo	101
4.7 Prueba de bondad de ajuste para la frecuencia	102
4.8 Prueba de bondad de ajuste para la severidad	102
4.9 Parámetros de las distribuciones utilizadas para hallar la distribu- ción de pérdida	104
4.10 Distribuciones de mejor ajuste de las pérdidas económicas causadas por el Fraude Externo	104
4.11 Parámetros de las Distribuciones que mejor ajustan a las Pérdidas Económicas	104

4.12 Pérdidas esperadas e inesperadas método LDA 105

4.13 Calculo de las Pérdidas Agregadas promedio y del VaR Operacional 106

4.14 Parámetros de los valores por encima del umbral seleccionado . . . 109

4.15 Estimación de parámetros de la distribución de Pérdida de la cola
derecha 110

4.16 Intervalos de Confianza para el parámetro de forma 110

4.17 Medidas OpVaR y Expected Shortfall, ES_α para cuantificar el riesgo 111

4.18 Valores muestrales que están por encima del umbral 113

4.19 Algunos posibles resultados obtenidos a partir de la función de Dis-
tribución y función de supervivencia teórica 115

4.20 Pérdidas para varios percentiles y frecuencia de pérdida 115

4.21 Comparación de las pérdidas esperadas en la distribución de la cola 116

4.22 Estimación de las pérdidas esperadas en la cola 116

**ANEXO C VaR PARA DIFERENTES PERCENTILES Y VALOR DEL ES-
TADÍSTICO DE PRUEBA SMIRNOV-KOLMOGOROV 123**

C.1 123

ÍNDICE DE FIGURAS

CAPÍTULO 1 RIESGO Y RIESGO OPERATIVO	25
1.1 Distribución de pérdidas agregadas	31
CAPÍTULO 3 MÉTODO DE DISTRIBUCIÓN DE PÉRDIDA AGREGADA (LDA)	87
3.1 Modelos agregados de severidad y de frecuencia	89
3.2 Distribución de pérdidas: esperadas e inesperadas	93
CAPÍTULO 4 APLICACIÓN	97
4.1 Distribución frecuencia de los diferentes eventos del riesgo operativo	99
4.2 Distribución del impacto económico de cada uno de los eventos de Riesgo operativo	100
4.3 Histograma de la frecuencia del fraude externo	100
4.4 Histograma de la frecuencia del fraude externo	101
4.5 Convolución de la distribución de Frecuencia y la Severidad	103
4.6 Modelos Agregados de Severidad y de Frecuencia	103
4.7 Ajuste de las pérdidas económicas a la distribución Logísticas	104
4.8 Número de iteraciones Vs. Promedio de pérdidas agregadas	106

4.9 Número de iteraciones Vs. OpVaR 107

4.10 Pérdidas esperadas, inesperadas y extremas 108

4.11 Ajuste de las pérdidas económicas en los valores extremos a la Beta 110

4.12 Función de Distribución Empírica versus Función de Distribución
teórica 112

4.13 Cuantil empírico versus cuantil teórico 113

4.14 Función de Distribución Empírica versus Función de Distribución
teórica 115

ANEXO A FUNCIÓN DE DISTRIBUCIÓN ACUMULADA PARA LA PÉRDIDA POR FRAUDE EXTERNO CON SEVERIDAD NORMAL 119

A.1 Función de distribución Acumulada para las pérdidas de Fraude Ex-
terno 119

ANEXO B GRÁFICO Q-Q PARA LA PÉRDIDA PARA FRAUDE EXTERNO CON SEVERIDAD NORMAL 121

B.1 Gráfico Q-Q para la pérdida para fraude externo con severidad nor-
mal 121

LISTA DE ABREVIATURAS Y SIMBOLOS

$\stackrel{d}{=}$: Igualdad de distribución
\rightarrow^d	: Converge en distribución
\rightarrow^p	: Converge en probabilidad
$a(t) \sim b(t)$: $\lim_{t \rightarrow 0} \frac{a(t)}{b(t)} = 1$
α	: Índice de cola
ξ	: Índice de valor extremo
Γ	: Función Gamma
$1 - F^n$: Función de distribución empírica continua a la izquierda
AMA	: Modelos de medición avanzada
a_+	: $\max(a, 0)$
a_-	: $\min(a, 0)$
$a \vee b$: $\max(a, b)$
$a \wedge b$: $\min(a, b)$
$\llbracket a \rrbracket$: El entero más grande menor o igual a a
$\lceil a \rceil$: El entero más pequeño menor o igual a a
a.s.	: Casi seguro
a_n	: Constante para centrar
b_n	: Constante para centrar
$\beta(x)$: Función positiva que depende del umbral
$D(G_\xi)$: Dominio de atracción de G_z
$E(X)$: Valor esperado de X
f^-	: Versión de la función f continua a la izquierda
F	: Función de distribución acumulada
F_n	: Función de distribución empírica continua a la derecha
$F_X^n(X)$: n -ésima convolución de la distribución de severidad
$F_u(y)$: Función de distribución de los excesos
G_ξ	: Función de distribución de valor extremo
$G_s(X)$: Función de distribución de pérdidas agregadas

GP	:	Generalizada de Pareto
G_1	:	Distribución tipo Gumbel
G_2	:	Distribución tipo Frechet
G_3	:	Distribución tipo Weibull
\overline{G}	:	Distribución de la cola superior
G_S^{-1}	:	Cuantil de las pérdidas agregadas
i.i.d.	:	Independiente e idénticamente distribuida
L	:	Pérdida total
LDA	:	(Loss Distribution Approach) Distribución de pérdidas
$\max(X_1, \dots, X_n)$:	Máximo de una muestra
$\min(X_1, \dots, X_n)$:	Mínimo de una muestra
n	:	Tamaño de la muestra
N_{ij}	:	Variable aleatoria que representa el número de eventos de riesgo en la celda i, j (Frecuencia de los eventos)
N_u	:	Número de excesos u
$OpVaR(\alpha)$:	Carga de capital por riesgo operacional
R_+	:	$[0, \infty]$
RV_α	:	Variación regular con índice α
S	:	Pérdidas agregadas
S_n	:	Suma de las n variables aleatorias de la muestra
S_{ij}	:	Pérdida total de la celda i, j de la matriz de pérdidas
$VaR(\text{var})$:	Value at risk (valor de riesgo)
u	:	Umbral
U	:	Inversa de $\frac{1}{1-F}$ (usualmente continua a la izquierda)
x^*	:	$\sup \{x : F(x) < 1\} = U(\infty)$
$*x$:	$\inf \{x : F(x) > 0\}$
\hat{x}_p	:	Cuantil estimado
Y_i	:	Exceso sobre un umbral
μ	:	Media de la población
σ	:	Desviación estándar de la población

INTRODUCCIÓN

A partir del 2004 el Nuevo Acuerdo de Basilea emitido por el Banco de Pagos Internacionales obliga a las instituciones financieras a identificar, cuantificar y controlar el riesgo operacional con el fin de controlar los procesos, estimar los requerimientos de capital, satisfacer las exigencias de las entidades regulatorias, disminuir las pérdidas potenciales y aumentar el valor de la empresa.

El acuerdo de Basilea define el riesgo operativo como la pérdida potencial de una entidad financiera por fallas o deficiencias en los sistemas internos, en los procesos, en las personas, o en algunos factores externos. Se incluye el riesgo legal, pero se excluye el riesgo de reputación. La norma colombiana, de la Superintendencia Financiera Colombiana, incluye además el riesgo de reputación.

El riesgo legal es la pérdida posible por sanciones o indemnizaciones como consecuencia del incumplimiento de normas, regulaciones u obligaciones contractuales o por fallas en los contratos o en las transacciones. El riesgo de reputación corresponde a pérdidas por desprestigio, mala imagen o publicidad negativa.

Históricamente, el trabajo en problemas de valor extremo se remota a 1709 cuando Nicolas Bernoulli discutió la distancia promedio más grande de un origen dado a n puntos que se extienden al azar en una línea recta de una longitud t fija.

La teoría del valor extremo tiene su origen principalmente en la necesidad de los astrónomos en aceptar o rechazar observaciones periféricas. Los artículos más antiguos son los de Fuller (1914) y Griffith (1920) ambos especializados en campos de aplicación y en métodos de análisis matemático. Un desarrollo sistemático de la teoría general puede considerarse tuvo su comienzo con el artículo de Von Bortkiewicz (1922) que se concentra en la distribución del rango de muestras aleatorias de una distribución normal. La importancia del artículo de Bortkiewicz es por el hecho de haber introducido claramente por primera vez el concepto de distribución del valor más grande. (En su libro clásico, E. J. Gumbel (1958) dedi-

ca un capítulo a la memoria de L. von Bortkiewicz. En el siguiente año von Mises (1923) evalúa el valor esperado de esta distribución, y Dodd (1923) calcula su mediana, también discute algunas distribuciones no normales. De relevancia más directa es el artículo de Frechet (1927) en el cual se consideran las distribuciones asintóticas de los valores más grandes. En los siguientes años Fisher y Tippet (1928) publicaron resultados de una investigación independiente dentro del mismo problema. Mientras Frechet (1927) ha identificado una posible distribución límite para los estadísticos de orden superior, Fischer y Tippet (1928) muestran que las distribuciones límites de extremos pueden ser solamente de tres tipos. Tippet (1925) había estudiado anteriormente la función de distribución acumulada exacta y los momentos de los estadísticos de orden superior y del rango muestral de una población normal. Von Mises (1936) presenta algunas simples y usuales condiciones suficientes para la convergencia débil de los estadísticos de orden superior para cada una de los tres tipos de distribuciones límite dadas anteriormente por Fischer y Tippet (1928). En 1943 Gnedenko presentó un fundamento riguroso para la teoría del valor extremo y suministro condiciones necesarias y suficientes para la convergencia débil de los estadísticos de orden extremo.

Mejzler (1949), Marcus y Pinsky (1969) (inconsciente del resultado de Mejzler) y de Haan (1970) (1971) refinaron el trabajo de Gnedenko. Un importante pero muy descuidado trabajo de Juncosa (1949) extendió los resultados de Gnedenko al caso de variables aleatorias no necesariamente idénticamente distribuidas e independientes. Aunque de fuerte interés teórico los resultados de Jucosa no parecen tener mucha utilidad práctica. El hecho de que las distribuciones asintóticas de una naturaleza muy general puedan darse no proporciona mucha guía para aplicaciones prácticas.

Los desarrollos teóricos de 1920 a mitad de 1930 fue seguidos al final del 1930 y 1940 por una gran cantidad de de artículos concentrados en las aplicaciones prácticas de los estadísticos de valor extremo en las distribuciones del tiempo de vida humana, emisiones radiactivas (Gumbel (1937 a,b) resistencia de materiales (Weibull 1939) análisis de flujo (Gumbel, 1941, 1944, 1945 1949 a), Rantz y Riggs (1949)) análisis sísmico (Nordquist (1945) , análisis de lluvias (Potter (1949) para mencionar unos pocos ejemplos. Desde el punto de vista de la aplicación Gumbel, hizo varias contribuciones significativas al análisis de valor extremo; muchas de ellas están detalladas en su largo libro de los estadísticos de extremos (Gumbel (1958).

Gumbel fue el primero en llamar la atención a ingenieros y estadísticos de las posibles aplicaciones de la teoría del valor extremo de ciertas distribuciones las cuales habían sido tratadas empíricamente. El primer tipo de problema tratado de esta manera en los Estados Unidos han sido los fenómenos meteorológicos, flujos de corrientes anuales, máxima precipitación., etc. Esto ocurrió en 1941.

En esencia todos los modelos propuestos en el estudio de las fracturas tuvieron como punto de arranque la teoría de Griffith's , la cual estableció la diferencia entre el esfuerzo calculado a los materiales y aquellas observaciones reales reside en el hecho de que existen fallas en el cuerpo que lo debilitan.

El primer escritor que realizó la conexión entre el esfuerzo de un espécimen y la distribución de valores extremos parece ser F. T. Peirce (1926) de la asociación Británica de la industria del algodón. La aplicación de las mismas ideas esenciales para el estudio del esfuerzo de materiales fue conducido por el bien conocido físico Sueco e ingeniero W. Weibull (1939). Los Físicos Rusos, Frenkel y Kontorova (1943) fueron los siguientes en estudiar estos problemas, otra publicación importante no despreciable relacionada con las distribuciones del análisis de valor extremo de esfuerzos factibles del caucho se debe a S. Kase (1953).

Hay varios excelentes libros que se concentran en la teoría asintótica de valor extremo y sus aplicaciones estadísticas. Se citan los siguientes. David (1981) y Arnold, Balakrishnan, y Nagaraja (1992) suministran una explicación compacta de la teoría asintótica de extremos; Galambos (1978, 1987), Resnik (1987), Leadbetter, Lingren, y Rootzen (1983) presentan un tratamiento de este tópico. Reiss (1989) discutió varios conceptos de convergencia asociados con extremos (y también con estadísticos de orden) Castillo (1988) tiene actualizaciones exitosas de Gumbel (1958) y presenta algunas aplicaciones estadísticas de la teoría del valor extremo con énfasis en ingeniería. Harter (1978) preparo una bibliografía autorizada sobre teoría del valor extremo la cual es todavía de gran valor científico. Beirlant, Teugels y Vynckier (1996) suministra un análisis práctico de valores extremos con énfasis en aplicaciones actuariales.

Para el desarrollo a nivel mundial de estudios de la modelación y cuantificación del riesgo operacional a partir del método de distribución de pérdidas agregadas (LDA) ver la introducción de la tesis del docente Luis Ceferino Franco Arbeláez Análisis y comparación de alternativas para cuantificar el riesgo operacional.

En este trabajo se desarrollan dos métodos de medición avanzada (AMA) para

cuantificar el riesgo operativo: El método de distribución de pérdidas agregadas (LDA: Loss Distribution Approach) y la Teoría del Valor Extremo.

Para el método LDA se usa una de las alternativas de modelación y cuantificación la Simulación de Montecarlo del cual se implementa un algoritmo en MatLab cuyo código se anexa en el Apéndice 3 y para la teoría del valor extremo el método Picos sobre un Umbral. Para el desarrollo de estas técnicas se usa una base de datos de riesgo operacional extraída de la Superintendencia Financiera en cuanto a la severidad de los eventos y de una Institución Financiera en cuanto a la frecuencia de los mismos. Al final se hacen comparaciones y se sacan conclusiones sobre el desempeño de ambos métodos.

CAPÍTULO 1

RIESGO Y RIESGO OPERATIVO

1.1 Contextualización del riesgo y el riesgo operativo

Se define riesgo operacional u operativo como la pérdida potencial de una entidad por fallas o deficiencias en los sistemas internos, en los procesos, en las personas, o en algunos factores externos. Se incluye el riesgo legal, pero se excluye el riesgo de reputación. La norma colombiana, de la Superintendencia financiera Colombiana, incluye además el riesgo reputación.

El riesgo legal es la pérdida posible por sanciones o indemnizaciones como consecuencia del incumplimiento de normas, regulaciones u obligaciones contractuales o por fallas en los contratos o en las transacciones. El riesgo de reputación corresponde a pérdidas posibles por desprestigio, mala imagen o publicidad negativa.

Existen varias maneras de medir el riesgo operativo: a través del ajuste de una función de distribución probabilidad a las pérdidas agregadas y luego estimar los cuantiles ($q_{1-\alpha}$) o los parámetros de la distribución tal como la desviación estándar (σ).

Las medidas más modernas del riesgo son cantidades estadísticas que describen las distribuciones de pérdida condicionales o incondicionales sobre un horizonte determinado de tiempo: El valor en riesgo, el Expected Shortfall y la varianza.

1.2 Valor en Riesgo

El valor en riesgo (VaR) es probablemente la medida del riesgo más ampliamente usada en instituciones financieras

Definición 1.1 (Valor en Riesgo). Se define como la máxima pérdida posible para un nivel de confianza y un periodo de tiempo.

Se calcula en base al $(1 - \alpha)$ % cuantil ($q_{1-\alpha}$). El VaR es el menos malo de los $(1 - \alpha)100$ % peores casos de la distribución de pérdidas agregadas.

Dado un nivel de confianza $1 - \alpha$ con $\alpha \in (0,1)$ el VaR a un nivel de confianza α esta dado por el número más pequeño I tal que la probabilidad de que las perdidas excedan L no exceden I no es mayor que α formalmente

$$\text{VaR}_{1-\alpha} = \inf \{I \in \mathbb{R} : P(L > I) \leq \alpha\} = \inf \{I \in \mathbb{R} : F_L(I) \geq 1 - \alpha\}$$

En términos probabilísticos, el VaR es luego simplemente un cuantil de la distribución de pérdida.

Siempre que se habla del VaR se asume sobre un periodo de tiempo Δ que para riesgo de crédito y operativo es usualmente de un año.

Si se denota por μ la media de la distribución de pérdida. Algunas veces el estadístico $\text{VaR}_{1-\alpha}^{\text{media}} = \text{VaR}_{1-\alpha} - \mu$ es usado como el capital adecuado en vez del VaR ordinario este valor determina el capital económico necesario como un amortiguador contra las pérdidas inesperadas.

1.3 Calculo del VaR de forma paramétrica

Si la distribución de pérdidas sigue una distribución normal el VaR para estas pérdidas está dado por:

$$P[L \geq \text{VaR}] = P\left[\frac{L - \mu}{\sigma} \geq \frac{\text{VaR} - \mu}{\sigma}\right] = \alpha$$
$$\frac{\text{VaR} - \mu}{\sigma} = \Phi^{-1}(1 - \alpha) = Z_{1-\alpha} \implies \text{VaR} = \mu + \sigma Z_{1-\alpha}$$

Donde $\Phi^{-1}(\cdot)$ es la función de distribución acumulada y $Z_{1-\alpha}$ esimo cuantil de una distribución normal estándar.

De una manera más general el VaR puede ser calculado a través de la función de distribución acumulada inversa, la cual identifica el cuantil asociado a una probabilidad dada una distribución:

$$\text{VaR}_{1-\alpha} = F_X^{-1}(1 - \alpha)$$

Donde $F_X^{-1}(1 - \alpha)$ es la función inversa de la función de distribución acumulada F_X y está definida como

$$F_X^{-1}(1 - \alpha) = \inf \{X \in \mathbb{R} : F_X(X) \geq 1 - \alpha\}$$

La solución analítica de esta ecuación no siempre es posible.

El Expect Shortfall está estrechamente relacionado con el VaR

Definición 1.2 (Expected shortfall). Para una pérdida L con $E(|L|) < \infty$ y una función de distribución acumulada F_L el Expected Shortfall a un nivel de confianza $1 - \alpha$ con $\alpha \in (0, 1)$ está definido como

$$ES_{1-\alpha} = \frac{1}{\alpha} \int_{1-\alpha}^1 q_u(F_L) du$$

donde $q_u(F_L) = F_L^{\leftarrow}(u)$ es la función cuantil de F_L .

El Expected Shortfall es luego relacionado con el VaR

$$ES_{1-\alpha} = \frac{1}{\alpha} \int_{1-\alpha}^1 \text{VaR}_u(L) du$$

1.4 Desviación estándar

La varianza de una variable aleatoria X está definida como:

$$V(X) = \sigma^2 = E(X^2) - [E(X)]^2$$

La desviación estándar mide la volatilidad de la variable objeto de estudio.

En caso de varias líneas de negocio y eventos para cada una se calcula la varianza y covarianza y se puede obtener el valor esperado y la varianza de las pérdidas agregadas:

$$\begin{aligned} E(L) &= w' \mu \\ V(L) &= w' \Sigma w \end{aligned}$$

Donde $w' = [w_1, w_2, \dots, w_k]$ son las pérdidas de cada una de las líneas de negocio o de cada uno de los eventos de pérdidas operacionales, y $\mu' = [\mu_1, \mu_2, \dots, \mu_k]$ es el vector de las pérdidas esperadas y Σ es la matriz de varianza y covarianza de las pérdidas agregadas de cada uno de los eventos o líneas de negocio.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1k} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2k} \\ \vdots & \vdots & & \vdots \\ \sigma_{k1} & \sigma_{k2} & \dots & \sigma_k^2 \end{bmatrix}$$

La desviación estándar de todas las pérdidas agregadas de la línea de negocios o de los eventos de riesgo operativo es:

$$\sigma = \sqrt{V(L)}$$

Existen tres aspectos relevantes en la administración de riesgos operativos. El primero consiste en la asignación de capital para hacer frente a eventos relacionados con fallas operativas. El segundo toma en cuenta la supervisión y control para evitar que se presenten dichas fallas. El tercero considera los modelos y métodos utilizados para cuantificar el riesgo operativo. Este trabajo se ocupará del tercer aspecto, acá abordaremos dos Métodos: El que surge de la Teoría del Valor Extremo y el Método de distribución de pérdidas agregadas.

El Comité de Basilea ha presentado dos acuerdos llamados Basilea I o Acuerdo de Capital y Basilea II. El primero determinó los primeros parámetros regulatorios sobre la Gestión del riesgo crediticio de los bancos comerciales, el comité definió que la metodología sería la cobertura del riesgo a través de un capital mínimo a mantenerse. Ya en este acuerdo se esclareció que algunos bancos habían logrado sofisticados desarrollos en materia de medición del riesgo. Esto condujo a que el Comité autorizara que los bancos tendrían la posibilidad de implementar sus propios modelos de riesgo para calcular el capital mínimo requerido.

En cuanto al riesgo operacional se establecen tres métodos determinísticos para el cálculo del capital necesario:

Indicador Básico: Asigna el capital mínimo a partir de los ingresos brutos de la entidad.

Método estándar: Asigna capital mínimo por línea de negocio en base a los ingresos brutos de cada una de las líneas.

Método de Medición avanzada: Permite la definición de líneas de negocio por parte de la entidad y asignar el capital mínimo a partir de modelos estadísticos de estimación o indicadores de gestión significativos. Los dos métodos que se van a implementar corresponden a esta categoría.

La implementación de un modelo matemático implica identificar las variables objeto de estudio y todos los aspectos que implica construirlo en nuestro caso el primer aspecto que debe quedar claro es el concepto de factores de riesgo. Se entiende por factores de riesgo las fuentes generadoras de riesgos operativos que pueden generar o no pérdidas. Para el riesgo operativo son factores de riesgo el recurso humano, la tecnología, la infraestructura y los acontecimientos.

Los factores de riesgo se clasifican en internos y externos.

Los factores internos pueden ser:

Recurso humano: Es el conjunto de personas vinculadas directamente o indirectamente con la ejecución de los procesos de la entidad.

Procesos: Es el conjunto interrelacionado de actividades para la transformación de elementos de entrada en productos o servicios, para satisfacer una necesidad.

Tecnología: Es el conjunto de herramientas empleadas para soportar los procesos de la entidad. Incluye: Hardware, software y telecomunicaciones.

Infraestructura: Es el conjunto de elementos de apoyo para el funcionamiento de una organización. Entre otros se incluye: edificios, espacios de trabajo, almacenamiento Y transporte.

Los factores externos pueden ser:

Son situaciones asociadas a la fuerza de la naturaleza u ocasionadas por terceros, que escapan en cuanto a su causa y origen al control de la entidad.

Hay otros dos aspectos fundamentales en los modelos del riesgo operativo que se deben tener en cuenta: los eventos y las pérdidas que ocasionan los eventos.

Los diferentes eventos de riesgo operacional pueden ser estudiados en términos de su frecuencia (el número de eventos que producen pérdidas en un cierto intervalo de tiempo) y su severidad (el impacto del evento en término de pérdidas económicas).

Evento:

Incidente o situación que ocurre en un lugar particular durante un intervalo de tiempo determinado.

Pérdidas:

Cuantificación económica de la ocurrencia de un evento de riesgo operativo, así como los gastos derivados de su atención.

Evento de pérdida:

Son aquellos incidentes que generan pérdidas de riesgo operativo a las entidades.

Los riesgos operativos se clasifican de la siguiente manera:

Fraude interno:

Actos que de forma intencionada buscan defraudar o apropiarse indebidamente de activos de la entidad o incumplir normas o leyes, en los está implicado, al menos, un empleado o administrador de la entidad.

Fraude externo:

Actos, realizados por una persona externa a la entidad, que buscan defraudar, apropiarse indebidamente de activos de la misma o incumplir normas o leyes.

Relaciones laborales:

Actos que son incompatibles con la legislación laboral, con los acuerdos internos de trabajo, y en general, la legislación vigente sobre la materia.

Clientes:

Fallas negligentes o involuntarias de las obligaciones frente a los clientes y que impiden satisfacer una obligación.

Fallas tecnológicas:

Pérdidas derivadas por fallas tecnológicas.

Ejecución y administración de procesos

Pérdidas derivadas de errores en la ejecución y administración de los procesos. El estudio del riesgo operativo también requiere construir una matriz donde se incluyan las siete líneas de negocios y los siete eventos de riesgo y en donde a cada celda se le asigne el capital mínimo de reserva que debe ser asignado para protegerse de los riesgos inesperados.

Tabla 1.1. Matriz de pérdidas

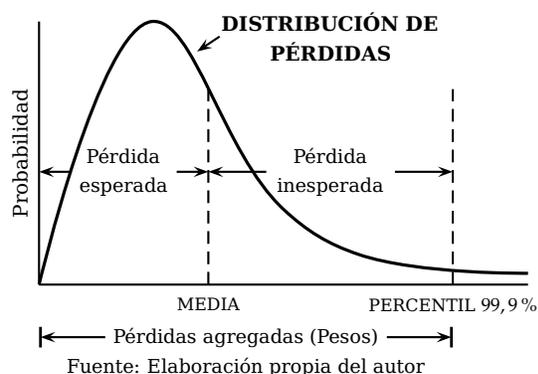
		Líneas de negocio				Total
		L1	L2	...	L8	
Eventos de Riesgo	Fraude interno					
	Fraude externo					
	Relaciones laborales					
	Clientes			$VaR(i,j)$		$\sum VaR(E_i)$
	Daños a activos físicos					
	Fallas tecnológicas					
	Ejecución y administración de procesos					
Total				$\sum VaR(L_j)$		$\sum_{i=1}^7 \sum_{j=1}^8 VaR(i,j)$

La matriz anterior muestra en las filas los 7 tipos de riesgo y en las columnas las 8 líneas de negocio (Finanzas Corporativas, Negociación y Ventas, Banca Minorista, Banca Comercial, Pagos y Liquidación, Servicios de Agencia, Administración de Activos e Intermediación Minorista)² en la última columna se obtiene la suma de todos los VaR del tipo de riesgo o evento E_i y por columna la sumatoria del VaR de la línea de negocio j y la última celda la sumatoria de todas las celdas dándonos el VaR total para las 56 celdas.

La suma total debe incluir las pérdidas esperadas y las pérdidas inesperadas debidas al riesgo operativo.

La gráfica siguiente muestra la curva sobre la distribución de pérdidas donde se observan las pérdidas esperadas y las pérdidas inesperadas.

Figura 1.1. Distribución de pérdidas agregadas



²Esta es la matriz propuesta por Basilea pero cada firma tiene la libertad de considerar sus propias líneas y eventos de tal forma de adecuarla a la estructura y necesidad de la organización.

CAPÍTULO 2

TEORÍA Y MODELOS DE VALORES EXTREMOS

2.1 Teoría Básica de Valor Extremo

2.1.1 Introducción

Inicialmente se presentan dos teoremas que dan varias relaciones de equivalencia las cuales son la base para probar que una función de distribución de las comunes que se conocen en estadística pertenece al dominio de atracción de una distribución límite: la Distribución de valor Extremo o Distribución Generalizada de Pareto.

El teorema de Fisher, Tippett y Gnedenko de los extremos de una muestra, es fundamental ya que a partir de él se reconoce la distribución límite de valores extremos con el parámetro de forma que es el esencial dentro de todo el proceso.

Posteriormente se presentan varios teoremas sobre las condiciones necesarias y suficientes para que una distribución determinada pertenezca al dominio de atracción de Gumbel, Fréchet o Weibull.

Los teoremas de Von Mises que da luego condiciones suficientes para el proceso anterior y el cómo se calculan las constantes que estabilizan las distribuciones a una distribución de valor extremo que no son únicas pero que se requieren para que la distribución en el límite no se degeneren.

Más adelante se presentan teoremas sobre la distribución generalizada de Pareto que es fundamental en el ajuste de los valores de variables aleatorias que están por encima de un umbral alto.

Finalmente se presenta el método para estimar el parámetro de forma de las distribuciones de valor extremo denominado el Método de Hill para cuando el parámetro de forma toma el valor positivo ($\xi > 0$). Se destaca este método porque

en el área de la finanzas este parámetro es un valor real positivo. Del parámetro de forma se analiza específicamente propiedades como consistencia débil y la normalidad asintótica, con el fin de tener elementos para formular pruebas e intervalos de confianza de este parámetro que es el que define el tipo de cola de la distribución.

La teoría sobre valor extremo descrita en esta tesis es bastante reciente: Solamente en 1980s los contornos de la teoría estadística tomaron forma. La teoría de valor extremo probabilística en una dimensión fue desarrollada por M Frechet (1927), R. Fisher y L. Tippett (1928) y R von Mises (1936) y culminó en el trabajo de B. Gnedenko (1943). La teoría estadística fue iniciada por J. Pickands III (1975).

La teoría asintótica de los extremos de la muestra ha sido desarrollada en paralelo con la teoría central del límite y en efecto las dos teorías tienen en común que mientras en la teoría central del límite la distribución a la cual converge la media es a una distribución normal estándar después de restarle la media de la población y dividir por la varianza de la media que es igual a $\frac{\sigma}{\sqrt{n}}$ valor que depende de n el máximo o el mínimo es decir los valores extremos.

Sea $X_1, X_2, X_3 \dots$ variables aleatorias independientes e idénticamente distribuidas. La teoría central del límite está concentrada con el comportamiento límite de las sumas parciales³ $S_n = X_1 + X_2 + \dots + X_n$ cuando $n \rightarrow \infty$ mientras la teoría de los extremos de una muestra está concentrada con el comportamiento límite del máximo y mínimo, extremos muestrales $\max(X_1, \dots, X_n)$ o $\min(X_1, \dots, X_n)$ cuando $n \rightarrow \infty$.

La teoría de valores extremos es el área de la estadística dedicada a desarrollar modelos y técnicas para estimar el comportamiento de eventos inusuales o raros. Los eventos raros son los que pertenecen a las colas de la distribución y son los que están alejados del grueso de la distribución (de la media o mediana). Otra manera de definir los eventos extremos son aquellos valores que están por encima de un valor denominado umbral, pero también se puede definir como el máximo o el mínimo de una variable en un cierto período.

Uno de los principales teoremas basados en la teoría asintótica un tipo de teorema central de límite versión valores extremos derivado en principio por Fisher y

³Para el caso del teorema central del límite con respecto a \bar{X} , $a_n = \frac{\sigma}{\sqrt{n}}$ y $b_n = \mu$

Tippet (1928) de manera heurística y luego desde un punto de vista más riguroso por Genedenko (1943) establecieron que hay solamente tres tipos de distribuciones (Distribución tipo Gumbel, Tipo Frechet y tipo Weibull) que se alcanzan como distribuciones límite de los máximos y mínimos de algún proceso aleatorio sobre intervalos de tiempo fijos.

Durante los años cincuenta Von Misses (1954) y Jenkisson (1955) unificaron las tres distribuciones en una sola denominada distribución generalizada de valores extremos (GEV) la cual depende de tres parámetros. Esta unificación de las tres distribuciones en una sola permitió simplificar los procedimientos de modelaje.

Otro desarrollo que se inició a partir de la década del setenta por Pickands (1975) y que se continuo en las décadas de los ochenta y noventa se basó en los métodos basados en los excesos por encima de un umbral en lugar del máximo o el mínimo de un periodo pero acá se origina otro tipo de distribución conocida como la distribución generalizada de Pareto la cual fue introducida por Pickands (1975), su aplicación se hace en el análisis de eventos extremos en el modelamiento de grandes derechos en seguros, en distribuciones de tiempos de falla en estudios de confiabilidad y en cualquier situación en la cual se requiera debido a la presencia de colas pesadas o colas ligeras.

Otro de los avances en teoría de Valor Extremo que se ha producido en los últimos años es extender esta teoría a procesos no estacionarios y dependientes además de desarrollar la inferencia para extremos multivariados.

2.1.2 Formulación de la relación límite para el máximo y el mínimo

Las distribuciones de valores extremos surgen formalmente como distribuciones límite para el máximo y el mínimo de una sucesión de variables aleatorias. Los conceptos fundamentales en los que se basa la teoría de valores extremos son los modos de convergencia: Convergencia en distribución, convergencia en probabilidad, convergencia casi segura y convergencia de funciones inversas generalizadas. Por ahora definamos lo que es convergencia en probabilidad.

Definición 2.1 (Convergencia en probabilidad). Se dice que una sucesión $\{X_n\}$ converge en probabilidad a la variable aleatoria X si para $\forall \varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0$$

lo cual equivale a que $\forall \varepsilon > 0 P(|X_n - X| > \varepsilon) \rightarrow 0$, cuando $n \rightarrow \infty$ la convergencia

en probabilidad se denota por $X_n \xrightarrow{P} X$.

Supongamos que son variables aleatorias independientes e idénticamente distribuidas con función de densidad de probabilidad

$$p_{x_j}(x) = f(x), \quad j = 1, 2, \dots, n$$

La función de distribución acumulada F , y X^* su punto final superior esto es $x^* = \{x : F(x) < 1\}$ el cual puede ser infinito entonces $\max(X_1, \dots, X_n) \xrightarrow{P} x^*$, $n \rightarrow \infty$ donde

$$F(x) = P(X_j \leq x) \quad \forall_{j,x} \quad (2.1)$$

Luego para el máximo

$$X_n = M_n = \max\{X_1, X_2, \dots\} \quad (2.2)$$

La función de distribución acumulada, en teoría, viene dada por

$$\begin{aligned} P(M_n \leq x) &= P(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) \\ &= [P(X_i \leq x)]^n = F^n(x) \end{aligned} \quad (2.3)$$

Donde $F(x) = \int_{-\infty}^x f(t) dt$

Para resolver el problema del desconocimiento de la función de distribución acumulada se pueden implementar dos soluciones la primera, es ajustar una distribución con los datos observados y así estimar aproximadamente el valor de F pero si hay algún error este se hará muy grande ya que F esta elevado a la n .

La otra solución es aceptar que F es desconocido y tratar de obtener la distribución aproximada que puede tener $F^n(x)$ y que solo se puede lograr a partir de datos extremos con una teoría análoga al teorema central del límite.

Pero hay que tener en cuenta que se puede tener el siguiente caso: en el límite la distribución puede ser degenerada.

Definición 2.2 (Distribución Degenerada). Una variable aleatoria X se dice que tiene una distribución degenerada en un punto h si su función de probabilidad es

$$p_x(x) = \begin{cases} 1 & \text{si } x = h \\ 0 & \text{si } x < h \end{cases}$$

La función de distribución de una variable aleatoria degenerada es:

$$F(x) = \begin{cases} 0 & \text{si } x < h \\ 1 & \text{si } x \geq h \end{cases}$$

Por ejemplo para cualquier $x < x^*$ donde x^* es el punto superior de F , es decir x^* es el valor más pequeño de x tal que se cumple que $\lim_{n \rightarrow \infty} F^n(x) = 0$ por lo que la función de distribución de M_n degenera en un punto de masas en x^* .

Cuando n tiende a infinito es claro que para cualquier valor fijo de x

$$\lim_{n \rightarrow \infty} F_{x_n'}(x) = \begin{cases} 1 & \text{si } F(x) = 1 \\ 0 & \text{si } F(x) < 1 \end{cases}$$

El cual es una función degenerada. Para evitar esta dificultad, se normaliza y centraliza la variable M_n :

$$M_n^* = \frac{M_n - b_n}{a_n} \quad (2.4)$$

El máximo de la variable aleatoria puede ser estabilizado a través de las sucesiones $\{a_n\}$ y $\{b_n\}$ con $a_n > 0$ y $b_n \in \mathbb{R}$ que corresponden a los parámetros de escala y de localización tales que

$$\begin{aligned} P\left(\frac{M_n - b_n}{a_n} \leq x\right) &= P(M_n \leq a_n x + b_n) \\ &= F^n(a_n x + b_n) \xrightarrow{n \rightarrow \infty} G(x) \Leftrightarrow \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x) \end{aligned} \quad (2.5)$$

La escogencia apropiada de $\{a_n\}$ y $\{b_n\}$ estabiliza la escala y la localización de M_n^* cuando se incrementa n , y así se evita la degeneración de M_n . De esta manera M_n^* converge en distribución a alguna de las familias, Gumbel, Frechet o Weibull.

2.1.3 Formulaciones alternativas de la relación límite

Se pueden obtener las siguientes relaciones equivalentes a la anterior que servirán como base para posteriores demostraciones.

Tomando logaritmos a la derecha y a la izquierda de la relación (2.5) se cumple para todos los puntos donde hay continuidad y para el cual aplicando la regla de L'Hopital

$$\lim_{n \rightarrow \infty} \frac{-\log F(a_n x + b_n)}{1 - F(a_n x + b_n)} = \lim_{n \rightarrow \infty} \frac{\left(x \frac{d a_n}{d n} + \frac{d b_n}{d n}\right) f(a_n x + b_n)}{F(a_n x + b_n)} = \lim_{n \rightarrow \infty} \frac{1}{F(a_n x + b_n)}$$

$$\lim_{n \rightarrow \infty} \frac{-\log F(a_n x + b_n)}{1 - F(a_n x + b_n)} = \frac{\lim_{n \rightarrow \infty} 1}{\lim_{n \rightarrow \infty} F(a_n x + b_n)} = \frac{1}{1} = 1$$

y en efecto (2.6) es equivalente a

$$\lim_{n \rightarrow \infty} n(1 - F(a_n x + b_n)) = -\log G(x) \quad (2.6)$$

o

$$\lim_{n \rightarrow \infty} \frac{1}{n(1 - F(a_n x + b_n))} = \frac{1}{-\log G(x)} \quad (2.7)$$

Teorema 2.1 (Equivalencia de enunciados)

Sean $\{a_n\}$ y $\{b_n\}$ sucesiones de números reales, y G una función de distribución no degenerada.

1. $\forall x \in G(x) \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x)$
2. $\forall x \in G(x)$, continuo y $0 < G(x) < 1$; $a(t) := a_{\llbracket t \rrbracket}$ (con $\llbracket t \rrbracket$ la parte entera de t)

$$\lim_{t \rightarrow \infty} t(1 - F(a(t)x + b(t))) = -\log G(x) \quad (2.8)$$

3. $\forall x > 0$ punto de continuidad de $D(x) = G^{\leftarrow}(e^{-\frac{1}{x}})$, $a(t) := a_{\llbracket t \rrbracket}$, y $b(t) := b_{\llbracket t \rrbracket}$

$$\lim_{t \rightarrow \infty} \frac{U(tx) - b(t)}{a(t)} = D(x), \quad (2.9)$$

entonces 1. \Leftrightarrow 2, 2. \Leftrightarrow 3.

Teorema 2.2 (Fisher y Tippett (1928), Gnedenko (1943))

El teorema de Fisher, Tippett y Gnedenko es un resultado importante por muchos aspectos, es el teorema fundamental de la Teoría de Valor Extremo. Muestra que la función de distribución límite forma una familia de un parámetro de forma simple y explícito aparte de los parámetros de localización y escala.

El teorema que se probará a continuación identifica la clase de distribuciones no degeneradas que puede ocurrir como límite de la relación (2.5).

Sean $\{a_n\}$ y $\{b_n\}$ sucesiones de números reales y $M_n = \{X_1, X_2, \dots\}$ si

$$\begin{aligned} P\left(\frac{M_n - b_n}{a_n} \leq x\right) &= P(M_n \leq a_n x + b_n) \\ &= F^n(a_n x + b_n) \xrightarrow{n \rightarrow \infty} G(x) \Leftrightarrow \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x) \end{aligned}$$

Entonces la distribución de valor extremo es

$$G_{\xi}(x) = G(x, \xi) = \begin{cases} \exp\left(-\left(1 + \xi \frac{x - D(1)}{H'(0)}\right)^{-\frac{1}{\xi}}\right), & 1 + \xi x > 0 \\ \exp\left(-e^{-\frac{x - D(1)}{H'(0)}}\right) & x \in \mathbb{R}, \xi = 0 \end{cases} \quad (2.10)$$

Demostración. Consideremos la distribución límite:

$$D(x) = \lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)}$$

Si 1 es un punto de continuidad de la función $D(x)$ se obtiene

$$D(1) = \lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = 0$$

Sea $E(x) = D(x) - D(1)$ entonces para los puntos donde hay continuidad $x > 0$

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = D(x) - D(1) =: E(x) \quad (2.11)$$

$y > 0$

$$\frac{U(txy) - U(t)}{a(t)} = \frac{U(txy) - U(ty)}{a(ty)} \frac{a(ty)}{at} + \frac{U(ty) - U(t)}{a(t)} \quad (2.12)$$

Afirmamos que $\lim_{n \rightarrow \infty} \frac{U(ty) - U(t)}{a(t)}$ y $\lim_{n \rightarrow \infty} \frac{a(ty)}{a(t)}$ existen.

Supongamos que no.

Entonces hay A_1, A_2, B_1, B_2 con $A_1 \neq A_2$ o $B_1 \neq B_2$, donde B_i son puntos límites de $\frac{U(ty) - U(t)}{a(t)}$ y A_i son puntos de $\lim_{t \rightarrow \infty} \frac{a(ty)}{a(t)}$ cuando $t \rightarrow \infty$

Se encuentra de (2.12) y de las definiciones de A_i y B_i que

$$E(xy) = E(x)A_i + B_i \quad i = 1, 2 \quad (2.13)$$

Para $i = 1$ $E(xy) = E(x)A_1 + B_1$

Para $i = 2$ $E(xy) = E(x)A_2 + B_2$

Restando y factorizando se tiene

$$E(x)(A_1 - A_2) = B_1 - B_2 \quad \forall x > 0$$

ya que E no puede ser constante (G no puede ser degenerada). Se puede concluir que $A_1 = A_2$ y $B_1 = B_2$, por lo tanto $A(y) := \lim_{t \rightarrow \infty} \frac{a(ty)}{a(t)}$ existe para $\forall x > 0$ y para $\forall x, y > 0$

$$E(xy) = E(x)A(y) + E(x)$$

por lo tanto para: $s = \log x$, $t = \log y$ ($x, y \neq 1$), como $H(x) = E(e^x)$

$$H(t + s) = E(e^{t+s}) = E(e^s e^t) = E(e^s)A(e^t) + E(e^t)$$

$E(x)(A_1 - A_2) = B_1 - B_2$ para todo $x > 0$. Ya que E no puede ser constante (recordar que G es no degenerada) nosotros debemos tener que $A_1 = A_2$ y por lo tanto también $B_1 = B_2$. Conclusión:

$$A(y) := \lim_{t \rightarrow \infty} \frac{a(ty)}{a(t)}$$

existe para $y > 0$, y para $x, y > 0$

$$E(xy) = E(x)A(y) + E(y)$$

Por lo tanto para $s := \log x$, $t = \log y$ ($x, y \neq 1$), y $H(x) := E(e^x)$, se tiene

$$\begin{aligned} H(t+s) &= E(e^{t+s}) = E(e^t e^s) = E(e^s e^t) = E(e^s)A(e^t) + Ee^t \\ &= H(s)A(e^t) + H(t) \end{aligned} \quad (2.14)$$

Lo cual puede ser escrito como (ya que $H(0) = 0$)

$$\frac{H(t+s) - H(t)}{s} = \frac{H(s) - H(0)}{s} A(e^t) \quad (2.15)$$

Hay ciertamente una t en la cual H es diferenciable (ya que H es monótona) por lo tanto (2.15) H es diferenciable en todo valor y

$$\begin{aligned} \lim_{s \rightarrow \infty} \frac{H(t+s) - H(t)}{s} &= \lim_{s \rightarrow \infty} \frac{H(s) - H(0)}{s} A(e^t) \\ H'(t) &= H'(0)A(e^t) \end{aligned} \quad (2.16)$$

La ecuación (2.15) puede ser escrita como:

$$H(t+s) = H(s)A(e^t) + H(t) \quad (2.17)$$

(ya que $H(0) = 0$). Definamos: $Q(t) = \frac{H(t)}{H'(0)}$. Note que $H'(0)$ no puede ser cero.

Por (2.16) se tiene $H'(t) = H'(0)A(e^t)$ para $t = 0$ se tiene

$$H'(0) = H'(0)A(e^0) \rightarrow 1 = A(1)$$

H no puede ser constante ya que G es no degenerada. Luego

$$Q(0) = 0 \quad , \quad Q'(0) = 1 \quad , \quad Q(t+s) - Q(t) = Q(s)A(e^t),$$

aplicando (2.14) y por (2.16) Sustrayendo la misma expresión con t y s intercambiados se consigue

$$Q(t) \frac{Q'(s) - 1}{s} = \frac{Q(s)}{s} (Q'(t) - 1)$$

Por lo tanto

$$\lim_{s \rightarrow 0} Q(t) \frac{Q'(s) - 1}{s} = \lim_{s \rightarrow 0} \frac{Q(s)}{s} (Q'(t) - 1) \Rightarrow Q(t)Q''(0) = Q'(t) - 1$$

ya que

$$\lim_{s \rightarrow 0} \frac{Q(s)}{s} = \lim_{s \rightarrow 0} \frac{Q(s) - Q(0)}{s} = Q'(0) = 0$$

ya que

$$Q(t) = \frac{H(t)}{H'(t)} \Rightarrow Q'(t) = \frac{[H'(t)]^2 - H(t)H''(t)}{[H'(t)]^2} \Rightarrow Q'(0) = 1$$

Derivando nuevamente se obtiene:

$$Q'(t)Q''(0) + Q(t)Q'''(0) = Q''(t) \Rightarrow Q''(0)Q(t) = Q''(t)$$

Pero como $(\log Q')'(t) = Q''(0) = \xi \in R$ para todo t . Se sigue que

$$Q(t) = \int_0^t e^{t\xi} ds \quad \text{ya que } Q(0) = 1$$

Esto significa que

$$H(t) = H'(0) \frac{e^{\xi t} - 1}{\xi} \quad \text{y} \quad D(t) = D(1) + H'(0) \frac{t^\xi - 1}{\xi}$$

Por lo tanto

$$D^{\leftarrow}(x) = \left(1 + \xi \frac{x - D(1)}{H'(0)}\right)^{\frac{1}{\xi}} \quad (2.18)$$

Ahora $D(x) = G^{\leftarrow}\left(e^{-\frac{1}{x}}\right)$, y por lo tanto

$$D^{\leftarrow}(x) = \frac{1}{-\log G(x)} \quad (2.19)$$

Combinando (2.18) y (2.19) se obtiene la prueba del teorema.

$$\frac{1}{-\log G(x)} = \left(1 + \xi \frac{x - D(1)}{H'(0)}\right)^{\frac{1}{\xi}}$$

$$G(x, \xi) = \exp\left(-\left(1 + \xi \frac{x - D(1)}{H'(0)}\right)^{-\frac{1}{\xi}}\right)$$

Definición 2.3 (Índice de Valor Extremo). El parámetro ξ es llamado el índice de valor extremo.

La parametrización del teorema 2.2 es debido a von Mises (1936) y Jenkinson (1955). Los tipos de distribución dependen del parámetro ξ así:

- a. Para $\xi > 0$, $G(x; \xi) < 1$ para todo x , es decir el punto final superior de la distribución es infinito. Además, cuando $x \rightarrow \infty$, $1 - G(x) \sim \xi^{-\frac{1}{\xi}} x^{-\frac{1}{\xi}}$, es decir la distribución tiene una cola pesada a la derecha. Y los momentos mayores o iguales a $\frac{1}{\xi}$ no existen.
- b. Para $\xi = 0$ el punto final superior de la distribución es infinito. La distribución es además de colas ligeras: $1 - G_0(x) e^{-x}$ cuando $x \rightarrow \infty$, y todos los momentos existen.
- c. Para $\xi < 0$ el punto final de la distribución es $-\frac{1}{\xi}$ así que tiene una cola corta que verifica $1 - G_\xi(-\xi^{-1} - x) \sim (-\xi x)^{-\frac{1}{\xi}}$, cuando $x \rightarrow 0$.

Una parametrización alternativa es la siguiente:

- a. Para $\xi > 0$ usando $G_\xi\left(\frac{x-1}{\xi}\right)$ y con $\alpha = \frac{1}{\xi} > 0$

$$\begin{aligned} G\left(\frac{x-1}{\xi}, \xi\right) &= \exp\left(-\left(1 + \xi \frac{x-1}{\xi}\right)^{-\frac{1}{\xi}}\right) \\ &= \exp\left(-(1+x-1)^{-\frac{1}{\xi}}\right) = \exp\left(-(x)^{-\alpha}\right) \\ G_2(x) &:= \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} \end{aligned}$$

Este tipo es llamada la distribución del tipo Frechet (Frechet (1927)).

- b. La función de distribución con $\xi = 0$

$$G_1(x) = \exp(-e^{-x}),$$

Para todo x real, es llamada la doble exponencial o distribución de Gumbel.

- c. Para $\xi < 0$ se usa $G_\xi\left(-\frac{1+x}{\xi}\right)$ y con $\alpha = -\frac{1}{\xi} > 0$

$$\begin{aligned} G_3(x) &:= \begin{cases} \exp(-(-x)^\alpha), & x < 0 \\ 1 & x \geq 0 \end{cases} \\ G\left(-\frac{1+x}{\xi}, \xi\right) &= \exp\left(-\left(1 + \xi \frac{-1-x}{\xi}\right)^{-\frac{1}{\xi}}\right) \\ &= \exp\left(-(-x)^{-\frac{1}{\xi}}\right) = \exp(-(-x)^\alpha) \end{aligned}$$

Este tipo es llamada la distribución de Weibull inversa.

Si la relación (2.5) se cumple con $G = G_{\xi}$ para algún $\xi \in R$ se dice que la función de distribución F esta en el dominio de atracción de G_{ξ} y se denota: $F \in D(G_{\xi})$

El teorema 2.2 conduce a la siguiente reformulación del teorema 2.1 y como en el siguiente teorema se requiere la inversa generalizada se definirá previamente:

Definición 2.4 (Ínfimo de un conjunto). Se dice que un conjunto A está acotado inferiormente si y solo si existe un número real c que es menor o igual a cualquier elemento de A . La mayor de las colas inferiores se denomina el ínfimo de A , y se denota $\inf(A)$

Definición 2.5 (Inversa Generalizada de una función monótona). Supongamos que h es una función no decreciente en los reales. La inversa generalizada de h está definida como:

$$h^{\leftarrow}(t) = \inf\{x \in R : h(x) \geq t\}$$

Usaremos la convención de que ínfimo de un conjunto vacío es ∞ .

Definición 2.6 (Función cuantil). La inversa generalizada de la función de distribución

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

Es llamada la función cuantil de la función de distribución. El cuantil $x_p = F^{\leftarrow}(p)$ define el cuantil p de F . Entre los cuantiles se tienen los percentiles.

Teorema 2.3 (Reformulaciones de las equivalencias del teorema 2.1)

Para todo $\xi \in R$ las siguientes afirmaciones son equivalentes:

1. Existen constantes reales $a_n > 0$ y b_n real tal que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_{\xi}(x) = \exp\left(- (1 + \xi x)^{-\frac{1}{\xi}}\right) \quad (2.20)$$

Para todo x con $1 + \xi x > 0$

2. Existe una función positiva tal que para todo $x > 0$

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = D_{\xi}(x) = \frac{x^{\xi} - 1}{\xi} \quad (2.21)$$

Donde para $\xi = 0$ el lado derecho es interpretado como $\log x$

3. Existe una función positiva tal que

$$\lim_{t \rightarrow \infty} t(1 - F(a(t)x + U(t))) = (1 + \zeta x)^{-\frac{1}{\zeta}} \quad (2.22)$$

para todo x con $1 + \zeta x > 0$

4. Existe una función positiva f tal que

$$\lim_{t \rightarrow x'} \frac{1 - F(t + xf(t))}{1 - F(t)} = (1 + \zeta x)^{-\frac{1}{\zeta}} \quad (2.23)$$

para todo x para el cual $1 + \zeta x > 0$, donde $x^* = \sup\{x : F(x) < 1\}$

Demostración. La equivalencia de 2 y 3 siguen del lema 1.1.1 (Texto Extreme Value Theory An Introduction, de Haan y Ferreira) y se aplica este mismo lema a la relación (2.7). Sea la función U la inversa continua a la izquierda de $\frac{1}{1-F}$. Se debe tener en cuenta que $U(t)$ está definida para $t > 1$. Se sigue (2.7) es equivalente a

$$\lim_{n \rightarrow \infty} \frac{U(nx) - b_n}{a_n} = G^{\leftarrow} \left(e^{-\frac{1}{x}} \right) := D(x)$$

para cada x positivo. Esto es alentador ya que la relación anterior es más simple que (2.7).

Así que es suficiente probar que la relación anterior implica 3. Sea x un punto de continuidad de

$$D_{\zeta}(x) = \lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^{\zeta} - 1}{\zeta}.$$

Para $t \geq 1$,

$$\frac{U(\llbracket t \rrbracket x - b_{\llbracket t \rrbracket})}{a_{\llbracket t \rrbracket}} \leq \frac{U(tx) - b_{\llbracket t \rrbracket}}{a_{\llbracket t \rrbracket}} \leq \frac{U\left(\llbracket t \rrbracket x \left(1 + \frac{1}{\llbracket t \rrbracket}\right)\right) - b_{\llbracket t \rrbracket}}{a_{\llbracket t \rrbracket}}$$

el lado derecho es eventualmente menor $D(x')$ para cualquier punto continuo $x' > x$ con $D(x') > D(x)$. Ya que D es continuo en x se tiene

$$\lim_{t \rightarrow \infty} \frac{U(tx) - b_{\llbracket t \rrbracket}}{a_{\llbracket t \rrbracket}} = D(x)$$

esto es 3 del teorema (2.8).

Se probará ahora que 2 implica 4.

Se puede ver que para $\varepsilon > 0$

$$g(h^{\leftarrow}(t) - \varepsilon) \leq t \leq g(h^{\leftarrow}(t) + \varepsilon),$$

donde g es una función no decreciente y h^{\leftarrow} es la inversa continua a la derecha. Se sigue que

$$\begin{aligned} \frac{(1 - \varepsilon)^{\zeta} - 1}{\zeta} &\leftarrow \frac{U\left(\frac{1 - \varepsilon}{1 - F(t)}\right) - U\left(\frac{1}{1 - F(t)}\right)}{a\left(\frac{1}{1 - F(t)}\right)} < \frac{t - U\left(\frac{1}{1 - F(t)}\right)}{a\left(\frac{1}{1 - F(t)}\right)} \\ &< \frac{U\left(\frac{1 + \varepsilon}{1 - F(t)}\right) - U\left(\frac{1}{1 - F(t)}\right)}{a\left(\frac{1}{1 - F(t)}\right)} \rightarrow \frac{(1 + \varepsilon)^{\zeta}}{\zeta} \end{aligned}$$

$t \rightarrow x^*$ Luego,

$$\lim_{t \rightarrow x^*} \frac{t - U\left(\frac{1}{1 - F(t)}\right)}{a\left(\frac{1}{1 - F(t)}\right)} = 0$$

Por lo tanto por 2 para todo $x > 0$

$$\lim_{t \rightarrow x^*} \frac{t - U\left(\frac{1}{1 - F(t)}\right)}{a\left(\frac{1}{1 - F(t)}\right)} = \frac{x^{\zeta} - 1}{\zeta}$$

y por el lemma 1.1.1 (Texto Extreme Value Theory An Introduction, de Haan y Ferreira)

$$\lim_{t \rightarrow x^*} \frac{1 - F(t)}{1 - F\left(t + xa\left(\frac{1}{1 - F(t)}\right)\right)} = (1 + \zeta x)^{\frac{1}{\zeta}}$$

Es decir 4 se cumple. Lo mismo 4 implica 2.

El rango completo de posibles distribuciones límite para M_n^* esta dado por el teorema de los tipos de Extremos (Fisher y Tipper (1928) Gnedenko, (1943)).

Definición 2.7 (Distribución Max-estable). Sea X_1, X_2, \dots variables aleatorias independientes e idénticamente distribuidas con una función de distribución F .

La función de distribución F es llamada max-estable si para alguna selección de constantes $a_n > 0$, y b_n real

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) = P(X_1 \leq x)$$

Para toda x y $n = 1, 2, \dots$. La clase de distribuciones max estables o estable en los máximos es la misma clase de distribuciones de valor extremo.

Una distribución para la cual el máximo muestral re-escalado, posee la misma distribución que la variable aleatoria inicial.

Teorema 2.4 (Distribución Max-Estable)

Una distribución es max-estable si y solo si es una de las distribuciones de valores extremos.

Demostración.

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = G(x) \Leftrightarrow \lim_{n \rightarrow \infty} P\left(\frac{M_{nk} - b_{nk}}{a_{nk}} \leq x\right) = G(x) \quad (2.24)$$

$$\Leftrightarrow P\left(\frac{M_{nk} - b_{nk}}{a_{nk}} \leq x\right) = \left[P\left(\frac{M_n - b_n}{a_n} \leq x\right) \right]^k \quad (2.25)$$

En consecuencia de (2.24) y (2.25)

$$\begin{aligned} \Leftrightarrow P(M_{nk} \leq x) &\approx G\left(\frac{x - b_{nk}}{a_{nk}}\right) \\ \Leftrightarrow P(M_{nk} \leq x) &\approx G^k\left(\frac{x - b_{nk}}{a_{nk}}\right) \end{aligned}$$

Por lo tanto

$$\Leftrightarrow G(x) = G^k\left(\tilde{a}_k x + \tilde{b}_k\right)$$

2.1.4 Distribución Generalizada de Valores Extremos

Von Mises (1954) y Jenkinson (1955) derivaron cada uno por su lado, la distribución generalizada de extremos que unificó las tres distribuciones limites de valores extremos establecidas en el teorema 2.2.

Estas tres distribuciones limites tienen comportamientos diferentes que están relacionados con las distintas formas que pueda tener la cola de la distribución

F de los X_i , y esto se puede precisar más considerando el comportamiento de la distribución límite G en el punto final superior x^* . Así para la distribución Weibull, x^* es finito, mientras que la distribución Frechet y Gumbel es infinito, $x^* = \infty^+$.

Sin embargo, la función de densidad para G decae exponencialmente para el caso de la distribución de Gumbel, y en forma polinomial para la distribución Frechet.

En las primeras aplicaciones de la teoría de valores extremos, era usual adoptar una de estas tres familias y luego estimar los parámetros relevantes de la distribución escogida. Pero esta metodología tiene dos debilidades, primero se necesita una técnica adecuada para escoger cual familia de distribución es la más apropiada para manejar los datos; segundo, una vez tomada la decisión, la inferencia asume que no hay incertidumbre en la selección de la familia de distribución más apropiada.

Los problemas anteriores se resuelvan al combinar en una sola distribución las tres distribuciones de valores extremos conocida como la distribución generalizada de valores extremos (GEV) dada por la siguiente expresión:

$$G(x, \mu, \xi, \sigma) = \exp \left\{ - \left[\xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} \quad (2.26)$$

Definida sobre el conjunto

$$\left\{ x : 1 + \xi \left(\frac{x - \mu}{\sigma} \right) > 0 \right\}$$

Porque de lo contrario, G toma el valor de 0 o 1 y entonces es una distribución degenerada.

Los parámetros satisfacen las condiciones

- $-\infty < \mu < \infty$
- $\sigma > 0$
- $-\infty < \xi < \infty$

Donde μ es el parámetro de localización, σ es el parámetro de escala, y ξ es el parámetro de forma, que es el más importante porque determina el comportamiento de la cola de distribución de los datos X_i .

1. Si $\xi > 0$ la cola de la distribución tiene un comportamiento tipo Frechet con $\alpha = \frac{1}{\xi}$
2. Si $\xi < 0$ la cola de la distribución tiene un comportamiento del tipo Weibull con $\alpha = -\frac{1}{\xi}$
3. Si $\xi \rightarrow 0$ la cola de la distribución tiene un comportamiento del tipo de Gumbel, con parámetros de localización y escala, arbitrarios.

La unificación de las tres familias de distribuciones de valores extremos en una sola, implica una gran simplificación en las implementaciones estadísticas; además, la inferencia de ξ permite que sean los mismos datos que determinen cual es el modelo más apropiado para el comportamiento de la cola de la distribución y así no hay necesidad de introducir en la inferencia incertidumbre debido a la escogencia del modelo más adecuado.

El siguiente teorema unifica en una sola distribución las tres distribuciones anteriores.

Teorema 2.5 (Distribución unificada de Valor Extremo⁴)

Sean $\{a_n\}$ y $\{b_n\}$ sucesiones de números reales, $M = \{X_1, X_2, \dots, X_n\}$ y si

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) \xrightarrow{n \rightarrow \infty} G(x) \Leftrightarrow \lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = G(x) \quad (2.27)$$

Para alguna distribución no degenerada G entonces G es un miembro de la familia GEV con

$$G(x, \sigma, \xi, \mu) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \quad (2.28)$$

Definida en el conjunto

$$\left\{x : 1 + \xi\left(\frac{x - \mu}{\sigma}\right) > 0\right\} \quad (2.29)$$

donde $-\infty < \mu < \infty$, $\sigma > 0$ y $-\infty < \xi < \infty$

Las siguientes son propiedades importantes de la distribución generalizada de valores extremos.

⁴Para un algoritmo para la estimación de los parámetros de la distribución de valor extremo generalizada ver el artículo titulado Algorithm AS 215: Maximum-Likelihood of the parameters of Generalized Extreme-Value Distribution escrito por J. R. M. Hosking.

a. La media existe si $\xi < 1$ y esta dada por:

$$\mu_1 = E(X) = \mu + \frac{\sigma}{\xi} [\sigma(1 - \xi) - 1] \quad (2.30)$$

b. La varianza existe si $\xi < \frac{1}{2}$ y está dada por:

$$\mu_2 = E(X - \mu)^2 = \frac{\sigma^2}{\xi^2} [\Gamma(1 - 2\xi) - \Gamma(1 - \xi)] \quad (2.31)$$

Para el caso limite $\xi \rightarrow 0$, estas ecuaciones se reducen a

$$\lim_{\xi \rightarrow 0} \mu_1 = \lim_{\xi \rightarrow 0} \left\{ \mu + \frac{\sigma}{\xi} [\sigma(1 - \xi) - 1] \right\} = \mu + \sigma\gamma \quad (2.32)$$

$$\lim_{\xi \rightarrow 0} \mu_2 = \lim_{\xi \rightarrow 0} \frac{\sigma^2}{\xi^2} [\Gamma(1 - 2\xi) - \Gamma(1 - \xi)] = \frac{\sigma^2 \pi^2}{6} \quad (2.33)$$

En donde $\gamma = 0.5772$ es la constante de Euler.

Otra característica importante a destacar de la distribución generalizada de extremos, es la estimación de los cuantiles extremos los cuales se pueden obtener invirtiendo la ecuación (2.26):

Si $\xi \neq 0$, como la función de distribución acumulada de la Generalizada de Pareto es:

$$G(x, \xi, \mu, \sigma) = \exp \left\{ - \left[\xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}$$

Pero para hallar el percentil $1 - p$ igualamos:

$$\begin{aligned} 1 - p &= \exp \left(- \left[\xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right) \Rightarrow -\log(1 - p) = \left[\xi \frac{x - \mu}{\sigma} \right]^{-\frac{1}{\xi}} \\ &\Rightarrow \{-\log(1 - p)\}^{-\xi} = \xi \frac{x - \mu}{\sigma} \\ &\Rightarrow x_p = \mu + \frac{\sigma}{\xi} \{-\log(1 - p)\}^{-\xi} \end{aligned}$$

Si $\xi = 0$ la función de distribución acumulada es:

$$G(x) = \exp(-e^{-x})$$

Pero para hallar el percentil $1 - p$ igualamos:

$$\begin{aligned} 1 - p &= \exp \left(-e^{\frac{x - \mu}{\sigma}} \right) \Leftrightarrow \log(1 - p) = -e^{\frac{x - \mu}{\sigma}} \Leftrightarrow -\log(1 - p) = e^{\frac{x - \mu}{\sigma}} \\ \log\{-\log(1 - p)\} &= \frac{x - \mu}{\sigma} \Rightarrow \sigma \log\{-\log(1 - p)\} = x - \mu \\ &\Rightarrow x_p = \mu + \sigma \log\{-\log(1 - p)\} \end{aligned}$$

Si definimos $x_p = -\log(1-p)$ entonces

$$x_p = \begin{cases} \mu + \frac{\sigma}{\xi} \left[1 - \{-\log(1-p)\}^{-\xi} \right] & \text{para } \xi \neq 0 \\ \mu + \sigma \log \{-\log(1-p)\} & \text{para } \xi = 0 \end{cases} \quad (2.34)$$

Un concepto que se requerirá en seguida es la convergencia débil el cual pasamos a definir:

Definición 2.8 (Convergencia en distribución o convergencia débil). Se dice que $\{X_n\}$ converge en distribución o converge débilmente a la variable aleatoria X si para toda función f continua y acotada $\lim_{n \rightarrow \infty} E(f(X_n)) = E(f(X))$ lo cual se anota así $X_n \xrightarrow{d} X$

2.1.5 Dominios de Atracción y Convergencia Débil

El interés principal de los investigadores en la teoría de valores extremos, es tratar de relacionar la función de distribución G con la función de distribución F , es decir dado F , ¿Podemos encontrar G ?. La respuesta a esta pregunta está directamente relacionada con el concepto de dominio de atracción.

En este sentido, se dice que una función de distribución F , discreta o absolutamente continua, pertenece al dominio de máxima atracción de una función de distribución G no degenerada, si existen sucesiones $\{a_n\}$, $a_n > 0$ y $\{b_n\}$ tales que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G(x)$$

Para todos los x , puntos de continuidad de G ; y se denota por $F \in D(G)$.

El dominio de atracción caracteriza un conjunto de distribuciones para las cuales sus máximos normalizados M_n^* , tienen distribución límite G , donde G pertenece a la familia de las distribuciones generalizadas de extremos.

Para poder caracterizar los dominios de atracción, es necesario establecer un conjunto de condiciones necesarias y suficientes sobre F que garanticen que para $i = 1, 2, 3$; o en otras palabras, que condiciones debe cumplir F que garanticen la existencia de sucesiones $\{a_n\}$, $a_n > 0$ y $\{b_n\}$ tales que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_i(x) \quad , \quad i = 1, 2, 3 \quad (2.35)$$

Este tipo de convergencia se llama convergencia débil de funciones de distribuciones o de variables aleatorias.

Corolario 2.5.1 (Base para pruebas de condiciones necesarias)

Sean $\{a_n\}$ y $\{b_n\}$ sucesiones de números reales tales que, si para $\forall x$

$$\lim_{n \rightarrow \infty} n(1 - F(a_n x + b_n)) = \mu(x)$$

entonces

$$\lim_{n \rightarrow \infty} P(M_n < a_n x + b_n) = \exp(-\mu(x))$$

A partir del corolario 2.5.1 se puede establecer los teoremas que determinan que condiciones debe cumplir F para que existan sucesiones normalizadoras $\{a_n\}$, $a_n > 0$ y $\{b_n\}$ que garanticen que M_n^* converge débilmente a una función de distribución no degenerada G .

En los teoremas siguientes $*x$ es el punto final inferior de la distribución F , definida como

$$*x = \alpha(F) = \inf\{x : F(x) > 0\} \quad (2.36)$$

y x^* es el punto final superior de la distribución F donde

$$x^* = W(F) = \sup\{x : F(x) < 1\} \quad (2.37)$$

Teorema 2.6 (Condición necesaria Distribución de Frechet)

Sea $w(F) = \infty^+$ y $a \in R$, $\alpha > 0$ tal que, si $\forall x > 0$

$$\lim_{n \rightarrow \infty^+} \frac{1 - F(x)}{1 - F(t)} = x^{-\alpha} \quad (2.38)$$

entonces existe una sucesión $\{a_n\}$ tal que

$$\lim_{n \rightarrow \infty} P(M_n < a_n x) = G_2(x, \alpha) \quad (2.39)$$

la sucesión normalizada $\{a_n\}$ se puede escoger como

$$a_n = \inf \left\{ x : 1 - F(x) \leq \frac{1}{n} \right\} \quad (2.40)$$

Demostración. Para nuestro propósito se escoge $b_n = 0$ y a_n como aparece en la ecuación (2.40). El supuesto de que $w(F) = \infty^+$ implica que $\{a_n\}$ tiende a ∞^+ , por lo tanto

1. Para $x < 0$ tenemos que

$$a_n x \xrightarrow{n \rightarrow \infty^+} \infty^- \Rightarrow 1 - F(a_n x) \xrightarrow{n \rightarrow \infty^+} 1 \Rightarrow \lim_{n \rightarrow \infty^+} n(1 - F(a_n x)) = \infty^+$$

Y por el corolario 2.5.1 se cumple

$$\lim_{n \rightarrow \infty^+} P(M_n \leq a_n x) = e^{-\infty^+} = 0$$

En consecuencia, el teorema 2.6

2. Para $x > 0$ aplicamos la hipótesis (2.38) con $t = a_n$, entonces obtenemos para $a_n \rightarrow \infty^+$

$$\lim_{n \rightarrow \infty} \frac{1 - F(a_n)}{1 - F(a_n x)} = x^{-\alpha} \quad (2.41)$$

Para $\alpha > 0$ y $x > 0$, en consecuencia

$$\begin{aligned} \lim_{n \rightarrow \infty^+} n(1 - F(a_n x)) &= \lim_{n \rightarrow \infty^+} n \frac{(1 - F(a_n))(1 - F(a_n x))}{1 - F(a_n)} \\ &= \lim_{n \rightarrow \infty^+} n(1 - F(a_n))x^{-\alpha} \quad \text{de acuerdo a (2.41)} \\ &= x^{-\alpha} \lim_{n \rightarrow \infty^+} n(1 - F(a_n)) \end{aligned} \quad (2.42)$$

Ahora en vista de la definición de a_n en (2.40), se tiene que

$$1 - F(a_n + 0) \leq \frac{1}{n} \leq 1 - F(a_n)$$

O equivalentemente

$$1 \leq n(1 - F(a_n)) \leq \frac{1 - F(a_n)}{1 - F(a_n + 0)} \quad (2.43)$$

y como $1 - F$ decrece, entonces,

$$1 - F(a_n + 0) \geq 1 - F(a_n x) \quad \text{para } x > 1$$

Y aplicando esta última desigualdad en (2.43) y usando de nuevo (2.38) obtenemos

$$1 \leq n(1 - F(a_n)) \leq \frac{1 - F(a_n)}{1 - F(b_n + 0)} \leq (1 - \varepsilon)x^\alpha \quad (2.44)$$

En donde $\varepsilon \rightarrow 0$ cuando $n \rightarrow \infty^+$. Como $x > 1$ es arbitrario, (2.44) implica

$$\lim_{n \rightarrow \infty^+} n(1 - F(a_n)) = 1$$

En consecuencia, en (2.42) se tiene

$$\lim_{n \rightarrow \infty^+} n(1 - F(a_n x)) = x^{-\alpha} \lim_{n \rightarrow \infty^+} n(1 - F(a_n)) = x^{-\alpha}$$

Luego por el corolario 2.5.1

$$\lim_{n \rightarrow \infty^+} P(M_n \leq x) = \begin{cases} e^{-x^{-\alpha}} & \text{si } x > 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Teorema 2.7 (Condición necesaria Distribución Weibull)

Sea $w(F)$ finito y supongamos que la función de distribución

$$F^*(x) = F(wF(x) - 1),$$

con $x > 0$, satisface la ecuación (2.42). Entonces, existen sucesiones $\{a_n\}$ y $\{b_n\}$ tal que

$$\lim_{n \rightarrow \infty} P(M_n < a_n x + b_n) = G_3(x, \alpha)$$

Las sucesiones normalizadoras, $\{a_n\}$ y $\{b_n\}$, se pueden escoger como

$$b_n = w(F)$$

$$a_n = w(F) - \inf \left\{ x : 1 - F(x) \leq \frac{1}{n} \right\}$$

Demostración. Como $w(F^*) = \infty^+$ y $F^*(x) = (w(F(x)) - 1)$ con $w(F)$ finito, entonces como (2.42) se satisface para F^* , la conclusión del teorema 2.6 establece que para cualquier $x > 0$

$$\begin{aligned} \lim_{n \rightarrow \infty^+} (F^*)^n(a_n^* x) &= \lim_{n \rightarrow \infty^*} F^n \left(w(F) - \frac{1}{a_n^* x} \right) \\ &= G_2(x; \alpha) \end{aligned} \quad (2.45)$$

Donde

$$\begin{aligned} a_n^* x &= \inf \left\{ x : 1 - F^*(x) \leq \frac{1}{n} \right\} = \inf \left\{ x : 1 - F \left(w(F) - \frac{1}{x} \right) \leq \frac{1}{n} \right\} \\ &= \left(w(F) - \inf \left\{ x : 1 - F(x) \leq \frac{1}{n} \right\} \right)^{-1} \end{aligned}$$

En consecuencia, cuando se escoge $b_n = w(F)$ y $a_n = \frac{1}{a_n^*}$, la ecuación (2.45) se transforma en

$$\begin{aligned} \lim_{n \rightarrow \infty^*} F \left(b_n - \frac{a_n}{x} \right) &= G_2(x; \alpha) \quad \text{si } x > 0 \\ \lim_{n \rightarrow \infty^*} F(b_n - a_n x) &= G_2 \left(-\frac{1}{x}; \alpha \right) \quad \text{si } x < 0 \end{aligned} \quad (2.46)$$

pero para $x < 0$

$$G_2 \left(-\frac{1}{x}; \alpha \right) = \exp(-(-x)^\alpha) = G_3(x; \alpha)$$

Para completar la prueba, se debe notar que como $b_n = w(F)$ y $a_n > 0$

Teorema 2.8 (Condición necesaria Distribución Gumbel)

Se asume que para algún γ finito, se cumple que

$$\int_{\gamma}^{w(F)} (1 - F(x)) dx < \infty^+$$

y para $\alpha(F) < t < w(F)$, definamos

$$R(t) = (1 - F(t))^{-1} \int_t^{w(F)} (1 - F(x)) dx \quad (2.47)$$

si para $\forall x \in \mathbb{R}$

$$\lim_{t \rightarrow w(F)} \frac{1 - F(t + xR(t))}{1 - F(t)} = \exp(-x) \quad (2.48)$$

entonces existen sucesiones $\{a_n\}$ y $\{b_n\}$ tal que

$$\lim_{n \rightarrow \infty^+} (M_n < a_n x + b_n) = G_1(x)$$

Las constantes normalizadoras se pueden escoger como

$$b_n = \inf \left\{ x : 1 - F(x) \leq \frac{1}{n} \right\} \quad (2.49)$$

$$a_n = R(b_n) \quad (2.50)$$

Demostración. Para la prueba se escogerá $\{b_n\}$ y $\{a_n\}$ como lo indica (2.49) y (2.50) respectivamente. Se observa que

$$\lim_{n \rightarrow \infty} b_n = w(F)$$

Luego para $t = b_n$ y $R(t) = R(b_n) = a_n$ según (2.50), y siguiendo (2.48), tenemos:

$$\lim_{n \rightarrow \infty^+} \frac{1 - F(b_n + a_n x)}{1 - F(b_n)} = \exp(-x) \quad \forall x \quad (2.51)$$

Por lo tanto para un x arbitrario

$$\begin{aligned} \lim_{n \rightarrow \infty^+} n(1 - F(a_n x + b_n)) &= \lim_{n \rightarrow \infty^+} n \frac{(1 - F(b_n))(1 - F(a_n x + b_n))}{1 - F(b_n)} \\ &= \exp(-x) \lim_{n \rightarrow \infty^+} n(1 - F(b_n)) \end{aligned} \quad (2.52)$$

pero por (2.49)

$$1 - F(b_n + 0) \leq \frac{1}{n} \leq 1 - F(b_n)$$

Por otra parte, para cualquier $\varepsilon > 0$

$$1 - F(b_n + \varepsilon a_n) \leq 1 - F(b_n + 0)$$

Estas dos últimas desigualdades se pueden combinar y producir el siguiente resultado

$$1 \leq n(1 - F(b_n)) \leq \frac{1 - F(b_n)}{1 - F(b_n + \varepsilon a_n)} \xrightarrow{n \rightarrow \infty^+} \exp(\varepsilon)$$

Usando el límite (2.51). Pero como $\varepsilon > 0$ es arbitrario, entonces

$$\lim_{n \rightarrow \infty} n(1 - F(b_n)) = 1$$

En consecuencia, volviendo al límite en (2.52)

$$\lim_{n \rightarrow \infty^+} n(1 - F(b_n + a_n x)) = \exp(-x) \lim_{n \rightarrow \infty^+} n(1 - F(b_n)) = \exp(-x)$$

Luego usando el corolario (2.5.1) se tiene que

$$\lim_{n \rightarrow \infty^+} P(M_n \leq b_n + a_n x) = \exp(-\exp(-x)) \quad \text{para } -\infty < x < \infty$$

Así que bajo las condiciones de uno de estos teoremas, la función de distribución F pertenece al dominio de atracción de la función de distribución Frechet o de la función de distribución Weibull o de la función de distribución Gumbel respectivamente. Además, estos resultados destacan el hecho de que la escogencia de las constantes normalizadoras no son únicas. También es importante destacar que las condiciones necesarias y suficientes dadas por los teoremas 2.6, 2.7 y 2.8 son equivalentes a las condiciones originales usadas por Gnedenko (1943).

Sin embargo las condiciones necesarias y suficientes de estos teoremas probados son por lo general muy difíciles de verificar. Por esta razón, se utiliza un resultado alternativo que es el teorema de Von Mises (1936), cuyas condiciones de suficiencia son más fáciles de probar y no son tan restrictivas, de hecho son aplicables solo a funciones de distribución F absolutamente continuas.

A continuación se derivan condiciones suficientes de la función de distribución F para asegurar que hay una sucesión de constantes $a_n > 0$ y b_n tales que

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_{\zeta}(x)$$

para algún ζ y todo x . Estas condiciones requieren la existencia de la primera y segunda derivada de F .

Todo teorema de condición suficiente tiene su corolario adicional en donde esta condición suficiente se expresa en términos de la función inversa U en vez de la función de distribución F , que es una manera más sencilla de expresarla.

Teorema 2.9 (Condición suficiente)

Sea F una función de distribución y x^* su extremo derecho. Supongamos que $F''(x)$ existe y $F'(x) > 0$ para todo x en alguna vecindad a la izquierda de x^* . Si

$$\lim_{t \rightarrow x^*} \left(\frac{1-F}{F'} \right)'(t) = \zeta \quad (2.53)$$

o equivalentemente

$$\lim_{t \rightarrow x^*} \frac{(1-F(t))F''(t)}{(F'(t))^2} = -\zeta - 1 \quad (2.54)$$

Entonces F esta en el dominio de atracción de G_ζ .

Demostración. Consideremos $\frac{1}{1-F(U(t))} = t$ por lo tanto

$$U'(t) = \frac{[1-F(U(t))]^2}{F'(U(t))}$$

Luego

$$\frac{U''(t)}{U'(t)} = -2[1-F(U(t))] - \frac{F'(U(t))[1-F(U(t))]^2}{[F'(U(t))]^2}$$

de modo de

$$\frac{tU''(t)}{U'(t)} = -2 - \frac{F'(U(t))[1-F(U(t))]^2}{[F'(U(t))]^2}$$

La relación probada es equivalente a

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^\zeta - 1}{\zeta} \quad (2.55)$$

Para todo $x > 0$. Así que se necesita demostrar que

$$\lim_{t \rightarrow \infty} \frac{tU''(t)}{U'(t)} = \zeta - 1$$

Implica (2.55) para la misma ζ , de modo que $1 < x_0 < x$

$$\log U'(x) - \log U'(x_0) = \int_{x_0}^x \frac{U''(s)}{U'(s)} ds$$

Se tiene que para $x > 0$, $tx > 1$

$$\log U'(tx) - \log U'(t) = \int_1^x A(ts) \frac{ds}{s} \quad \text{con} \quad A(t) := \frac{tU''(t)}{U'(t)}.$$

Se sigue que para $0 < a < b < \infty$,

$$\limsup_{t \rightarrow \infty} \left| \log \frac{U'(tx)}{U'(t)} - \log x^{\zeta-1} \right| = 0$$

Por lo tanto también, ya que $|e^s - e^t| < c|s - t|$ en un intervalo compacto para una constante positiva c ,

$$\limsup_{t \rightarrow \infty} \left| \frac{U'(tx)}{U'^t} - x^{\zeta-1} \right| = 0$$

Esto implica que

$$\frac{U(tx) - U(t)}{tU(t)} - \frac{x^{\zeta} - 1}{\zeta} = \int_1^x \left(\frac{U'(ts)}{U'(t)} - s^{\zeta-1} \right) ds$$

Converge a cero. Por lo tanto por la parte b del teorema 2.3 se concluye que F pertenece al dominio de atracción de G_{ζ} ($F \in D(G_{\zeta})$).

Corolario 2.9.1

La condición (2.53) es equivalente a:

$$\lim_{t \rightarrow \infty} \frac{tU''(t)}{U'(t)} = \zeta - 1$$

Lo cual implica

$$\lim_{t \rightarrow \infty} \frac{U'(tx)}{U'(t)} = x^{\zeta-1}$$

Localmente uniforme en $(0, \infty)$ y finalmente

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{tU'(t)} = \frac{x^{\zeta} - 1}{\zeta}$$

Así que por el teorema 2.2 se obtiene $F \in D(G_{\zeta})$.

Condiciones más simples son posibles para $\zeta \neq 0$

Cuando $\zeta > 0$ se está en el dominio de atracción de la distribución límite de Fréchet. Algunas funciones que caen en esta clase son: Distribución Burr, Distribución Generalizada de Pareto, la distribución log-gamma, la distribución de Cauchy, la distribución F y la t .

Teorema 2.10 (Condición suficiente. Distribución de Frechet)

($\xi > 0$) Si se supone $x^* = \infty$ y F' Si

$$\lim_{t \rightarrow \infty} \frac{tF'(t)}{1 - F(t)} = \frac{1}{\xi} \quad (2.56)$$

para algún $\xi > 0$, entonces F esta en el dominio de atracción de G_{ξ} .

Demostración. Como en la prueba del teorema 2.9 se ve que la condición (2.56) es equivalente a

$$\lim_{t \rightarrow \infty} \frac{tU'(t)}{U(t)} = \xi \quad (2.57)$$

Además

$$\log U(tx) - \log U(t) = \int_1^x \frac{tsU'(ts)}{U(ts)} \frac{ds}{s}$$

por lo tanto (2.57) para $x > 0$

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^{\xi} \quad (2.58)$$

o

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{\xi U(t)} = \frac{x^{\xi} - 1}{\xi}$$

El cuál es la condición (2.20) del teorema 2.3.

Corolario 2.10.1

($\xi > 0$) Condición (2.56) es equivalente a

$$\lim_{t \rightarrow \infty} \frac{tU'(t)}{U(t)} = \xi$$

Que implica (2.54) en vista (2.58).

Cuando $\xi < 0$ se está en el dominio de atracción de la distribución límite de Weibull. Algunas distribuciones que caen en esta clase son: Uniforme, Beta, la Weibull de valor Extremo.

Teorema 2.11 (Condición suficiente. Distribución Weibull)

($\xi < 0$) Se supone $x^* < \infty$ y $F'(x)$ existe para $x < x^*$. Si

$$\lim_{t \rightarrow \infty} \frac{(x^* - t)F'(t)}{1 - F(t)} = -\frac{1}{\xi} \quad (2.59)$$

para algún ξ negativo, entonces F es el dominio de atracción de G_{ξ}

Demostración. Como antes, se puede ver que la condición (2.59) es equivalente a

$$\lim_{t \rightarrow \infty} \frac{tU'(t)}{U(\infty) - U(t)} = -\xi \quad (2.60)$$

$$\log(U(\infty) - U(tx)) - \log(U(\infty) - U(t)) = - \int_1^x \frac{tsU'(ts)}{U(\infty) - U(ts)} \frac{ds}{s}$$

La relación (2.60) implica

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

esto es para $x > 0$

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{-\xi(U(\infty) - U(t))} = \frac{x^{\xi} - 1}{\xi}$$

Corolario 2.11.1

($\xi < 0$) la condición (2.59) es equivalente a

$$\lim_{t \rightarrow \infty} \frac{tU'(t)}{U(\infty) - U(t)} = -\xi$$

que implica (2.60)

Las distribuciones de valores extremos se obtienen como distribuciones límite de los valores más grandes o más pequeños de una muestra aleatoria de tamaño creciente. Para obtener una distribución límite no degenerada es necesario reducir el valor más grande real aplicando una transformación lineal con coeficientes que dependen del tamaño de la muestra. Este proceso es análogo a la estandarización aunque no se restringe para esta sucesión particular de transformaciones lineales. Si hay una distribución límite de interés se debe encontrar como la distribución límite de alguna sucesión de valores reducidos transformados, tales como donde a_n, b_n pueden depender de n pero no de x . Para distinguir el límite de la distribución acumulada del valor más grande reducido de $F(x)$ será denotado por $G(x)$ esta debe satisfacer la ecuación

$$|G(x)|^n = G(a_n x + b_n)$$

Esta ecuación fue obtenida por Frechet (1927) y también por Fisher y Tippett (1928). Es algunas veces llamado *el postulado de estabilidad*.

El teorema siguiente presenta los valores de a_n y b_n para las tres familias de distribuciones límite de valores extremos: la distribución tipo Gumbel, la tipo Frechet y la tipo Weibull.

Teorema 2.12 (Constantes Normalizadas)**Hipótesis:**

1. Si $G = G_2(x; \alpha) = \begin{cases} \exp(-x^{-\alpha}) & x > 0 \\ 0 & x \leq 0 \end{cases}$
2. Si $G = G_3(x; \alpha) = \begin{cases} \exp(-(-x^\alpha)) & x < 0 \\ 0 & x \geq 0 \end{cases}$
3. Si $G = G_1(x; \alpha) = \exp(-e^{-x})$ para todo $x \in \mathbb{R}$

Tesis:

Se escoge

1. $b_n = 0$ y $a_n = F^{-1}\left(1 - \frac{1}{n}\right)$ para la distribución tipo 2 (Distribución tipo Frechet).
2. $b_n = F^{-1}(1)$ y $a_n = F^{-1}(1) - F^{-1}\left(1 - \frac{1}{n}\right)$ para una distribución tipo 3 (Distribución tipo weibull).
3. $a_n = F^{-1}\left(1 - \frac{1}{n}\right) - F^{-1}\left(1 - \frac{1}{n}\right)$, $b_n = F^{-1}\left(1 - \frac{1}{n}\right)$ Para una distribución tipo 1 (Distribución tipo Gumbel).

Demostración.

1. Distribución tipo 1 (Distribución tipo Gumbel)

$$[G(x)]^n = G(a_n x + b_n) = G(x + b_n), \text{ para un } a_n = 1$$

$$\begin{aligned} x = a_n x + b_n &\Rightarrow x - a_n x = b_n \Rightarrow (1 - a_n)x = b_n \\ &\Rightarrow x = \frac{b_n}{1 - a_n} \Rightarrow x = b_n(1 - a_n)^{-1} \end{aligned}$$

$$\text{Por lo tanto } G(x) = G(b_n(1 - a_n)^{-1})$$

$$[G(x)]^{nm} = [G(x + b_n)]^m = G(x + b_n + b_m) = G(x + b_{nm})$$

por lo tanto $b_n + b_m = b_{nm}$

$$[G(x)]^n = G(x + b_n)$$

$$n \log G(x) = \log G(x + b_n)$$

$$\log n + \log [-\log G(x)] = \log [-\log(x + \sigma \log n)]$$

$$h(x) = \log [-\log G(x)]$$

$$h(x) = h(0) \frac{x}{\sigma} \Rightarrow h(x) = \frac{\sigma h(0) - x}{\sigma}$$

$$-\frac{x - \sigma h(0)}{\sigma} = -\log\{-\log G(x)\}$$

$$e^{-\frac{x - \sigma h(0)}{\sigma}} = -\log G(x) \Rightarrow -e^{-\frac{x - \sigma h(0)}{\sigma}} = \log G(x)$$

$$G(x) = \exp\left\{-e^{-\frac{x - \mu}{\sigma}}\right\}$$

dado que $\mu = \sigma \log(-\log G(0))$

2.2 Métodos de umbral y Distribución Generalizada de Pareto

2.2.1 Introducción

Sea X_1, X_2, \dots una sucesión de variables aleatorias independientes e idénticamente distribuidas, que tienen una función de distribución marginal F . Es natural considerar como valores extremos aquellos X_i que exceden algún umbral u grande. De esta manera, el comportamiento estocástico de estas variables está dado por la probabilidad condicional

$$F_u(x) = P(X > u + x / X > u) = \frac{1 - F(u + x)}{1 - F(u)} \quad x > 0 \quad (2.61)$$

Si la distribución F fuera conocida, se podría determinar la función de distribución de (2.61) de los excesos sobre un umbral; pero en la práctica no se conoce, de tal manera que se aplican aproximaciones y en este punto se preserva el paralelismo con la aproximación que se realiza en el modelo de la función de distribución generalizada de los extremos (GEV).

Definición 2.9 (Función del promedio de los excesos). La función

$$e(u) = E(X - u / X > u) \quad (2.62)$$

Es llamada la función del promedio de los excesos de X .

2.2.2 Ajuste de excesos sobre un umbral

Definición 2.10 (Función de distribución de excesos, función del promedio de los excesos). Sea X_1, \dots, X_n son i. i. d. con una $F \in D(G_\xi)$ para algún $\xi \in \mathbb{R}$

Primero, seleccionamos un umbral alto u y se denota por

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

El número de excesos de u de X_1, \dots, X_n se denota los correspondientes excesos por Y_1, \dots, Y_n . La función de distribución de los excesos de X está dada por

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

Es la función de distribución de los excesos sobre el umbral u . La función anterior se puede escribir también como:

$$\bar{F}(u + y) = \bar{F}(u)\bar{F}_u(y)$$

La función del promedio de los excesos juega un papel muy importante en muchos campos. Diferentes nombres surgen de aplicaciones específicas. Por ejemplo, F_u es conocida como la distribución de probabilidad del exceso de vida o tiempo de vida residual en la teoría de la confiabilidad y en estadísticas médicas. En el contexto de seguros, F_u es usualmente referida como la función de distribución del exceso de pérdida.

Por lo tanto, el siguiente teorema permite caracterizar el modelo asintótico de los excesos sobre un umbral.

A continuación se demuestran cinco propiedades de la Distribución de Pareto Generalizada. La primera propiedad a su vez consta de tres numerales.

Teorema 2.13 (Propiedades de la Distribución de Pareto Generalizada GPD)
Supongamos que X tiene una Distribución Generalizada de Pareto con parámetro ξ y β . Entonces $E(X) < \infty$ sí y solo si $\xi < 1$.

Se demuestra con el enunciado anterior los siguientes tres literales: a, b y c.

a. Dado el supuesto (a) demostrar que: Si $r = -\frac{1}{\xi}$ entonces

$$E \left[\left(1 + \frac{\xi}{\beta} x \right)^{-r} \right] = \frac{1}{1 + \xi r} \quad (2.63)$$

Demostración.

$$\begin{aligned} E \left[\left(1 + \frac{\xi}{\beta} x \right)^{-r} \right] &= \int_0^\infty \left(1 + \frac{\xi}{\beta} x \right)^{-r} \frac{1}{\beta} \left(1 + \frac{\xi}{\beta} x \right)^{-\frac{1}{\xi}-1} dx \\ &= \frac{1}{\beta} \lim_{b \rightarrow \infty} \int_0^b \left(1 + \frac{\xi}{\beta} x \right)^{-r-\frac{1}{\xi}-1} dx \end{aligned}$$

Haciendo la siguiente sustitución: $u = 1 + \frac{\xi}{\beta} x \Rightarrow \frac{\beta}{\xi} du = dx$

$$= \frac{1}{\xi} \lim_{b \rightarrow \infty} \int_1^{u_1} u^{-r-\frac{1}{\xi}-1} du = \frac{1}{\xi} \lim_{b \rightarrow \infty} \frac{\left(1 + \frac{\xi}{\beta} x \right)^{\left(-r-\frac{1}{\xi}\right)} \Big|_0^b}{-r-\frac{1}{\xi}} = \frac{1}{1+\xi r'}$$

siempre y cuando $r > -\frac{1}{\xi}$

b. Dado el supuesto (a) Demostrar que si $k \in N$ entonces

$$E \left[\left(\ln \left(1 + \frac{\xi}{\beta} x \right) \right)^r \right] = \xi^k k!, \quad (2.64)$$

c. Dado el supuesto (a) Demostrar que si $\xi < \frac{1}{r}$ con $r \in R$, entonces

$$E \left[\left(X \left(\overline{G}(X; \xi, \beta) \right) \right)^r \right] = \frac{\beta}{(r+1-\xi)(r+1)} \frac{r+1}{|\xi|} > 0 \quad (2.65)$$

$$\begin{aligned} E \left[\left(X \left(\overline{G}_{\xi, \beta}(x) \right) \right)^r \right] &= \int_0^\infty x \left(1 + \frac{\xi}{\beta} x \right)^{-\frac{r}{\xi}} \frac{1}{\beta} \left(1 + \frac{\xi}{\beta} x \right)^{-\frac{1}{\xi}-1} dx \\ &= \frac{1}{\beta} \int_0^\infty x \left(1 + \frac{\xi}{\beta} x \right)^{-\frac{r}{\xi}-\frac{1}{\xi}-1} dx = \end{aligned}$$

$$u = 1 + \frac{\xi}{\beta} x \Rightarrow dx = \frac{\beta}{\xi} du \quad x = \frac{\beta}{\xi}(u-1)$$

Por lo tanto queda la siguiente integral:

$$\begin{aligned} &= \frac{\beta}{\xi^2} \lim_{b \rightarrow \infty} \int_0^b (u-1) u^{-\frac{r}{\xi}-\frac{1}{\xi}-1} du = \frac{\beta}{\xi^2} \lim_{b \rightarrow \infty} \int_0^b \left(u^{-\frac{r}{\xi}-\frac{1}{\xi}} - u^{-\frac{r}{\xi}-\frac{1}{\xi}-1} \right) du \\ &= \frac{\beta}{(r+1-\xi)(r+1)}, \end{aligned}$$

siempre y cuando $\frac{r+1}{|\xi|} > 0$

$$E(X^r) = \frac{\beta^r \Gamma(\xi^{-1} - r)}{\xi^{r+1} \Gamma(1 + \xi^{-1})} r! \quad (2.66)$$

1. Para todo $\zeta \in R$, $F \in D(G_\zeta)$ si y solo si

$$\limsup_{u \rightarrow x^*} |F_u(x) - G_{\zeta, \beta(u)}(x)| = 0$$

para alguna función β positiva.

2. Si suponemos que $x_i \in D(\zeta, \beta)$, $i = 1, 2$, entonces

$$\frac{\overline{G}(x_1 + x_2; \zeta, \beta)}{\overline{G}(x_1, \beta, \beta)} = \overline{G}(x_2; \zeta, \beta + \zeta x_1)$$

Demostración.

$$\begin{aligned} \frac{\overline{G}(x_1 + x_2; \zeta, \beta)}{\overline{G}(x_1; \zeta, \beta)} &= \frac{\left(1 + \zeta \frac{x_1 + x_2}{\beta}\right)^{-\frac{1}{\zeta}}}{\left(1 + \zeta \frac{x_1}{\beta}\right)} = \left(\frac{1 + \zeta \frac{x_1 + x_2}{\beta}}{1 + \zeta \frac{x_1}{\beta}}\right)^{-\frac{1}{\zeta}} = \left(\frac{1 + \zeta \frac{x_1}{\beta} + \zeta \frac{x_2}{\beta}}{1 + \zeta \frac{x_1}{\beta}}\right)^{-\frac{1}{\zeta}} \\ &= \left(1 + \frac{\zeta \frac{x_2}{\beta}}{1 + \zeta \frac{x_1}{\beta}}\right)^{-\frac{1}{\zeta}} = \left(1 + \frac{\zeta x_2}{\beta + \zeta x_1}\right)^{-\frac{1}{\zeta}} = \left(1 + \frac{\zeta}{\beta + \zeta x_1} x_2\right)^{-\frac{1}{\zeta}} \\ &= \overline{G}(x_2; \zeta, \beta + \zeta x_1) \end{aligned}$$

3. Si se asume que n es Poi (λ) independiente de una sucesión $\{X_n\}$ i. i. d. con una Distribución Generalizada de Pareto con parámetros ζ, β . Escrito $M_n = \max(X_1, \dots, X_n)$. Entonces

$$P(M_n \leq x) = \exp \left\{ -\lambda \left(1 + \zeta \frac{x}{\beta}\right)^{-\frac{1}{\zeta}} \right\} = G(x; \zeta, \mu, \beta)$$

donde $\mu = \beta \zeta^{-1} (\lambda \zeta - 1)$ y $\sigma = \beta \lambda \zeta$

Demostración.

$$\begin{aligned} P(M_n \leq x) &= \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} G^n(x; \zeta, \beta) = \exp \{ -\lambda \overline{G}(x; \zeta, \beta) \} \\ &= \exp \left\{ -\lambda \left(1 + \zeta \frac{x}{\beta}\right)^{-\frac{1}{\zeta}} \right\} \\ &= \exp \left\{ - \left(1 + \zeta \frac{x - \zeta^{-1} \beta (\lambda \zeta - 1)}{\beta \lambda \zeta}\right)^{-\frac{1}{\zeta}} \right\}, \quad \zeta \neq 0 \end{aligned}$$

Para el caso $\zeta = 0$ se reduce a $P(M_n \leq x) = \exp \left\{ -e^{-\frac{x - \beta \ln \lambda}{\beta}} \right\}$

4. Supongamos que x tiene Distribución Generalizada de Pareto con parámetro $\xi < 1$ y β , entonces para $u < x^*$

$$e(u) = E(X - u/X > u) = \frac{\beta + \xi u}{1 - \xi} \quad , \quad \beta + u\xi > 0$$

Demostración.

$$e(u)E(X - u/X > u) = \int_u^\infty (x - u) \frac{dF(x)}{\bar{F}(u)} = \frac{1}{\bar{F}(u)} \int_u^\infty \bar{F}(x) dx$$

$$\bar{F}(x) = \bar{G}(x; \xi, \beta) = \left(1 + \xi \frac{x}{\beta}\right)^{-\frac{1}{\xi}} \quad \text{reemplazando se tiene}$$

$$e(u) = \frac{1}{\left(1 + \xi \frac{u}{\beta}\right)^{-\frac{1}{\xi}}} \int_u^\infty \left(1 + \xi \frac{x}{\beta}\right)^{-\frac{1}{\xi}} dx$$

$$v = 1 + \xi \frac{x}{\beta} \Rightarrow dv = \frac{\xi}{\beta} dx \Rightarrow dx = \frac{\beta}{\xi} dv$$

$$e(u) = \frac{1}{\left(1 + \xi \frac{u}{\beta}\right)^{-\frac{1}{\xi}}} \lim_{b \rightarrow \infty} \int_u^b v^{-\frac{1}{\xi}} \frac{\beta}{\xi} dv = \frac{\beta}{\xi \left(1 + \xi \frac{u}{\beta}\right)^{-\frac{1}{\xi}}} \lim_{b \rightarrow \infty} \left[v^{-\frac{1}{\xi}+1}\right]$$

$$e(u) = \frac{\beta + \xi u}{1 - \xi},$$

siempre y cuando $\beta + \xi u > 0$.

La propiedad (3) anterior es a menudo reformulada como sigue: *la clase de GPDs es cerrada con respecto al cambio del umbral*. Es más las variables aleatorias en la ecuación de (3) es la probabilidad condicional que, dado de que las variables aleatorias excedan a x_1 , es también que excedan el umbral $x_1 + x_2$. La regla establece que esta probabilidad es otra vez de tipo Pareto generalizada. Esta propiedad cerrada es importante en reaseguros, donde las GDP son básicas cuando se tratan con contratos con excesos de pérdidas.

La propiedad (2) anterior sugiere una GPD como una aproximación apropiada de los excesos F_u para u grandes. Este resultado se remota a Pickands y es a menudo formulado como sigue. Para alguna función β que sea estimada de los datos,

$$\bar{F}(x; u) = P(X - u > x/X > u) \approx \bar{G}(x; \xi, \beta(u)) \quad , \quad x > 0$$

Alternativamente uno considera para $x > u$,

$$P(X > x/X > u) \approx \bar{G}(x; \xi, u, \beta(u)) \quad (2.67)$$

En ambos casos u debe ser tomado lo suficientemente grande. Juntos (2) y (5) dan una técnica gráfica maravillosa para seleccionar u tan alto que una aproximación de la función de los excesos a una GPD está justificada: dada una muestra aleatoria X_1, \dots, X_n i. i. d. construir la función del promedio de los excesos empírica $e(u)$. De (5) se sabe que la función del promedio de los excesos de una GPD es lineal, por lo tanto chequear para una región u donde el gráfico de $e_n(u)$ llegue a ser rigurosamente lineal. Para tal u una aproximación de F_u por GPD parece razonable.

De las observaciones anteriores sugerimos el siguiente modelo aproximado para los excesos de una muestra i. i. d:

- El número de excesos de un umbral alto sigue un proceso de Poisson.
- Los excesos sobre un umbral alto pueden ser modelados por una Distribución Generalizada de Pareto (GPD).
- Un valor apropiado para el umbral alto puede ser encontrado graficando la función del promedio de los excesos empírica.
- La distribución del máximo de un número de Poisson de excesos iid sobre un umbral alto es una Distribución de valor extremo generalizada.

La propiedad (2) del teorema (2.14) da un resultado límite para $\bar{F}(y; u)$ como

$$\lim_{u \rightarrow x^*} \sup |F(x; u) - G(x; \xi, \beta(u))| = 0$$

Para una función positiva β . Basada en esta relación, para un u grande la siguiente aproximación se sugiere:

$$\bar{F}(y; u) \approx \bar{G}(y; \xi, \beta(u)) \quad (2.68)$$

Es importante resaltar que β es una función del umbral u . En la práctica u deberá tomar un valor suficientemente grande. Dado un valor de u , ξ , y $\beta = \beta(u)$ son estimados a partir de los datos de los excesos, por lo tanto los datos estimados dependerán de u .

La ecuación (2.68) luego sugiere un método para estimar el extremo de la cola de F estimado $\bar{F}_u(y)$ y $\bar{F}(u)$ separadamente. Un estimador natural para $\bar{F}(u)$ es dado por la función de distribución empírica

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

Por otro lado la aproximación de la generalizada de Pareto motiva un estimador de la forma

$$(\bar{F}(u))(\bar{F}_u(y)) = \bar{G}(y; \hat{\xi}, \beta)$$

Para un $\hat{\xi} = \hat{\xi}_{N_u}$ y $\beta = \beta_{N_u}$. Un estimador para la cola para cuando $y > 0$ tiene la forma:

$$(\bar{F}(u + y)) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{y}{\beta} \right)^{-\frac{1}{\hat{\xi}}} \quad (2.69)$$

En el caso de Frechet y de Gumbel ($\xi \geq 0$) la restricción del dominio es $y \geq 0$ se destaca claramente que se estima \bar{F} en la cola superior. Un estimador del cuantil x^* es el siguiente:

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

Además para $\hat{\xi} < 0$ un estimador del punto superior de la derecha x^* de F está dado por $\hat{x}^* = u - \frac{\beta}{\hat{\xi}}$. Lo anterior es obtenido poniendo $\hat{x}^* = \hat{x}_1$ es decir ($p = 1$) en la ecuación anterior.

Teorema 2.14 (Distribución Generalizada de Pareto⁵)

Sea X_1, X_2, \dots una sucesión de variables aleatorias i.i.d. con función de distribución común y sea $M_n = \{X_1, \dots, X_n\}$ sea X un término X_i cualquiera de la sucesión y supongamos que F satisface el teorema 3.4.

Si $P(M_n \leq x) \approx G(x; \xi, \mu, \sigma)$, $n \rightarrow \infty$ donde

$$G(x; \xi, \mu, \sigma) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}$$

para algún μ , $\sigma > 0$ y ξ .

Entonces para algún u suficientemente grande la función de distribución de $X - u$ condicionado a que $X > u$ definida sobre $\left\{ y : y > 0 \text{ y } \left(1 + \frac{\xi y}{\sigma} \right) > 0 \right\}$ donde

$$\hat{\sigma} = \sigma + \xi(u - \mu) \quad (2.70)$$

es

$$H(y) = 1 - \left(1 + \frac{\xi y}{\sigma} \right)^{-\frac{1}{\xi}} \quad (2.71)$$

⁵Para un desarrollo de la estimación de parámetros y cuantiles ver el artículo *Parameter and Quantile Estimation for the Generalized Pareto Distribution* escrito por los autores J. R. M Hosking and J. R. Wallis.

Demostración. Se considera la v. a. X lo cual tiene una función de distribución F que cumple la suposición del teorema 3.4, entonces

$$F^n \approx \exp \left\{ - \left[1 + \zeta \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\zeta}} \right\}$$

Para algunos parámetros $\mu, \sigma > 0$ y ζ Por lo tanto,

$$n \log(F) \approx - \left[1 + \zeta \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\zeta}} \quad (2.72)$$

Luego para valores grandes de x , una expansión de Taylor indica que:

$$\log F(x) \approx -\{1 - F(x)\}$$

Y sustituyendo está última expresión en (2.72), obtenemos después de re arreglar

$$1 - F(u) \approx \frac{1}{n} \left[1 + \zeta \left(\frac{u - \mu}{\sigma} \right) \right]^{-\frac{1}{\zeta}} \quad (2.73)$$

Para u grande. Similarmente, para $y > 0$

$$1 - F(u + y) \approx \frac{1}{n} \left[1 + \zeta \left(\frac{u + y - \mu}{\sigma} \right) \right]^{-\frac{1}{\zeta}}$$

por lo tanto,

$$\begin{aligned} P \left(X > u + \frac{y}{X} > u \right) &\approx \frac{n^{-1} \left[1 + \frac{\zeta(u + y - \mu)}{\sigma} \right]^{-\frac{1}{\zeta}}}{n^{-1} \left[1 + \frac{\zeta(u - \mu)}{\sigma} \right]^{-\frac{1}{\zeta}}} = \left[1 + \frac{\frac{\zeta(u + y - \mu)}{\sigma}}{1 + \frac{\zeta(u - \mu)}{\sigma}} \right]^{-\frac{1}{\zeta}} \\ &= \left[1 + \frac{\zeta y}{\sigma} \right]^{-\frac{1}{\zeta}} \end{aligned} \quad (2.74)$$

Donde $\hat{\sigma} = \sigma + \zeta(u - \mu)$

La familia de distribuciones definida por la ecuación (2.70) se conoce distribución generalizada de Pareto (GPD). Por lo tanto para un umbral u lo suficientemente grande y existe algún $\hat{\sigma}$ que depende de u y algún ζ para los cuales la distribución generalizada de Pareto es una muy buena aproximación de los excesos sobre un umbral u .

2.3 Estimación del índice de valor extremo y pruebas

2.3.1 Introducción

Cuando aplicamos la teoría de valor extremo la primera decisión que se debe tomar es si se aplica el método de bloques o picos sobre un umbral de acuerdo a la disponibilidad de datos que se tenga. Si se elige el primer método se trabaja con la distribución de valores extremos y si se elige el segundo se trabaja con la distribución generalizada de Pareto. Identificada la distribución con la cual se va a trabajar se procede a estimar los parámetros de la distribución: Existen tres parámetros uno de forma, uno de escala y otro de localización. El parámetro más importante es el de forma o índice de valor extremo que define el tipo de distribución si $\zeta > 0$ es la distribución tipo 2 Frechet, si $\zeta < 0$ la distribución tipo 3 Weibull, y si $\zeta \rightarrow \infty, -\infty$ es la distribución tipo 1 Gumbel.

Hay varios métodos para estimar el índice de valor extremo: Hill ($\hat{\zeta}_H$) Bajo condiciones apropiadas, el estimador de Hill es consistente solamente para valores positivos de ζ , Pickands ($\hat{\zeta}_P$), Máximo Verosimilitud ($\hat{\zeta}_{MV}$), Momentos ponderados por probabilidad⁶ ($\hat{\zeta}_{MPP}$), y Hill Negativo ($\hat{\zeta}_F$), el de Máxima Verosimilitud (MV) está definido para $\zeta > -\frac{1}{2}$, el de Pickands y el de momentos están definidos y son consistentes para todos los valores reales de ζ , el Momentos Ponderados por probabilidad (MPP) es consistente para $\zeta < 1$, y el Hill Negativo para $\zeta < -\frac{1}{2}$.

Para el cálculo del parámetro de forma existen diferentes métodos para estimarlo el estimador de Hill del cual hay dos posibilidades según el valor que tenga mayor que cero o menor de $-\frac{1}{2}$ y el estimador de Pickands que sirve para el caso general su valor es posiblemente cualquier valor real, Un estimador de momentos para un valor dentro de los reales, un estimador de momentos ponderados según una probabilidad cuando el valor es menor de 1 en nuestro caso se elegirá el estimador de Hill.

También deben estimarse las constantes de normalización, los cuantiles y la probabilidad de la cola.

Pero cuando se usa un método de estimación: Método de Hill, Método de Máxima

⁶Para un desarrollo amplio del Método de estimación de parámetros de Momentos ponderados por probabilidad en el artículo Estimation of Generalized Extreme-Value Distribution by the Method of probability-Weighted Moments. Revista American Society for Quality.

verosimilitud⁷, Método de Momentos Ponderados por Probabilidad o cualquier otro la diferencia entre uno y otro son insesgades, la eficiencia, la consistencia y la suficiencia.

También se debe saber cuál es la distribución del parámetro porque a partir de esta información se construyen las pruebas estadísticas y los intervalos de confianza pertinentes.

Las dos propiedades en las que se concentrará este trabajo son la consistencia y la distribución asintótica de los estimadores.

Definición 2.11 (Estimador consistente). Se dice que $\hat{\theta}$ es un estimador consistente de θ si, $\forall \varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| \leq \varepsilon) = 1 \quad \text{o} \quad \lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| > \varepsilon) = 0$$

La condición alternativa al teorema 2.3 sirve como una base para las aplicaciones de la teoría del valor extremo.

Considere la relación 2.22; existe una función no decreciente positiva f tal que

$$\lim_{t \rightarrow x^*} \frac{1 - F(t + xf(t))}{1 - F(t)} = (1 - \zeta x)^{-\frac{1}{\zeta}} \quad (2.75)$$

Para todo x para el cual $1 + \zeta x > 0$ donde $x^* = \sup\{x : F(x) < 1\}$

Sea X una v. a. con función de distribución F y sea $F \in D(G_{\zeta})$ para algún $\zeta \in \mathbb{R}$.

Entonces (2.75) nos dice que $x > 0$ y $x < (0 \vee (-\zeta))^{-1}$

$$\lim_{t \rightarrow x^*} P\left[\frac{x-t}{f(t)} > x/X > t\right] = (1 + \zeta x)^{\frac{1}{\zeta}} \quad (2.76)$$

Esto es, la distribución condicional de $\frac{x-t}{f(t)}$ dado $X > t$ tiene una distribución límite cuando $t \rightarrow x^*$.

$$H_{\zeta}(x) := 1 - (1 + \zeta x)^{-\frac{1}{\zeta}}, \quad 0 < x < (0 \vee (-\zeta))^{-1} \quad (2.77)$$

Donde para $\zeta = 0$ el lado derecho es interpretando $1 - e^{-x}$ Esta clase de funciones de distribuciones es llamada la función generalizada de Pareto.

⁷Para un desarrollo amplio ver el artículo de *Maximum Likelihood Estimation in a Class of Nonregular Cases* de Richard L. Smith. Revista Bimetrika Trust.

La relación (2.75) significa en términos generales que de algún umbral alto t en adelante (es decir, $X > t$) la función de distribución puede ser escrita como

$$P(X > x) = 1 - F(x) \cong (1 - F(t)) \left\{ 1 - H_{\zeta} \left(\frac{x-t}{f(t)} \right) \right\}, \quad x > t$$

La cual es una familia paramétrica de distribución de la cola.

Se puede esperar que esta aproximación se cumpla para estadísticos de orden intermedios y extremos. Sea X_1, \dots, X_n v. a. independientes e idénticamente distribuidas con función de distribución F y F_n la correspondiente distribución empírica, esto es, $F_n(x) := n^{-1} \sum_{i=1}^n \mathbf{1}_{(X_i \leq x)}$

Se aplica la última aproximación con $t := X_{n-k,n}$ donde se selecciona

$$K = K(n) \rightarrow \infty, \quad \frac{k}{n} \rightarrow 0, \quad n \rightarrow \infty$$

entonces

$$1 - F(x) \approx (1 - F(X_{n-k,n})) \left\{ 1 - H_{\zeta} \left(\frac{x - X_{n-k,n}}{f\left(\frac{n}{k}\right)} \right) \right\}$$

Y, ya que $1 - F(X_{n-k,n}) \approx 1 - F_n(X_{n-k,n}) = \frac{k}{n}$

$$1 - F(x) \approx \frac{k}{n} \left\{ 1 - H_{\zeta} \left(\frac{x - X_{n-k,n}}{f\left(\frac{n}{k}\right)} \right) \right\} \quad (2.78)$$

Para poder usar esta aproximación se necesita estimar ζ y la función f en el punto $\frac{n}{k}$.

Esta aproximación es válida para cualquier x más grande que $X_{n-k,n}$ y puede ser usada siempre que $x > X_{n-k,n}$ y puede ser usada siempre que $x > X_{n,n}$, esto es, fuera del rango de observaciones y este es en efecto la base para aplicaciones de valores extremos.

Posteriormente se considerará una aplicación de la relación (2.20) ya que existe una función positiva tal que $\forall x > 0$

$$\lim_{t \rightarrow \infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^{\zeta} - 1}{\zeta} \quad (2.79)$$

La relación (2.77) se convierte en la siguiente aproximación

$$U(x) \approx U(t) + a(t) \frac{\left(\frac{x}{t}\right)^{\zeta} - 1}{\zeta}, \quad x > t \quad (2.80)$$

Esta aproximación es usual cuando uno desea estimar un cuantil

$$F^{\leftarrow}(1-p) = U\left(\frac{1}{p}\right)$$

Con p muy pequeño, ya que este cuantil está luego relacionado con muchos bajos cuantiles.

$U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right)$ El cual puede ser estimado por un estadístico de orden intermedio. Por lo tanto se selecciona $t := \frac{n}{k}$ con $k = k(n) \rightarrow \infty$ y $\frac{k}{n} \rightarrow 0$, $n \rightarrow \infty$ Entonces para grandes y y por ejemplo $y = \frac{1}{p}$ con p pequeño,

$$U(y) \approx U\left(\frac{n}{k}\right) + a\left(\frac{n}{k}\right) \frac{\left(\frac{y}{n}\right)^{\zeta} - 1}{\zeta} \quad (2.81)$$

En orden a hacer esta aproximación aplicable se necesita estimar ζ la función a en el punto $\frac{n}{k}$ y $U\left(\frac{n}{k}\right)$. El cuantil anterior puede ser estimado por un estadístico de orden intermedio, el cual puede ser usado para $y < n$ y también para extrapolar fuera de la muestra.

2.3.2 El estimador de Hill $\zeta > 0$

Para introducir el estimador de Hill, se comienza presentando el teorema 1.2.1 (Ver en Haan y Ferreira (2006))

Si $\zeta > 0$ y $F \in D(G)$ sí y solo si

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\frac{1}{\zeta}}, \quad \zeta > 0$$

El parámetro $\alpha = \frac{1}{\zeta}$ es llamado el índice de colas de F el teorema 2.3 nos permite escribir esta expresión de la siguiente manera:

$$\lim_{t \rightarrow \infty} \frac{\int_t^\infty (1 - F(x)) \frac{dx}{x}}{1 - F(t)} = \zeta$$

Ahora la integral parcial da

$$\int_t^\infty (1 - F(s)) \frac{ds}{s} = \int_t^\infty (\log u - \log t) dF(u)$$

Por lo tanto se tiene

$$\lim_{t \rightarrow \infty} \frac{\int_t^\infty (\log u - \log t) dF(u)}{1 - F(t)} = \zeta \quad (2.82)$$

En orden a desarrollar un estimador basado en este resultado asintótico reemplazando en (2.82) el parámetro t por el estadístico de orden intermedio $X_{n-k,n}$ y F por la función de distribución empírica F_n . Entonces se logra el estimador de Hill (1975) $\hat{\gamma}_H$ definido por

$$\hat{\gamma}_H = \frac{\int_{X_{n-k,n}}^{\infty} \log u - \log X_{n-k,n} dF_n(u)}{1 - F_n(X_{n-k,n})}$$

o

$$\hat{\gamma}_H = \frac{1}{k} \sum_{i=0}^{k-1} \log X_{n-i,n} - \log X_{n-k,n} \quad (2.83)$$

Teorema 2.15 (Consistencia débil)

Sea X_1, X_2, \dots variables aleatorias i. i. d con función de distribución F . Suponga que $F \in D(G_{\xi})$ con $\xi > 0$ Entonces cuando $n \rightarrow \infty$, $k = k(n) \rightarrow \infty$ y $\frac{k}{n} \rightarrow 0$

$$\hat{\xi}_H \xrightarrow{P} \xi$$

Demostración. Por el corolario 1.2.10 (Ver en Haan y Ferreira (2006)) $F \in D(G_{\xi})$ con $\xi > 0$ implica $\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^{\xi}$ para $x > 0$, esto es (Proposición B.1.9) (Ver en Haan y Ferreira (2006)) para $x \geq 1$, y $t \geq t_0$

$$(1 - \xi)x^{\xi - \xi'} < \frac{U(tx)}{U(t)} < (1 + \xi)x^{\xi + \xi'}$$

O equivalentemente

$$\log(1 - \xi) + (\xi - \varepsilon') \log x < \log U(tx) - \log U(t) < \log(1 + \xi) + (\xi + \varepsilon') \log x \quad (2.84)$$

Sea Y_1, Y_2, \dots v. a. i. d. con distribución común $1 - \frac{1}{y}$, $y \geq 1$

Note que

$$U(Y_i) =^d X_1 \quad , \quad i = 1, 2$$

Así que es suficiente probar el resultado para

$$\hat{\xi}_H = k^{-1} \sum_{i=1}^{k-1} \log U(Y_{n-i,n}) - \log U(Y_{n-k,n})$$

Aplicar (2.84) con $t \cong Y_{n-k,n}x - \frac{Y_{n-i,n}}{Y_{n-k,n}}$ ya que por el lema 1 (ver Haan y Ferreira (2006)) $Y_{n-k,n} \rightarrow \infty$ a. s. $n \rightarrow \infty$ eventualmente

$$\begin{aligned} \log(1 - \varepsilon) + (\xi - \varepsilon') \log \frac{Y_{n-i,n}}{Y_{n-k,n}} &< \log U(Y_{n-i,n}) - \log U(Y_{n-k,n}) \\ &< \log(1 + \varepsilon) + (\xi - \varepsilon') \log \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right) \end{aligned}$$

Para $i = 0, 1, 2, 3, \dots, k-1$, por lo tanto

$$\begin{aligned} \log(1 - \varepsilon) + (\xi - \varepsilon') \frac{1}{k} \sum_{i=1}^n \log \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right) &< \hat{\xi}_H \\ &< \log(1 + \varepsilon) + (\xi - \varepsilon') \frac{1}{k} \sum_{i=1}^n \log \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right) \end{aligned}$$

Además es suficiente probar que cuando $n \rightarrow \infty$

$$\frac{1}{k} \sum_{i=1}^{K-1} \log \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right) \xrightarrow{P} 1$$

Teorema 2.16 (Consistencia débil)

Sea X_1, X_2, \dots v. a. i. i. d con función de distribución F . Supongamos que para una sucesión de enteros $k = k(n) \rightarrow \infty$, $\frac{k}{n} \rightarrow 0$ y $\frac{k(n+1)}{k(n)} \rightarrow 1$ cuando $n \rightarrow \infty$,

$$\hat{\xi}_H \xrightarrow{P} \xi$$

entonces $F \in D(G_\xi)$

Demostración. Sea F_n la función de distribución empírica de X_1, \dots, X_n y G_n la función de distribución empírica de Y_1, \dots, Y_n las cuales son independientes e igualmente distribuidas $1 - \frac{1}{x}$, $x \geq 1$ luego para cada n

$$1 - F_n(x) =^d 1 - G_n \left(\frac{1}{1 - F(x)} \right)$$

Se escribe

$$\begin{aligned} \hat{\xi}_H &= \frac{n}{k} \int_{X_{n-k,n}}^{\infty} \left\{ 1 - G_n \left(\frac{1}{1 - F(u)} \right) \right\} \frac{du}{u} \\ &= \frac{n}{k} \int_{X_{n-k,n}}^{\infty} (1 - G_n(s)) d \log U(s) \end{aligned}$$

con $Y_{1,n} \leq Y_{2,n} \leq \dots \leq Y_{n,n}$ el estadístico de orden de Y_1, \dots, Y_n Se van a usar los siguientes resultados

1. $P(\sup s(1 - G_n(s)) > b) = \frac{1}{b}$, para $b > 1$

$$P(\inf s(1 - G_n(s)) < b) = \frac{e}{a} e^{-\frac{1}{a}} \quad \text{para } 0 < a < 1$$

(Ver Shoroack y Wellner (1986), pp. 345 y 415)

2. $\inf s(1 - G_n(s)) = \inf \frac{ks}{n} Y_{n-[ks],n} = \left(\inf s \frac{Y_{n-[ns],n}}{Y_{n-k,n}} \right) \left(\frac{k}{n} Y_{n-k,n} \right)$ donde los dos factores son independientes.

$$\left(\inf s Y_{k-[ks],n}^* \right) \frac{k}{n} Y_{n-k,n}$$

con los Y^* independientes e iguales en distribución a los Y esto

$$\left(\inf s(1 - G_n(s)) \right) \frac{k}{n} Y_{n-k,n}$$

3. Del corolario 2.2.2 (ver en Haan y Ferreira (2006))

$$\lim_{n \rightarrow \infty} P \left(1 - \varepsilon \leq \frac{k}{n} Y_{n-k,n} \leq 1 + \varepsilon \right) = 1$$

Para t suficientemente grandes sea $n = n(t)$ el entero que satisface

$$\frac{n(t)}{k(n(t))} \leq t \leq \frac{n(t) + 1}{k(n(t) + 1)}$$

Considere

$$\begin{aligned} P \left(\frac{1 - \varepsilon}{(1 + \varepsilon)^2} t(1 + \varepsilon) \int_{t(1 + \varepsilon)}^{\infty} \frac{d \log U(s)}{s} \leq \frac{n}{k} (1 - \varepsilon) \int_{\frac{n}{k}(1 + \varepsilon)}^{\infty} \frac{d \log U(s)}{s} \right. \\ \leq Y_{n-k,n} \int_{Y_{n-k,n}}^{\infty} \frac{d \log U(s)}{s} \\ \leq \frac{n}{k} \int_{Y_{n-k,n}}^{\infty} \frac{s(1 - G_n(s))}{1 - \varepsilon} \frac{d \log U(s)}{ds} \\ \left. \leq \frac{\tilde{\zeta} + \varepsilon}{1 - \varepsilon} \right) \end{aligned}$$

La primera desigualdad es verdad por definición: la segunda y la cuarta desigualdad son verdad con probabilidad tendiendo a 1. Para la segunda usamos el resultado (3); la cuarta es verdad por supuesto. Por los resultados (1) y (2) la tercera desigualdad es verdad con probabilidad al menos

$$1 - e(1 - \varepsilon)^{-1} \exp \left(-\frac{1}{1 - \varepsilon} \right) > 0$$

por lo tanto se alcanza la conclusión para cada $\varepsilon > 0$

$$P\left(t \int_t^\infty \frac{d \log U(s)}{s} \leq \varepsilon + \zeta\right) > 0$$

para t suficientemente grande. Por lo tanto para t suficientemente grande,

$$t \int_t^\infty s^{-1} d \log U(s) \leq \zeta + \varepsilon$$

se consigue la otra desigualdad en forma similar, Por tanto

$$\lim_{t \rightarrow \infty} t \int_t^\infty \frac{d \log U(s)}{s} = \zeta$$

Ahora por integración parcial

$$t \int_t^\infty \frac{1}{s} d \log U(s) = t \int_t^\infty \log U(s) \frac{ds}{s} - \log U(t)$$

Por lo tanto por la observación B.2.14(2) (Ver en Haan y Ferreira (2006)) se define para $x > 0$

$$\lim_{t \rightarrow \infty} (\log U(tx) - \log U(t)) = \zeta \log x$$

Es decir, U varía regularmente con índice ζ lo cual implica por la proposición B.1.9 (ver en Haan y Ferreira (2006)) que la función $1 - F$ varía regularmente con índice $-\frac{1}{\zeta}$

Teorema 2.17 (Normalidad asintótica para ζ_h)

Suponga que la función de distribución F satisface la condición de segundo orden, esto es, para $x > 0$

$$\lim_{t \rightarrow \infty} \frac{\frac{U(tx)}{U(t)} - x^\zeta}{A(t)} = x^\zeta \frac{x^\rho - 1}{\rho} \quad (2.85)$$

o equivalentemente

$$\lim_{t \rightarrow \infty} \frac{1 - \frac{F(tx)}{F(t)} - x^{-\frac{1}{\zeta}}}{A\left(\frac{1}{1-F(t)}\right)} = x^{-\frac{1}{\zeta}} \frac{x^{\frac{\rho}{\zeta}} - 1}{\zeta^\rho} \quad (2.86)$$

Donde $\zeta > 0$, $\rho \leq 0$ y A es una función positiva o negativa con $\lim_{t \rightarrow \infty} A(t) = 0$ y $k = k(n) \rightarrow \infty$, $\frac{k}{n} \rightarrow \infty$, $n \rightarrow \infty$, y

$$\lim_{n \rightarrow \infty} \sqrt{k} A\left(\frac{n}{k}\right) = \lambda \quad (2.87)$$

entonces $\sqrt{k} (\hat{\zeta}_H - \zeta) \xrightarrow{d} N\left(\frac{\lambda}{1-\rho}, \zeta\right)$

Demostración. Se puede escribir la ecuación de segundo orden como

$$\lim_{t \rightarrow \infty} \frac{\frac{x^{-\xi} U(tx)}{U(t)} - 1}{A(t)} = \frac{x^\rho - 1}{\rho}$$

Por hipótesis $\lim_{t \rightarrow \infty} A(t) = 0$, esto es equivalente a

$$\lim_{t \rightarrow \infty} \frac{\log U(tx) - \log U(t) - \xi \log x}{A(t)} = \frac{x^\rho - 1}{\rho}$$

Se aplica la desigualdad dada en el teorema B.2.18 (ver en Haan y Ferreira (2006)): para una función diferencial posible A_0 , con $A_0(t) \sim A(t)$, $t \rightarrow \infty$ y para cada $\varepsilon > 0$ existe un t_0 tal que para $t \geq t_0$, $x \geq 1$,

$$\left| \frac{\log U(tx) - \log U(t) - \xi \log x}{A_0(t)} - \frac{x^\rho - 1}{\rho} \right| \leq \varepsilon x^{\rho+\varepsilon} \quad (2.88)$$

como se prueba en el teorema 2.15 se nota que

$$\hat{\xi}_H = \frac{1}{k} \sum_{i=0}^{k-1} \log U(Y_{n-1,n}) - \log U(Y_{n-k,n}),$$

Donde los Y_i son independientes y tienen una función de distribución común $1 - \frac{1}{x}$, $x \geq 1$ Por lo tanto se trabaja con esta representación para $\hat{\xi}$.

Aplicando (2.88) con $t = Y_{n-k,n} \rightarrow \infty$ a.s. $n \rightarrow \infty$ (Lema 1) y $x = \frac{Y_{n-i,n}}{Y_{n-k,n}}$ y entonces se consigue eventualmente, como en el teorema 2.15.

$$\begin{aligned} \hat{\xi}_H &= \frac{1}{k} \sum_{i=0}^{k-1} \log U(Y_{n-i,n}) - \log U(Y_{n-k,n}) \\ &= \frac{\xi}{k} \sum_{i=0}^{k-1} \log \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right) + A_0(Y_{n-k,n}) \frac{1}{k} \sum_{i=0}^{k-1} \frac{\left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right) - 1}{\rho} \\ &O_\rho(1) |A_0(Y_{n-k,n})| \frac{1}{k} \sum_{i=0}^{k-1} \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right)^{\rho+\varepsilon} \end{aligned}$$

por lo tanto

$$\begin{aligned} \sqrt{k} (\hat{\xi}_H - \xi) &= \xi \sqrt{k} \left(\frac{1}{k} \sum_{i=0}^{k-1} \log \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right) - 1 \right) \\ &+ \sqrt{k} A_0(Y_{n-k,n}) \frac{1}{k} \sum_{i=0}^{k-1} \frac{\left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right)^\rho - 1}{\rho} + O_\rho(1) \sqrt{k} |A_0(Y_{n-k,n})| \frac{1}{k} \sum_{i=0}^{k-1} \left(\frac{Y_{n-i,n}}{Y_{n-k,n}} \right)^{\rho+\varepsilon} \end{aligned}$$

El primer término, normalizado, es asintóticamente normal por el lema 2. Como en la prueba del lema 2 se tiene

$$\frac{1}{k} \sum_{i=0}^{k-1} \frac{\left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right)^\rho - 1}{\rho} = \frac{1}{k} \sum_{i=0}^{k-1} \frac{Y_i^\rho - 1}{\rho} \quad (2.89)$$

Por lo tanto (2.89) se tiene a

$$\frac{E(Y_i^\rho - 1)}{\rho} = (1 - \rho)^{-1}$$

Por la ley de los grandes números.

Similarmente,

$$\frac{1}{k} \sum \left(\frac{Y_{n-i,n}}{Y_{n-k,n}}\right)^{\rho+\varepsilon} \xrightarrow{P} EY_1^{\rho+\varepsilon} = \frac{1}{1 - \rho - \varepsilon}$$

Resta probar que

$$\frac{A_0(Y_{n-k,n})}{A_0\left(\frac{n}{k}\right)} \xrightarrow{P} 1$$

Esto sigue del lema 2.2.3 (ver en Haan y Ferreira (2006)) el hecho que la función $|A_0|$ varía regularmente y las desigualdades de Protter's (Proposición B.1.9 ver en Haan y Ferreira (2006)).

El marco teórico de segundo orden usado en el teorema (2.17) suministra la aproximación más natural a la normalidad asintótica de estimadores como el estimador de Hill. El parámetro ρ controla la velocidad de convergencia para la normalidad asintótica de $\hat{\xi}_H$. Función de distribución F satisface

$$1 - F(x) = c_1 x^{-\frac{1}{\xi}} + c_2 x^{-\frac{1}{\xi} + \frac{\rho}{\xi}} (1 + O(1)) \quad , \quad x \rightarrow \infty$$

y la segunda condición de orden (2.86) con $\tilde{\xi} = 1$ y $\rho = -2$, en efecto es fácil resolver ver si

$$1 - F(x) = c_1 x^{-\frac{1}{\xi}} + c_2 x^{-\frac{1}{\xi} + \frac{\rho}{\xi}} (1 + O(1)) \quad , \quad x \rightarrow \infty$$

para constantes $c_1 > 0$, $c_2 \neq 0$, $\xi > 0$ y $\rho < 0$, entonces, las condiciones de segundo orden (2.86) cumplen con la indicación $\tilde{\xi}$ y ρ con

$$A\left(\frac{1}{1 - F(t)}\right) = \rho \tilde{\xi}^{-1} c_2 c_1^{-1} t^{\frac{\rho}{\tilde{\xi}}}$$

Y por lo tanto (2.85) $A(t) = \rho \zeta^{-1} c_2 c_1^{\rho-1} t^\rho$. Además, si (2.87) se cumple con $\lambda \neq 0$ entonces un simple cálculo muestra que esto es verdad si y solo si

$$k(n) \sim \left(\frac{\lambda \zeta}{\rho c_2} c_1^{1-\rho} \right)^{\frac{2}{1-2\rho}} \eta^{-\frac{2\rho}{1-2\rho}} \quad (2.90)$$

Entonces la tasa de convergencia en el teorema 2.17 es del orden $\eta^{\frac{\rho}{1-2\rho}}$.

Ahora se considerara tres tipos de sucesiones:

1. Suponga $\frac{\sqrt{k}}{|A(\frac{n}{k})|} \rightarrow \infty$ Entonces no es difícil ver usando la desigualdad en la prueba del teorema 2.17 que

$$\frac{\widehat{\zeta}_H - \zeta}{A\left(\frac{n}{k}\right)} \xrightarrow{P} \frac{1}{1-\rho}$$

Ya que para n, k grande debe ser mucho más grande que $n^{-\frac{2\rho}{1-2\rho}}$, se tiene para n grande, $\frac{n}{k}$ mucho más pequeño que

$$n^{\frac{1+2\rho}{1-2\rho}} = n^{\frac{1}{1-2\rho}}$$

por lo tanto la tasa de convergencia $|A\left(\frac{n}{k}\right)|$ es más suave que la tasa $n^{\frac{\rho}{1-2\rho}}$ encontrada después de (2.90).

2. Supongamos que $\sqrt{k}A\left(\frac{n}{k}\right) \rightarrow \infty$. Entonces $k(n) = O\left(n^{-\frac{2\rho}{1-2\rho}}\right)$, y la tasa de convergencia $\frac{1}{\sqrt{k}}$ es otra vez más baja que $\eta^{\frac{\rho}{1-2\rho}}$

3. Supongamos $\sqrt{k}A\left(\frac{n}{k}\right) \rightarrow \lambda \neq 0, \infty$ luego por (2.90) la tasa de convergencia $\frac{1}{\sqrt{k}}$ es del orden $\eta^{\frac{\rho}{1-2\rho}}$. Esta es la situación óptima.

¿Cuál es la mejor selección de λ ? El teorema (2.17) dice que si $\sqrt{k}A\left(\frac{n}{k}\right) \rightarrow \lambda$, entonces

$$\sqrt{k}\left(\widehat{\zeta}_H - \zeta\right) \xrightarrow{d} \zeta N + \frac{\zeta}{1+\rho} \quad (2.91)$$

Con N la normal estándar y por lo tanto

$$\widehat{\zeta}_H - \zeta \approx^d \frac{\zeta N}{\sqrt{k}} + \frac{\zeta}{(1-\rho)\sqrt{k}} \approx^d \frac{\zeta N}{\sqrt{k}} + \frac{A\left(\frac{n}{k}\right)}{1-\rho} \quad (2.92)$$

Para que valor de $k = k(n)$ esta aproximación es la mejor, es decir para que k su error cuadrado medio

$$\frac{\bar{\zeta}^2}{k} + \frac{A^2\left(\frac{n}{k}\right)}{(1-\rho)^2} \quad (2.93)$$

Es mínimo. Para que el tiempo sea continuo se considera el caso especial

$$A(t) = ct^\rho.$$

Escribiendo $r := \frac{n}{k}$. Esto conduce a

$$\arg \min_{\left(\frac{n}{r}\right)=1,2,\dots} \left(\frac{r\bar{\zeta}^2}{n} + \frac{c^2 r^{2\rho}}{(1-\rho)^2} \right)$$

Y para simplificar se considera

$$\arg \min_{t>0} \left(\frac{r\bar{\zeta}^2}{n} + \frac{c^2 r^{2\rho}}{(1-\rho)^2} \right)$$

El ínfimo se alcanza igualando la derivada a cero, es decir

$$\frac{\bar{\zeta}^2}{n} = \frac{-2\rho c^2 t^{2\rho-1}}{(1-\rho)^2}, \quad \text{o} \quad t = \left(\frac{\bar{\zeta}^2(1-\rho)^2}{-2\rho c^2} \right)^{\frac{1}{2\rho-1}} n^{\frac{1}{1-\rho}}$$

Equivalentemente, colocando $t = r = \frac{n}{k}$

$$r = \left\lceil \left[\left(\frac{\bar{\zeta}^2(1-\rho)^2}{-2\rho c^2} \right)^{\frac{1}{2\rho-1}} n^{-\frac{2\rho}{1-2\rho}} \right] \right\rceil$$

donde $\lceil x \rceil$ significa la parte entera de x . Por lo tanto

$$k_0(n) = \left\lceil \left[\left(\frac{\bar{\zeta}^2(1-\rho)^2}{-2\rho c} \right)^{\frac{1}{2\rho-1}} n^{-\frac{2\rho}{1-2\rho}} \right] \right\rceil \quad (2.94)$$

Y $k_0(n)$ es la selección óptima para la secuencia $k(n)$ bajo las condiciones dadas.

Ahora se considerara el mínimo de la siguiente expresión $E\left(\hat{\zeta}_H - \bar{\zeta}\right)^2$ o sustituyéndola por la siguiente

$$E\left(\frac{\bar{\zeta}N}{\sqrt{k}} + \frac{A\left(\frac{n}{k}\right)}{1-\rho}\right)^2 \quad (2.95)$$

Y la sucesión k_0 que optimiza (2.95) servirá como la selección óptima para el estimador $\widehat{\xi}_H$ también. Se nota que

$$\lim_{n \rightarrow \infty} \sqrt{k_0} A \left(\frac{n}{k_0} \right) = \lim_{n \rightarrow \infty} cn^\rho k^{2-\rho} = \frac{\text{sign}(c)\xi(1-\rho)}{\sqrt{-2\rho}} \quad (2.96)$$

Donde $\text{sign}(c) = 1$ si $c > 0$ y $\text{sign}(c) = -1$ si $c < 0$. Por lo tanto para esta selección de k se tiene

$$\sqrt{k_0} \left(\widehat{\xi}_H - \xi \right) \xrightarrow{d} N \left(\frac{\text{sign}(c)\xi}{\sqrt{-2\rho}}, \xi^2 \right)$$

y

$$\lim_{n \rightarrow \infty} k_0 \text{mín} E \left(\frac{\xi N}{\sqrt{k}} + \frac{A \left(\frac{n}{k} \right)}{1-\rho} \right)^2 = \lim_{n \rightarrow \infty} k_0 \text{mín} \left(\frac{\xi^2}{k} + \frac{A^2 \left(\frac{n}{k} \right)}{(1-\rho)^2} \right)$$

por (2.94), la cual es igual

$$\lim_{n \rightarrow \infty} k_0 \left(\frac{\xi^2}{k} + \frac{A^2 \left(\frac{n}{k} \right)}{(1-\rho)^2} \right) = \xi^2 + \frac{1}{(1-\rho)^2} \left(\frac{\text{sign}(c)\xi(1-\rho)}{\sqrt{-2\rho}} \right)^2 = \xi^2 \left(1 - \frac{1}{2\rho} \right)$$

Por (2.93)

Anteriormente solamente se ha considerado el caso especial $A(t) = ct^\rho$ con $\rho < 0$. Esto es lo que a menudo se asume en aplicaciones de la teoría de valor extremo. Sin embargo esta simplificación no es factible en el caso de que $\rho = 0$. Próximamente se considerará el problema de optimizar en el caso más general de condición de segundo orden no solamente el caso especial $A(t) = ct^\rho$.

Como se convertirá en el final si una secuencia $k_0(n)$ es óptimo entonces cualquier sucesión $k(n) = k_0(n)$ cuando $n \rightarrow \infty$ es también óptimo. Esto implica que se puede reemplazar la función A por cualquier función A^* con $A^*(t) \sim A(t)$, cuando $t \rightarrow \infty$ sin pérdida de generalidad.

Similarmente como antes se puede hallar

$$\arg \text{mín}_{t>0} \left(\frac{t\xi^2}{n} + \frac{A^2(t)}{(1-\rho)^2} \right) \quad (2.97)$$

Donde la función $|A|$ esta variando regularmente con índice $\rho \leq 0$. Para $\rho = 0$ es razonable asumir que existe un función A^* con $|A^*(t)| \sim |A(t)|$ cuando $n \rightarrow \infty$ es

monótona decreciente. En este caso se puede asumir sin pérdida de generalidad que la función A^2 satisface

$$\lim_{t \rightarrow \infty} \frac{A^2(tx) - A^2(t)}{q(t)} = -\log x, \quad x > 0$$

Con q una función deseablemente positiva, Entonces para cada $\rho \leq 0$ existe una función decreciente positiva $s \in RV_{2\rho-1}$ tal que $n \rightarrow \infty$

$$A^2(t) \int_1^\infty s(u) du \quad (2.98)$$

Se tiene para $c > 1$ y t suficientemente grande

$$\frac{t\tilde{\xi}}{n} + \frac{c^{-1}}{(1-\rho)^2} \int_t^\infty s(u) du < \frac{t\tilde{\xi}^2}{n} + \frac{A^2(t)}{1-\rho^2} \frac{t\tilde{\xi}^2}{n} + \frac{c}{(1-\rho)^2} \int_t^\infty s(u) du \quad (2.99)$$

El ínfimo sobre $t > 0$ para los lados derecho e izquierdo puede calcularse la derivada igualando a cero. Para el lado derecho se tiene

$$\frac{\tilde{\xi}^2(1-\rho)^2}{cn} = s(t)$$

lo cual es equivalente a

$$t = s^{\leftarrow} \left(\frac{\tilde{\xi}^2(1-\rho)^2}{cn} \right)$$

y el ínfimo es

$$\begin{aligned} & \frac{\tilde{\xi}^2}{n} s^{\leftarrow} \left(\frac{\tilde{\xi}^2(1-\rho)^2}{cn} \right) + \frac{c}{(1-\rho)^2} \int_\rho^\infty s(u) du \quad p = s^{\leftarrow} \left(\frac{\tilde{\xi}^2(1-\rho)^2}{cn} \right) \\ &= \frac{c}{(1-\rho)^2} \left\{ \frac{\tilde{\xi}^2(1-\rho)^2}{cn} s^{\leftarrow} \left(\frac{\tilde{\xi}^2(1-\rho)^2}{cn} \right) + \int_\rho^\infty s(u) du \right\} \\ &= \frac{c}{(1-\rho)^2} \int_0^r s(u) du \end{aligned}$$

donde en la última etapa se usó

$$v s^{\leftarrow}(v) + \int_{s^{\leftarrow}(v)}^\infty s(u) du = \int_0^v s^{\leftarrow}(u) du \quad (2.100)$$

Por el lado izquierdo de (2.99) se tiene el mismo resultado pero reemplazando ac por c^{-1} se sigue que el ínfimo (2.97) es

$$\frac{1}{(1-\rho)^2} \int_0^r s^{\leftarrow}(u) du$$

Y se alcanza cuando

$$t \sim s^{\leftarrow} \left(\frac{\zeta^2(1-\rho)^2}{n} \right)$$

Es decir, (ya que t se reemplaza por $\frac{n}{k}$)

$$k(n) \sim \frac{n}{s^{\leftarrow} \left(\frac{\zeta^2(1-\rho)^2}{n} \right)}$$

Ya que una secuencia óptima $k_0 = k_0(n)$ en el sentido de minimizar

$$\frac{\zeta^2}{k} + \frac{A^2 \left(\frac{n}{k} \right)}{(1-\rho)^2}$$

Es dado por

$$k_0 = \left[\frac{n}{s^{\leftarrow} \left(\frac{\zeta^2(1-\rho)^2}{n} \right)} \right]$$

¿Qué puede decirse acerca de la distribución asintótica de $\sqrt{k_0}(\widehat{\zeta}_H - \zeta)$?

$$\sqrt{k_0}(\widehat{\zeta}_H - \zeta) \xrightarrow{d} \zeta N + \frac{\sqrt{k_0} A \left(\frac{n}{k_0} \right)}{1-\rho}$$

Así que se tiene que evaluar $\sqrt{k_0} A \left(\frac{n}{k_0} \right)$ para n grande. Por (2.98) y (2.100)

$$\begin{aligned} k_0 A^2 \left(\frac{n}{k_0} \right) &\sim \frac{n}{s^{\leftarrow} \left(\frac{\zeta^2(1-\rho)^2}{n} \right)} \left\{ \int_{\rho}^{\infty} s(u) du \right\} \\ &= \zeta^2(1-\rho)^2 \left\{ \int_0^m s^{\leftarrow}(u) du - \frac{\zeta^2(1-\rho)}{n} s^{\leftarrow} \left(\frac{\zeta^2(1-\rho)^2}{n} \right) \right\} \\ &= \zeta^2(1-\rho)^2 \left\{ \frac{\int_0^m s^{\leftarrow}(u) du}{\frac{\zeta^2(1-\rho)^2}{n} s^{\leftarrow} \left(\frac{\zeta^2(1-\rho)^2}{n} \right)} - 1 \right\} \end{aligned}$$

Ahora por el teorema B.1.5 (ver en Haan y Ferreira (2006)) ya que

$$\left(\frac{1}{s} \right)^{\leftarrow} \in RV_{\frac{1}{(1-2\rho)'}}$$

$$\lim_{t \rightarrow \infty} \frac{\int_0^{\frac{1}{x}} s^{\leftarrow}(u) du}{\frac{1}{x} s^{\leftarrow} \left(\frac{1}{x} \right)} = \lim_{t \rightarrow \infty} \frac{\int_x^{\infty} \left(\frac{1}{s} \right)^{\leftarrow}(u) \frac{du}{u^2}}{\frac{1}{x} \left(\frac{1}{s} \right)^{\leftarrow}(x)} = \frac{1-2\rho}{-2\rho} \quad (2.101)$$

Se debe notar que $s^{\leftarrow} \left(\frac{1}{x} \right) = \left(\frac{1}{s} \right)^{\leftarrow} (x)$. Por lo tanto para $\rho < 0$

$$\lim_{n \rightarrow \infty} k_0 A^2 \left(\frac{n}{k_0} \right) = \xi^2 (1 - \rho)^2 \left(\frac{1 - 2\rho}{-2\rho} - 1 \right) = \frac{\xi^2 (1 - \rho)^2}{-2\rho}$$

Y

$$\sqrt{k_0} (\hat{\xi}_H - \xi) \longrightarrow N \left(\frac{\text{sign}(A)\xi}{\sqrt{-2\rho}}, \xi^2 \right)$$

Para $\rho = 0$ el límite en (2.101) debe ser interpretado como infinito. Esto significa que para $\rho = 0$ minimizando el error cuadrado medio conseguimos una secuencia óptima k_0 para el cual

$$\sqrt{k_0} (\hat{\xi}_H - \xi) + b_n \xrightarrow{d} N(0, \xi^2)$$

Donde b_n es una sucesión que varía suavemente tendiendo a más o menos infinito. Esta afirmación no es común para obtener intervalo de confianza para ξ . Todo lo que se puede decir es que

$$\frac{\sqrt{k_0}}{b_n} (\hat{\xi}_H - \xi) \longrightarrow 1$$

Si $\rho = 0$ se toma la sucesión $k(n)$ a bit más pequeño, no se logra una normalidad asintótica. Tomamos $k_\lambda := k_\lambda(n)$ tal que

$$\lim_{n \rightarrow \infty} k_\lambda A^2 \left(\frac{n}{k_\lambda} \right) = \lambda^2 > 0 \quad (2.102)$$

Si escribimos

$$f(t) = \frac{\lambda^2 t}{\int_t^\infty s(u) du}$$

con s como en (2.98). Entonces la expresión anterior es creciente y es RV_1 además se tiene

$$n \sim f \left(\frac{n}{k_\lambda} \right)$$

En contraste con el caso $\rho < 0$ se tiene para la selección óptima

$$n \sim \frac{\xi^2 (1 - \rho)^2}{s \left(\frac{n}{k_0} \right)}$$

Ahora, las funciones

$$f(t) = \frac{\lambda^2 t}{\int_t^\infty s(u) du} \quad \text{y} \quad \frac{1}{s(t)}$$

son ambos RV_1 pero por el teorema B.1.5. (ver en Haan y Ferreira (2006))

$$\lim_{t \rightarrow \infty} s(t)f(t) = 0$$

Propiedades Asintóticas

El estimador de Hill es asintóticamente normal. Esto es, para una muestra independiente e idénticamente distribuida de tamaño n y con un $k = K(n)$ una sucesión intermedia

$$\sqrt{k}(\hat{\xi} - \xi) \approx \sqrt{\text{var}_{\xi}}N + \lambda b_{\xi, \rho}$$

Con N normal estándar, donde las constantes λ , var_{ξ} y $b_{\xi, \rho}$ son conocidos (ver teorema 2.17) para el estimador de Hill.

El estimador de Hill tiene sistemáticamente la varianza asintótica más pequeña en su rango de posibles valores. Los estimadores de Hill y de momentos tienen la varianza asintótica más pequeña para valores positivos de ξ .

Los estimadores de Máxima Verosimilitud y de Hill Negativo tienen la varianza asintótica más pequeña para valores negativos de ξ .

Es más complicado comparar el sesgo de los estimadores, ya que en general el sesgo depende de ambos parámetros ξ y ρ entre otras características de la distribución subyacente.

Para determinar una secuencia óptima $k_0(n)$ se considera la representación

$$\hat{\xi} - \xi \approx \sqrt{\text{var}_{\xi}} \frac{N}{\sqrt{k}} + \frac{\lambda b_{\xi, \rho}}{\sqrt{k}} \approx \sqrt{\text{var}_{\xi}} \frac{N}{\sqrt{k}} + b_{\xi, \rho} A\left(\frac{n}{k}\right)$$

Con $k = k(n)$ una sucesión intermedia tal que $\sqrt{k}A\left(\frac{n}{k}\right) \rightarrow \lambda \in R$ normal estándar. Por simplicidad se supone $A(t) = ct^{\rho}$ para alguna constante c real y $\rho < 0$ buscamos una secuencia k para la cual

$$\frac{\text{var}_{\xi}}{k} + b_{\xi, \rho}^2 c^2 \left(\frac{n}{k}\right)^2 \text{ es mínimo}$$

Se obtiene

$$k_0(n) = \left[\left(\frac{\text{var}_{\xi}}{-2\rho c^2 b_{\xi, \rho}^2} \right)^{\frac{1}{1-2\rho}} n^{-\frac{2\rho}{1-2\rho}} \right]$$

Donde $\llbracket x \rrbracket$ denota la parte entera de x .

Luego para esta selección de k se tiene

$$\sqrt{k_0}(\hat{\xi} - \xi) \xrightarrow{d} N\left(\frac{\text{sign}(c)\sqrt{\text{var}_\xi}}{\sqrt{-2\rho}}, \text{var}_\xi\right)$$

Y

$$\lim_{n \rightarrow \infty} k_0 \min E\left(\frac{\text{var}_\xi N}{\sqrt{k}} + \frac{A\left(\frac{n}{k}\right)}{1-\rho}\right)^2 = \text{var}_\xi \left(1 - \frac{1}{2\rho}\right)$$

Se observa que con los supuestos dados una comparación de la calidad de los diferentes estimadores se reduce a una comparación de la varianza asintótica.

De la teoría asintótica uno obtiene intervalos de confianza correspondientes. La aproximación más común es asumir

$$\hat{\xi} - Z_{\frac{\alpha}{2}} \sqrt{\frac{\text{var}_\xi}{k}} < \xi < \hat{\xi} + Z_{\frac{\alpha}{2}} \sqrt{\frac{\text{var}_\xi}{k}}$$

Donde var_ξ es la respectiva varianza asintótica con ξ reemplazada por su estimador y $Z_{\frac{\alpha}{2}}$ es el cuantil $1 - \frac{\alpha}{2}$ para la distribución normal estándar.

Si formulamos la hipótesis Nula y alterna se tiene:

$$H_0 : \xi \geq 0 \quad \text{versus} \quad H_a : \xi < 0$$

La hipótesis nula es rechazada si para un nivel de significancia del 5%,

$$\hat{\xi}_P < \frac{-2.96}{\sqrt{k}}$$

para el estimador de Pickands.

CAPÍTULO 3

MÉTODO DE DISTRIBUCIÓN DE PÉRDIDA AGREGADA (LDA)

3.1 Introducción

El objetivo central del método LDA⁸ es hallar una función de las pérdidas agregadas, $S(x)$ para cada línea de negocio y tipo de riesgo combinando dos modelos: El modelo de la distribución de probabilidad de la frecuencia de las pérdidas y el modelo de la distribución de probabilidad de la severidad a través de métodos entre otros como el algoritmo de Recursión de Panjer, la aproximación Analítica de Bocker y Kluppelberg, o la simulación de Montecarlo.

Una vez se tiene la función de distribución de las pérdidas agregadas $S(x)$ se obtiene la distribución de probabilidad de esta variable, $G_S(x)$ la cual nos permite determinar la carga de capital por riesgo operacional, $\text{OpVar}(1 - \alpha)$ con la aplicación del concepto VaR valor en Riesgo para el cual se puede presentar dos casos según la situación de la empresa:

- $\text{OpVar}(1 - \alpha) = G_S^{-1}(1 - \alpha)$ para un nivel de confianza $1 - \alpha$ determinado, que Basilea II ha fijado en 99,9% donde G_S es la función de distribución de las pérdidas agregadas.
- Cuando la entidad demuestra hacer previsiones para las pérdidas esperadas, la carga de capital se obtiene restando las pérdidas esperadas del percentil $1 - \alpha$. Así que

$$\text{OpVaR}(1 - \alpha) = G_S^{-1}(1 - \alpha) - E(S)$$

⁸El método LDA (Loss Distribution Approach) es un método aplicado en el campo actuarial y combina la distribución de probabilidad de la severidad y la distribución de probabilidad de la frecuencia del evento y uno de los métodos a través del cuál se puede llevar a cabo esta operación es la Simulación de Montecarlo. Es un método muy usado en la industria de seguros.

Definición 3.1 (OpVaR). La carga de capital es el percentil 99,9 % de la distribución de probabilidad de Pérdida Agregada $1 - G_S = P\{S > \text{OpVaR}\} = 0,1 \%$

Definición 3.2 (OpVaR-pérdidas inesperadas). Este es el OpVaR anterior menos las pérdidas esperadas. Basilea acepta esta definición siempre y cuando la entidad haya provisionado las pérdidas esperadas.

$$P\{S > \text{OpVaR} - E(S)\} = 0,1 \%$$
 (3.1)

Donde $E(S)$ son las pérdidas totales esperadas, o sea:

$$E(S) = E \left[\sum_{i=0}^N X_i \right]$$
 (3.2)

En estas definiciones se está asumiendo un percentil del 99,9 % pero podría calcularse teniendo en cuenta otro percentil en algunos de los métodos que se han reportado en la literatura. En nuestro caso se usará el método de Hill y la gráfica de los excesos promedio según los datos que se tengan⁹.

El cálculo de las pérdidas agregadas se obtiene con base a la distribución de frecuencias de las pérdidas $p_n = P(N = n)$ y con la función de probabilidad de las pérdidas y la función de probabilidad de la severidad económica de las pérdidas $f_X(X)$ o con función de distribución acumulada $F_X(X) = P(X \leq x)$, entonces la n -ésima convolución de la distribución de severidad denotada por $F_X^{*n}(X)$, está dada por:

$$P(X_1 + X_2 + \dots + X_n \leq x) = F^*F^* \dots F^*(x) = F_X^{*n}(X)$$

de este modo la función de distribución de las pérdidas agregadas es:

$$G_S(X) = P(S \leq x) = \sum_{n=0}^{\infty} P(N = n)F_X^{*n}(X).$$

Donde $F_X^{*n}(X)$ se refiere a la n -ésima convolución de F_X consigo mismo, por ejemplo

$$F_X^{*n}(x) = P(X_1 + \dots + X_n \leq x)$$

la función de distribución de la suma de n variables aleatorias independientes con la misma distribución que X . Como se menciono antes, la formula anterior puede difícilmente ser resuelta analíticamente. Se debe confiar en aproximaciones, expansiones, recursiones o algoritmos numéricos.

⁹En este documento el símbolo α representa el nivel de significancia de una prueba estadística y $1 - \alpha$ el Nivel de confianza.

En resumen, las pérdidas operacionales agregadas pueden ser vistas como una suma (S) de un número aleatorio (N) de pérdidas operacionales individuales (X_1, \dots, X_N) Esta suma se puede presentar por:

$$S = X_1 + X_2 + \dots + X_N = \sum_{i=1}^N X_i ; \quad N = 1, 2, \dots \quad (3.3)$$

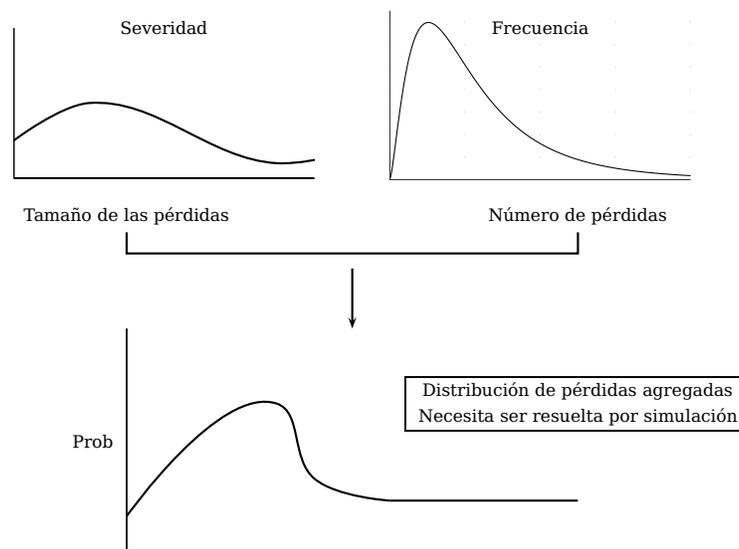
El modelo de riesgo colectivo asume que las pérdidas X_i son variables aleatorias independientes e igualmente distribuidas representan las pérdidas individuales. En este modelo la distribución de N (Frecuencia) es independiente de los valores de X_i (severidad).

El modelo de riesgo individual caracteriza el agregado como una suma de un número fijo de perdidas. Tales modelos son usados en seguros para la suma de las pérdidas de un contrato particular. En el riesgo operacional ellas pueden ser usadas en términos de un número fijo de tipo de pérdidas particulares.

La separación de los procesos de severidad y frecuencia es importante debido a que esto permite por ejemplo, un mejor entendimiento del efecto en la distribución agregada al cambiar el nivel del umbral.

Al tener calculado separadamente ambos procesos la severidad y la frecuencia, lo que se necesita es combinarlos dentro de una distribución de pérdida agregada que nos permita predecir una curva para la pérdida operacional con un grado de confianza (Ver figura 3.1).

Figura 3.1. Modelos agregados de severidad y de frecuencia



Si la frecuencia de riesgos operativos es muy grande, se debe suponer que el teorema central del límite es adecuado. Como una aproximación a gran escala:

$$F_{X(t)}(x) \approx \phi \left(\frac{x - EX(t)}{\sqrt{Var x(t)}} \right) \quad (3.4)$$

Donde $\phi(x)$ denota la función de distribución normal estándar la cual debe ser factible. Independientemente estas aproximaciones a gran escala a menudo han sido probadas no confiables por investigadores actuariales. Prácticos han tratado de evitar esta deficiencia aplicando versiones refinadas del teorema del límite central. Los ejemplos son aplicaciones de la expansión de Edgeworth y series de Gram-Charlier, pero ellas no siempre resultan en mejoras satisfactorias.

Es necesario entender que el cálculo de la pérdida total se define como la sumatoria de las pérdidas de cada una de las celdas de la matriz de 7×8 la cual se representa por la siguiente ecuación:

$$L = \sum_{i=1}^7 \sum_{j=1}^8 S_{ij}$$

Donde S_{ij} es la pérdida de la celda i, j de las matriz de pérdidas. Cada una de las celdas i, j se calculan de la siguiente manera:

$$S_{ij} = \sum_{N=1}^n X_{N_{ij}}$$

Donde N_{ij} es la variable aleatoria que representa el número de eventos de riesgo en la celda i, j (frecuencia de los eventos) y X_N es el monto de la pérdida en la celda i, j (severidad del evento). Por lo tanto las pérdidas son resultado de dos variables aleatorias la frecuencia y la severidad.

3.2 Distribuciones de frecuencias

El número de eventos de riesgo operacional, N por periodo es una variable aleatoria que sigue una distribución de probabilidad discreta. La pregunta que nos hacemos es cuál es la distribución de probabilidad de los eventos por periodo es decir la frecuencia los eventos. La hemos denominado $p(n) = P(N = n)$ siendo n un entero no negativo.

En la literatura especializada de riesgos operativos se consideran diferentes distribuciones de frecuencia de fallas en sistemas y procedimientos. La distribución

de Poisson, La distribución binomial, La distribución binomial negativa, la hipergeométrica, la Geométrica, Polya–Aeppli (Poisson–Geométrica).

Es necesario destacar que se pueden identificar las siguientes distribuciones discretas: Poisson, Binomial y Binomial negativas con los siguientes criterios:

Si $\mu = \sigma^2$ los datos distribuyen como una Poisson.

Si $\mu > \sigma^2$ los datos se distribuyen como una Binomial.

Si $\mu < \sigma^2$ los datos se distribuyen como una Binomial negativa.

La prueba bondad de ajuste más común es la chi cuadrado.

3.3 Distrubución de severidad

Una etapa importante en la aplicación del método LDA es saber cuál es la distribución de probabilidad de la severidad o pérdidas económicas de los eventos operacionales durante un determinado periodo según datos históricos, es decir tratar de ajustar una distribución de probabilidad continua a los datos y además estimar sus parámetros.

Las distribuciones más comunes para ajustar a la severidad son: *La normal, Lognormal, normal inversa, Exponencial, Weibull, Pareto, Gamma, Cauchy, Rayleigh, Logística, Loglogística.*

La distribución a escoger puede ser de colas pesadas o no lo cual determina el tipo de distribución que se debe escoger. Las distribuciones más comunes de colas pesadas son: La distribución generalizada de valor extremo y la distribución generalizada de Pareto.

Un método a seguir es ajustar una distribución al cuerpo de la distribución y otra distribución a la cola y hallar el VaR al cuerpo de la distribución y el VaR a la distribución de la cola.

La pruebas de bondad de ajuste más comunes son: *Kolmogorov Smirnov, Anderson Darling, Cramer–Von Mises.*

Cuantificación de las pérdidas agregadas

El método que se usará para las pérdidas agregadas es el método de simulación de Montecarlo¹⁰

¹⁰Las pérdidas agregadas se pueden calcular además de por el Método de Montecarlo por el Algoritmo de Recursión de Panger, y la aproximación Analítica de Bocker y Kluppelberg.

El procedimiento que se sigue es el siguiente:

1. Se generan una valor de $N = 1000$ hasta 10000 con incrementos de 1000, después de 10.000 con incrementos de 10000 hasta 100000 y de 100000 en adelante con incrementos de 100000 hasta un 1000000. Para un total de 28 simulaciones. Cada valor que se simule corresponde al número de eventos que producen pérdidas económicas.
2. Según el número de eventos que se generen en cada iteración anterior se generan el número de pérdidas económicas para cada uno. Cada una de las pérdidas se denotará por $L_1, L_2, L_3, \dots, L_n$.
3. Una vez se tienen las pérdidas se suman obteniéndose los valores S_1, S_2, \dots, S_n que representan las pérdidas agregadas por período.
4. Se calcula el percentil 99,9 % de la distribución de pérdidas agregadas.
5. Se obtiene la media de los valores de S para calcular las pérdidas esperadas (EL).
6. La carga de capital se calculará de la siguiente manera: Si la empresa no hace reservas es el OpVaR (99,9 %) y si la empresa tiene reservas al OpVaR (99,9 %) se le resta las pérdidas esperadas.

En el Anexo D se presenta un código del algoritmo implementado en MatLab para estimar las pérdidas agregadas aplicando la simulación de Montecarlo. Una vez calculada la distribución de las pérdidas agregadas se obtiene el OpVaR ($1 - \alpha$)

3.4 Valor en riesgo de mercado y operacional

Los modelos VaR (del término en inglés Value at Risk) comenzaron a desarrollarse a inicios de la década del 90, son actualmente considerados la medida estándar para el riesgo de mercado y son usados ampliamente en la administración del riesgo. En 1994, el banco estadounidense JP Morgan propuso, en su documento técnico denominado RiskMetrics¹¹, el concepto de valor en riesgo como modelo para medir cuantitativamente los riesgos de mercado en instrumentos financieros o portafolios con varios tipos de instrumentos. El valor en riesgo (VaR) es un modelo estadístico basado en la teoría de la probabilidad.

¹¹Una descripción completa de esta metodología se encuentra en J.P. Morgan.

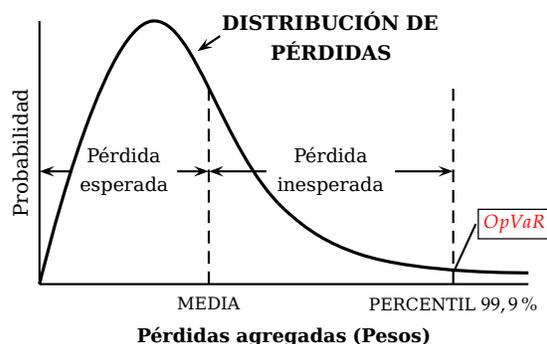
El VaR resume en un solo número la pérdida potencial máxima que se puede sufrir en una posición de riesgo dado un nivel de confianza elevado (usualmente 95 %, 99 % o 99.9 %) en un periodo de tiempo determinado.

Es importante destacar que la definición de valor en riesgo es válida únicamente en condiciones normales de mercado, ya que en momentos de crisis y turbulencia la pérdida esperada se define por pruebas de valores extremos.

Desde un punto de vista del riesgo de mercado el VaR mide la pérdida esperada máxima a precios de mercado de un portafolio, dado que puede esperarse que ocurra hasta que la posición pueda ser neutralizada. En otras palabras, en el riesgo de mercado, el VaR calcula una pérdida extrema eventual resultado de mantener un portafolio por un periodo determinado usando como medida de riesgo la volatilidad de los precios de un activo en los n últimos días. Más precisamente, pérdida extrema es el cuantil X_p para un porcentaje $100(1 - \alpha)$ % de la distribución¹². Como una consecuencia, el VaR estima el cuantil X_p para valores de α suficientemente pequeños.

El Valor en riesgo (VaR) resume la pérdida máxima esperada (o peor pérdida) a largo de un horizonte de tiempo objetivo dentro de un intervalo de confianza dado¹³.

Figura 3.2. Distribución de pérdidas: esperadas e inesperadas



Fuente: Elaboración propia del autor

Una aproximación similar basada en los mismos principios fundamentales de la administración del riesgo de mercado puede ser aplicado al riesgo operativo, desarrollando un VaR operacional el cual es estimando por combinación de los modelos de severidad y de frecuencia.

¹²Bajo en este contexto indican la máxima pérdida posible.

¹³Algunas comparaciones entre la simulación histórica y otros métodos para calcular el VaR se pueden encontrar en Mahoney (1996) y Danielsson y de Vries (2000).

Hay dos diferencias fundamentales entre los modelos de riesgo de mercado y operacional. La primera está relacionada con el hecho de que la aplicación de la teoría de valor extremo EVT (u otras distribuciones de colas pesadas) nos permiten relajar la hipótesis Gaussiana en la cual los modelos del VaR de mercado están basados. Dicho de otra manera, los procesos estocásticos subyacentes a las pérdidas operacionales no se explican por una distribución normal como se podría esperar en el riesgo de mercado (aunque este supuesto también es a menudo cuestionado de igual forma en el riesgo de mercado). La segunda diferencia principal es que los modelos VaR de mercado no tienen que ver con la frecuencia del evento sino que asume solamente que los precios de los activos siguen un proceso estocástico continuo. Hay siempre una cotización disponible del precio del activo mientras el mercado esté abierto. Las pérdidas operativas siguen procesos estocásticos discretos. Esto significa que los eventos pueden ser contados en un cierto periodo de tiempo, es decir un determinado evento operacional puede suceder n veces por día, lo cual no tiene sentido en la medición de riesgo de mercado. También nosotros medimos pérdidas observadas no cambios en el valor MTM (marca de mercado).

Una aproximación más simple establece que las reclamos totales esperados = frecuencia de los eventos \times severidad de la pérdida (o desde un punto de vista del riesgo reclamos totales inesperados = frecuencia de los eventos que están por encima de un umbral \times severidad de la pérdida).

Una de las más grandes ventajas del VaR es que trata de poner dentro de una sola cifra, fácil de comprender el riesgo de un banco basado en las variables de mercado financiero. En pocas palabras el VaR resume la pérdida máxima esperada sobre un horizonte objetivo de tiempo dentro de un intervalo de confianza dado.

En la medición del VaR de mercado se necesita definir: un intervalo de confianza, un horizonte de tiempo (10 días, 1 día, etc.), medir la volatilidad de los factores de riesgo (tasa de interés a seis meses, una acción particular, etc.) y saber la posición de cada activo¹⁴.

En el mercado financiero que los precios sean continuos significa que se puede tener una cotización del precio continuamente en la medida que el mercado esté abierto. En términos probabilísticos los eventos en general seguirán un proceso

¹⁴Existen otras formas de calcular el VaR usando métodos de simulación de Montecarlo, una descripción sobre estas metodologías se puede encontrar en McNeil et al. (2004).

estocástico conocido como movimiento browniano.

Por otro lado, en el riesgo operativo los eventos siguen un patrón discreto. Los procesos estocásticos en los cuales el riesgo operativo está basado son un proceso de Poisson, procesos Cox, etc. Se va a medir que tan grande es el riesgo agregado para un horizonte de tiempo y un cierto nivel de confianza.

Otra diferencia notable es que en el VaR de mercado se puede saber cómo cambia eventualmente el precio de los factores del riesgo que podrían cambiar completamente el VaR. En el riesgo operativo como los factores que cambian son exógenos al sistema se necesita un modelo auxiliar para hacer pruebas de valores extremos.

En la ciencia actuarial hay dos modelos básicos de riesgo para estimar la pérdida total para un periodo: Los modelos de riesgo individual y colectivo. El esquema puede ser válido también para riesgo operacional con unas pocas adaptaciones. Los detalles de tales modelos pueden ser revisados en Klugman et. Al. (1997).

CAPÍTULO 4

APLICACIÓN

4.1 Introducción

Una de las fases más importantes en la implementación de una metodología es la captura de datos confiables y un tamaño de muestra lo suficientemente representativo, este es uno de los primeros aspectos que debe encarar quien intenta medir el riesgo operativo. Según Basilea II las estimaciones de riesgo operativo generadas internamente y utilizadas para efecto de capital regulatorio, deberán basarse en un período de observación de por lo menos cinco años, excepto cuando el banco utilice por primera vez uno de los métodos de medición avanzada (AMA), caso en el cual se aceptará un período de observación de tres años. Los datos que se van a usar provienen de las quejas que hicieron los usuarios del sector financiero y que son remitidas a la Superfinanciera de Colombia en los últimos cinco años (con una periodicidad mensual), estas quejas fueron clasificadas dentro de los siete tipos de eventos de riesgo operativo que estableció la circular externa 041 de 2007 de la Superfinanciera. Los eventos que definieron fueron: Fraude Interno, Fraude Externo, Relaciones Laborales, Clientes, Daños a Activos Fijos, Fallas Tecnológicas y Ejecución y administración de procesos.

Los datos de la severidad de los eventos se recopilaron de una entidad financiera del Sector Cooperativo, es decir la pérdida económica de cada uno de los eventos teniendo en cuenta los riesgos operativos de esta entidad en la línea de la Banca minorista¹⁵.

Se privilegiara aquel evento que implique mayor cantidad de pérdidas económicas con el fin de que la entidad financiera controle de una manera más rigurosa este tipo de evento y así logre tener menos cantidad de pérdidas económicas.

¹⁵La empresa de la cual se ha extraído la información no se publica para respetar la confidencialidad de los datos.

Una vez recopilados los datos se realiza el análisis exploratorio para calcular las medidas de localización, de variabilidad, de asimetría y curtosis y aquellos gráficos que nos permitan identificar los aspectos más fundamentales del proceso generador de datos asociado a las variables estudiadas, y así poder tomar las decisiones para la implementación de los dos métodos: Pérdidas Agregadas, y Teoría Básica del Valor Extremo. En la Teoría Básica del Valor Extremo las medidas de asimetría y curtosis son más importantes que las medidas de tendencia central por lo tanto se le darán más importancia a estas ya que a través de ellas es que se detecta distribuciones con colas pesadas.

La aplicación de las pérdidas agregadas implicara desarrollar el cálculo de las pérdidas esperadas e inesperadas del cuerpo de la distribución como si supusiéramos que el comportamiento de las colas no fuera importante (aplicación de la teoría clásica de Riesgo).

La Teoría Básica del Valor Extremo consiste en un conjunto de técnicas estadísticas para la identificación y modelamiento de los máximos o mínimos de una variable aleatoria, es una disciplina estadística que desarrolla un conjunto de modelos y métodos tanto paramétricos como no paramétricos con objeto de describir, cuantificar y modelar los casos raros. No se puede asumir que una sola distribución de probabilidad explique todo el comportamiento de la variable aleatoria pérdidas agregadas sino que también se requiere ajustar otra distribución de probabilidad para los valores de la cola.

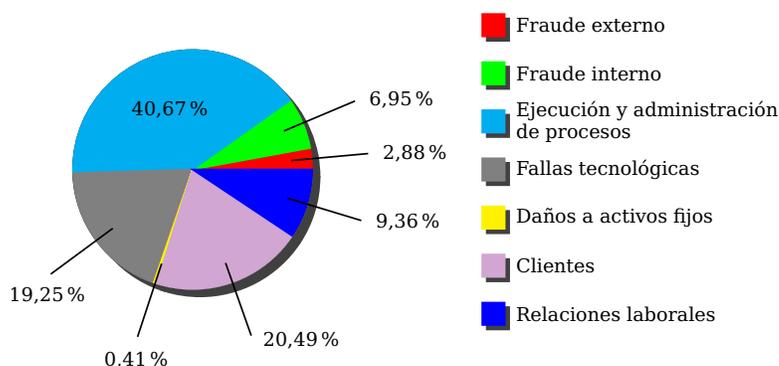
Se aplicará la simulación para darle solución parcial a la escasez de datos especialmente se hará con el fin de tener suficientes valores por encima de un determinado umbral y así poder aplicar pruebas de bondad de ajuste que nos permitan advertir que tipo de distribución mejor se ajusta a las colas de las distribución y así darle un tratamiento adecuado a los valores extremos.

4.2 Análisis exploratorio de los datos de frecuencia y severidad

Tabla 4.1. Eventos de riesgo durante los últimos cinco años

Riesgos Operacionales	Frecuencia Absoluta (Eventos de riesgo)	Frecuencia relativa (%)
Fraude interno	2662	6.95
Fraude externo	1102	2.88
Relaciones laborales	3586	9.36
Clientes	7853	20.49
Daños a activos fijos	157	0.41
Fallas tecnológicas	7378	19.25
Administración y ejecución de procesos	15587	40.67
Total	38325	100.00

Figura 4.1. Distribución frecuencia de los diferentes eventos del riesgo operativo

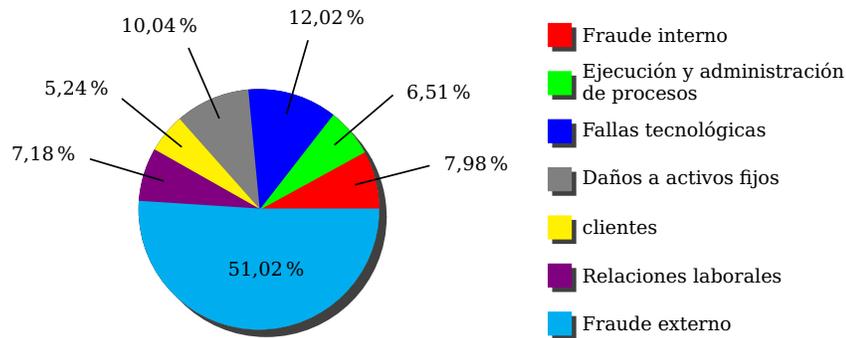


La Ejecución y Administración de procesos son los eventos de mayor ocurrencia con un 40.67 %, seguida por los eventos de Clientes (20.49 %), Fallas tecnológicas (19.25 %) estas tres categorías acumulan el 80.41 % lo cual sugiere que se les debe hacer el mayor control para lograr disminuirlos.

Tabla 4.2. Pérdidas económicas por eventos de riesgo en los últimos cinco años

Riesgos Operacionales	Pérdidas económicas (severidad)	Frecuencia relativa (%)
Fraude interno	182.062.136	7,98
Fraude externo	1.164.038.293	51,02
Relaciones laborales	163.858.258	7,18
Clientes	119.499.527	5,24
Daños a activos fijos	229.047.558	10,04
Fallas tecnológicas	274.332.722	12,02
Administración y ejecución de procesos	148.565.252	6,51
Total	2.281407746	100,00

Figura 4.2. Distribución del impacto económico de cada uno de los eventos de Riesgo operativo



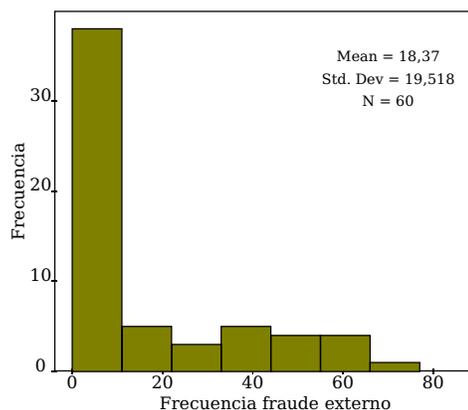
El Fraude externo representa el 51,02 % del total de las pérdidas económicas seguida por las pérdidas ocasionadas por Fallas Tecnológicas con un porcentaje de 12,02 % y Daños a activos fijos con un 10,04 % entre estos tres eventos se acumula un total del 73,08 % de las pérdidas.

Dado que el mayor porcentaje de pérdidas lo aporta Fraude Externo se analizará prioritariamente este evento se presentará un análisis exploratorio de la frecuencia y la severidad de esta variable.

Tabla 4.3. Tabla de frecuencia de la frecuencia del fraude externo

I	Lím inf	Lím sup	Marca de clase	Frec absoluta	Frec absoluta acumulada	Frec relativa	Frec relativa acumulada
1	0	11	5,5	38	38	0,633	0,633
2	11	22	16,5	5	43	0,083	0,717
3	22	33	27,5	3	46	0,050	0,767
4	33	44	38,5	5	51	0,083	0,850
5	44	55	49,5	4	55	0,067	0,917
6	55	66	60,5	4	59	0,067	0,983
7	66	77	71,5	1	60	0,017	1,000

Figura 4.3. Histograma de la frecuencia del fraude externo



El histograma presenta una asimetría positiva o hacia la derecha.

Tabla 4.4. Medidas Frecuencia de la Frecuencia del Fraude Externo

Medida	Valor de la medida
Media	18,37
Desviación Estándar (D.E.)	19,518
Coefficiente de Asimetría	0,527
Curtosis	1,348

Tabla 4.5. Frecuencia de la Severidad del Fraude Externo

I	Lím inf	Lím sup	Marca de clase	Frec absoluta	Frec absoluta acumulada	Frec relativa	Frec relativa acumulada
1	345149	5922345	3133747	12	12	0,200	0,200
2	5922345	11499541	8710943	11	23	0,183	0,383
3	11499541	17076737	14288139	7	30	0,117	0,500
4	17076737	22653933	19865335	3	33	0,050	0,550
5	22653933	28231129	25442531	8	41	0,133	0,683
6	28231129	33808325	31019727	7	48	0,117	0,800
7	33808325	39385521	36596923	12	60	0,200	1,000

Figura 4.4. Histograma de la frecuencia del fraude externo

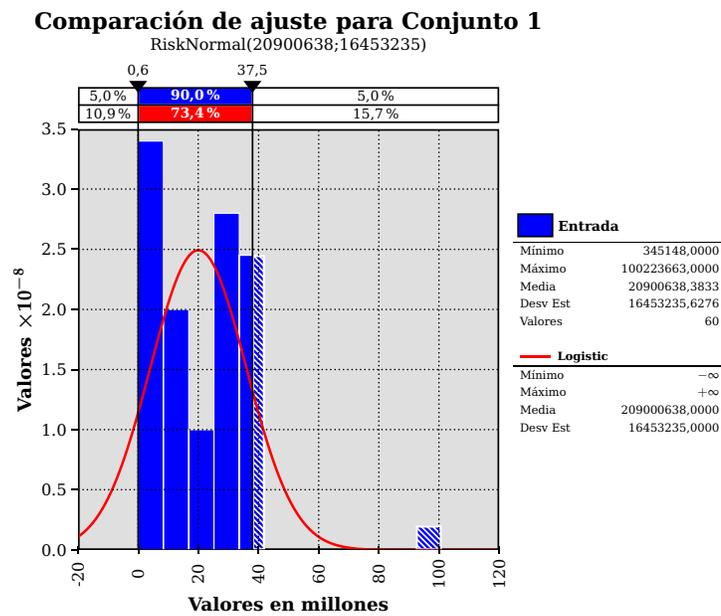


Tabla 4.6. Medidas de resumen de la Severidad del Fraude Externo

Medida	Valor de la medida
Media	19.400.638,22
Desviación Estándar (D.E.)	12.794.784
Coefficiente de Asimetría	1,538
Curtosis	-0,0027

Según el valor de la curtosis ($-0,0027$) se puede concluir que la distribución es mesocurtica es un valor muy cercano a cero, la distribución que tiene una curtosis igual a cero es la distribución normal, lo cual permite sospechar que la distribución a la cual se ajustan los datos de severidad es una distribución normal para confirmarlo se usará una prueba de bondad de ajuste.

4.3 Pruebas de bondad de ajuste para los datos

4.3.1 Pruebas de Bondad de ajuste para los datos de frecuencia

Tabla 4.7. Prueba de bondad de ajuste para la frecuencia

	Estadístico de prueba K-S alfa 0,05 p -value bilateral	
	Ajuste de los datos a la distribución de probabilidad de Poisson	Ajuste de los datos a la distribución de probabilidad binomial negativa
Fraude Externo	0,000	0,054

Para un nivel de significancia del 5 % no se rechaza la hipótesis nula de que la muestra proviene de una población que se distribuye como una binomial negativa y por el contrario se descarta que provenga de una distribución de Poisson.

4.3.2 Pruebas de bondad de ajuste para los datos de severidad

Tabla 4.8. Prueba de bondad de ajuste para la severidad

	Estadístico de prueba k-S alfa 0,05 p -value bilateral					
	Primera distribución		Segunda distribución		Tercera distribución	
	Distribución	p -valor	Distribución	p -valor	Distribución	p -valor
Fraude Externo	Normal	0,1191	Logística	0,1172	Weibull	0,1286

Para la severidad la distribución que mejor se ajusta aunque es la logística escogemos la normal ya que la diferencia entre los dos valores no es muy grande (ver Anexo A y B).

Una vez se conoce la distribución de la frecuencia y de la severidad se hace la convolución de estas dos distribuciones obteniéndose los valores de las pérdidas agregadas que servirán como una muestra para después obtener por simulación un número mayor de valores de estas pérdidas agregadas y así poder hacer un ajuste de esta variable aleatoria que se convierte en la más importante dentro del proceso y que se denominará la distribución de perdida la cual se representará por $S(x)$.

4.4 Distribución de pérdida agregada

A partir de la distribución binomial negativa de la frecuencia del Fraude Externo y de la distribución normal de la Severidad se construye la convolución de ambas para crear una nueva variable aleatoria las pérdidas económicas variable central del presente estudio. En las siguientes tablas y gráficas se muestra el proceso que se siguió.

Figura 4.5. Convolución de la distribución de Frecuencia y la Severidad

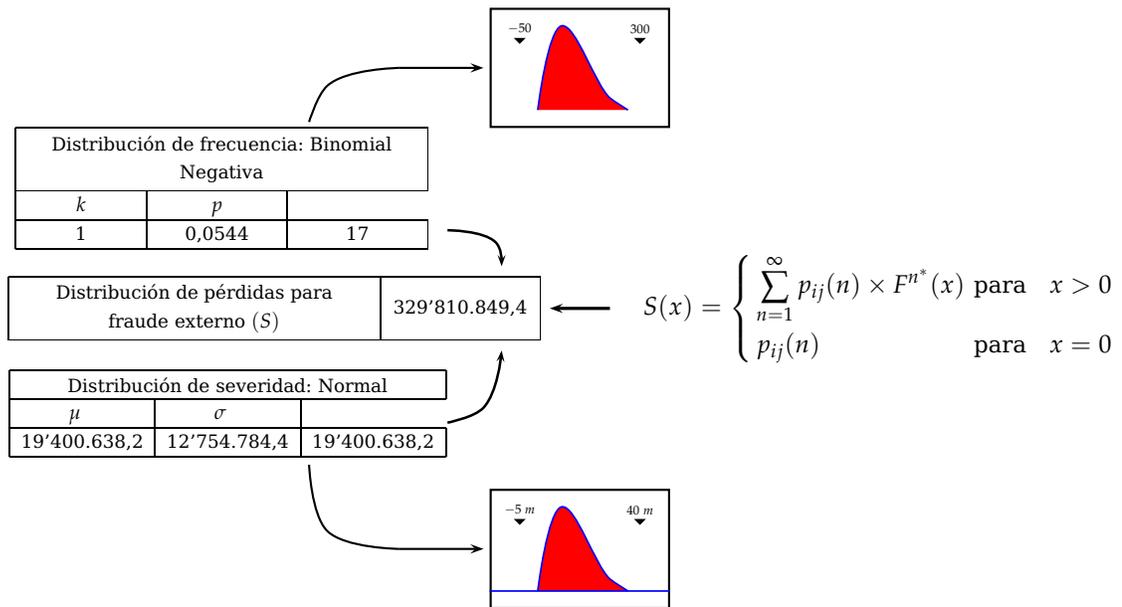
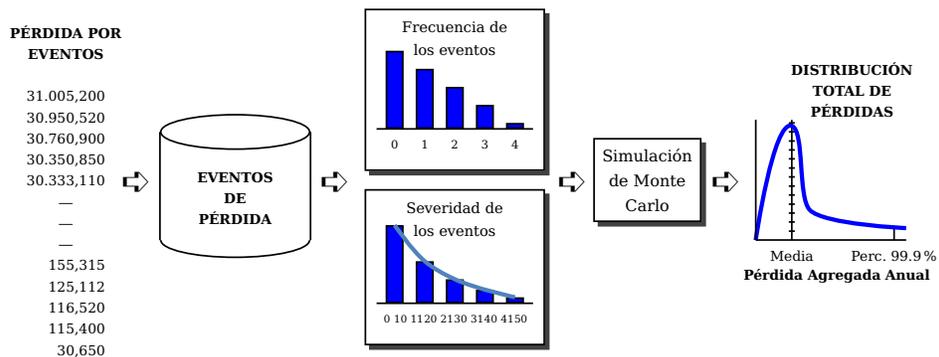


Figura 4.6. Modelos Agregados de Severidad y de Frecuencia



Los parámetros de las distribuciones de la Frecuencia y de la Severidad se reportan en la siguiente tabla:

Tabla 4.9. Parámetros de las distribuciones utilizadas para hallar la distribución de pérdida

Riesgo Opeacional	Distribución de probabilidad Binomial negativa Frecuencia		Distribución de probabilidad normal (severidad)	
	k	p	μ	σ
Fraude Externo	1	0,0544	19.400.638,2	12.794.784,4

A las pérdidas agregadas se le aplica la prueba de bondad de ajuste de Kolmogorov Smirnov para conocer el tipo de distribución de probabilidad a la cual se ajustan.

La distribución de probabilidad a la cual se ajusta la pérdida económica agregada es la Loglogística y como segundo lugar la Logística. Son tan cercanos los p -value que escogemos la Logística.

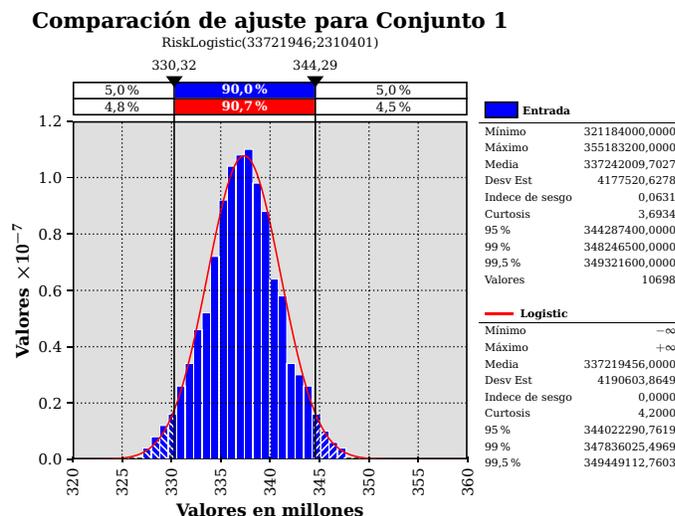
Tabla 4.10. Distribuciones de mejor ajuste de las pérdidas económicas causadas por el Fraude Externo

Riesgo operativo	Distribución de probabilidad Logística p -value	Distribución de probabilidad Loglogística p -value
Fraude Externo	0,292336	0,296346

Tabla 4.11. Parámetros de las Distribuciones que mejor ajustan a las Pérdidas Económicas

Riesgo Opeacional	Distribución de probabilidad Logistica		Distribución de probabilidad LogLogistica	
	μ	σ	ζ	Γ
Fraude Externo	337242009	4177520,62	140,735	337,202

Figura 4.7. Ajuste de las pérdidas económicas a la distribución Logísticas



4.4.1 Aplicando el Método de pérdidas agregadas (LDA)

Las pérdidas esperadas e inesperadas (ver tabla 4.12) para la distribución de pérdidas económicas que se ajustan a una Distribución Logística aplicando el método VaR para los percentiles más comúnmente usados 95 %, 99,5 % y 99,9 % corresponderán al VaR operacional suponiendo que lo importante es el cuerpo de la distribución y no el comportamiento de las colas.

Tabla 4.12. Pérdidas esperadas e inesperadas método LDA

Fraude Externo	Pérdidas esperadas (E(S))	OpVar	Pérdidas inesperadas (OpVaR-E(S))
Percentil 95 %	337.219.456	344.022.290,76	6.802.834,7
Percentil 99 %	337.219.456	347.836.025,50	10.616.569,5
Percentil 99,5 %	337.219.456	349.449.112,76	12.229.656,7

Si la entidad aprovisiona las pérdidas esperadas, Basilea acepta como carga de capital la diferencia entre el VaR operacional y las pérdidas esperadas E(S) lo correspondiente a la columna 4 que son las pérdidas inesperadas, en caso contrario, la entidad deberá tomar como carga de capital cualquiera de los percentiles según la norma que se establezca, es decir la pérdidas esperadas e inesperadas, $E(S)+U(S)$.

El valor en riesgo es una medida estadística de riesgo operativo que estima la pérdida máxima que podría registrarse en el Fraude externo en este caso en un intervalo de tiempo y con cierto nivel de probabilidad o confianza.

El VaR de un mes es de \$ 344.022.290,76 con un 95 % de confianza significa que la pérdida máxima esperada será de \$ 344.022.290,76 en 19 de cada 20 meses. En otras palabras, solo en un mes de cada 20 de operación de la empresa ($1/20 = 5\%$), en condiciones normales, la pérdida que ocurra puede ser mayor a \$ 344.022.290,76.

El VaR de un mes es de \$ 347.836.025,50 con un 99 % de confianza significa que la pérdida máxima esperada será de \$ 347.836.025,50 en 99 de cada 100 meses. En otras palabras, solo en un mes de cada 100 de operación de la empresa ($1/100 = 1\%$), en condiciones normales, la pérdida que ocurra puede ser mayor a \$ 347.836.025,50.

El VaR de un mes es de \$ 349.449.112,76 con un 99,5 % de confianza significa que la pérdida máxima esperada será de \$ 349.449.112,76 en 995 de cada 1000

meses. En otras palabras, solo en cinco meses de cada 1000 de operación de la empresa ($5/100 = 0,5\%$), en condiciones normales, la pérdida que ocurra puede ser mayor a \$ 349.449.112,76.

Aplicando el algoritmo con diferentes número de iteraciones (28 casos en total) para un nivel de confianza del 99,9% se obtuvieron los resultados de la tabla 14 la cual presenta el valor promedio de la pérdidas agregadas para 28 simulaciones cada una de diferente número de iteraciones y la carga de capital por riesgo operacional, destacamos que el valor con una probabilidad del 99,9% es de \$ 342.190.000 lo cual se interpreta de que existe una probabilidad del 99,9% de que la carga de capital no sea excedido por ese valor en un periodo de un mes.

Es decir la carga de capital converge a un valor cercano a \$ 342.190.000.

Tabla 4.13. Calculo de las Pérdidas Agregadas promedio y del VaR Operacional

Dato	iteraciones	Pérdidas promedio	OpVar 99,9 %	Dato	iteraciones	Pérdidas promedio	OpVar 99,9 %
1	1000	343060000	367900000	15	60000	337010000	346010000
2	2000	343820000	362820000	16	70000	338030000	345940000
3	3000	340030000	357570000	17	80000	337650000	345890000
4	4000	339630000	352310000	18	90000	337700000	345830000
5	5000	336000000	350270000	19	100000	337410000	345760000
6	6000	337790000	349130000	20	200000	337040000	344930000
7	7000	338920000	348840000	21	300000	336610000	344500000
8	8000	336260000	348500000	22	400000	337080000	344170000
9	9000	336490000	348110000	23	500000	337060000	343760000
10	10000	337090000	347840000	24	600000	337030000	343420000
11	20000	337590000	346420000	25	700000	336950000	343030000
12	30000	337080000	346240000	26	800000	336780000	342710000
13	40000	336580000	346140000	27	900000	336640000	342380000
14	50000	336800000	346070000	28	1000000	336690000	342190000

Figura 4.8. Número de iteraciones Vs. Promedio de pérdidas agregadas

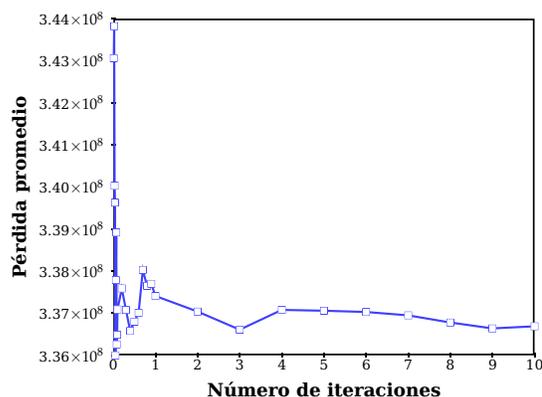
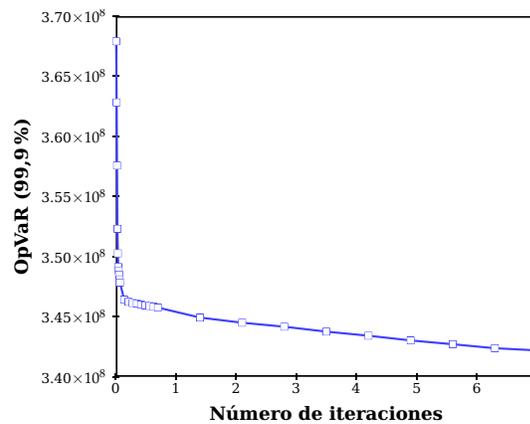


Figura 4.9. Número de iteraciones Vs. OpVaR

Con 1000000 de iteraciones los resultados indican, por ejemplo que OpVaR estimado con un nivel de confianza del 99,9% es \$ 342.190.000. La interpretación es la siguiente: Con un nivel de confianza del 99,9% se estima que las pérdidas operacionales por mes, ocasionadas por fraude externo en la Banca Comercial, no serán superiores a \$ 342.190.000

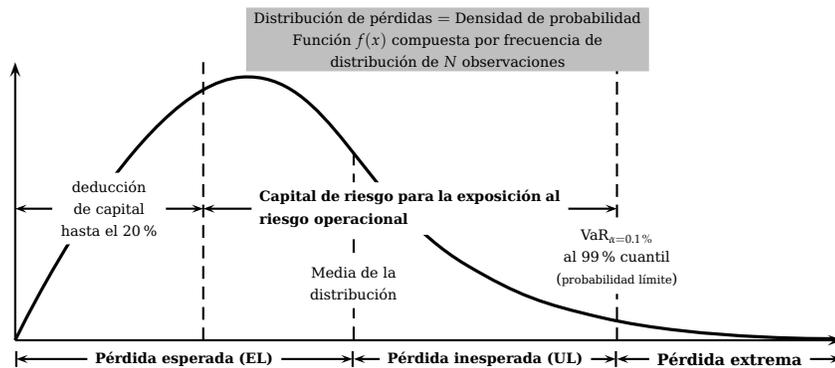
4.4.2 Aplicando la Teoría Básica del Valor Extremo

La Teoría básica del Valor Extremo consiste en un conjunto de técnicas estadísticas para la identificación y modelamiento de los máximos o mínimos de una variable aleatoria.

Uno de los objetivos centrales del presente trabajo al aplicar la Teoría Básica del Valor Extremo es advertir que para una institución financiera lograr la solvencia es decir, hacerle frente a los compromisos que la entidad asume es fundamental. El comportamiento inusual de la variable aleatoria de las pérdidas agregadas es lo más común en oposición al comportamiento normal que es más propio de la Teoría Clásica del Riesgo. Se debe tener en cuenta que se cuenta con una muestra pequeña lo cual limita la posibilidad de unos resultados de mayor alcance.

El otro objetivo es llegar a un ajuste que modele los valores de la muestra que excedan un determinado umbral o prioridad y ajustar la distribución generalizada de Pareto que es la más adecuada para modelar los excesos sobre un umbral cuando este es elevado.

El gráfico muestra la ubicación de las pérdidas extremas que se dan con baja frecuencia pero con alta severidad y allí el cuerpo de la distribución no aporta mucho.

Figura 4.10. Pérdidas esperadas, inesperadas y extremas

En la aplicación del método POT surgen tres cuestiones a resolver:

Lo primero es la selección del umbral óptimo por encima del cual la distribución de Pareto generalizada puede ser ajustada a los excesos sobre dicho valor. Determinar el umbral u . La segunda es determinar a qué tipo de distribución ajustan dichos excesos y la tercera es hallar la función de supervivencia, complementaria de la función de distribución.

En la primera etapa se generaron 10.698 valores de pérdidas agregadas haciendo uso de la simulación y se aplicaron varios métodos para hallar el parámetro de forma (ξ). Los procedimientos que se usaron para calcular el parámetro de forma fueron:

Primer procedimiento: Método de Hill

Con diferentes umbrales los cuales se tomaron de acuerdo a diferentes niveles de confiabilidad desde el 90% hasta el 99,5% con incrementos de 0,5 con los valores de los percentiles se determinó la pérdidas que corresponde a ese percentil y ese valor se toma como umbral y se busca que tipo de distribución se ajusta a los datos que están por encima de este umbral y se selecciona como distribución la correspondiente a la que mejor se ajuste de acuerdo al estadístico de prueba que surge después de aplicar la prueba de Smirnov Kolmogorov. Se selecciono como umbral el correspondiente al nivel de confiabilidad correspondiente a la confiabilidad cuya distribución presenta el mejor ajuste de todas en este caso se seleccionó la correspondiente al de una confiabilidad del 90% (ver Anexo C).

Una vez obtenido el umbral se estimó el parámetro de forma con base a los dos métodos siguientes:

Método I:

$$\hat{\zeta}_k = \left(\frac{1}{k-1} \sum_{i=1}^{k-1} \ln(x_i) \right) - \ln(x_k) \quad (4.1)$$

Método II:

$$\hat{\zeta}_k = \left(\frac{1}{k} \sum_{i=1}^k \ln(x_i) \right) - \ln(x_k) \quad (4.2)$$

Una vez calculado el valor de $\hat{\zeta}_k$ para los diferentes valores de k se obtiene un estimador promedio con las siguientes fórmulas:

Método I: $\hat{\zeta}_k = \frac{1}{n} \sum_{i=1}^n \theta_i$ donde,

$$\theta_k = \left(\frac{1}{k-1} \sum_{i=1}^{k-1} \ln(x_i) \right) - \ln(x_k) \quad \text{para } k = 2, \dots, n \quad (4.3)$$

Método II: $\hat{\zeta}_k = \frac{1}{n} \sum_{i=1}^n \theta_i$ donde,

$$\theta_k = \left(\frac{1}{k} \sum_{i=1}^k \ln(x_i) \right) - \ln(x_k) \quad \text{para } k = 1, 2, \dots, n \quad (4.4)$$

Tabla 4.14. Parámetros de los valores por encima del umbral seleccionado

Parámetro	Valor del Parámetro
Media	344.927.153
Desviación estándar (D.E.)	2.144.715,17
Índice de colas (Método I)	0,00613620
Índice de Colas (Método II)	0,00610887
Coficiente de Asimetría	1,291
Coficiente de Curtosis	1,520

El cálculo del parámetro de Hill por los dos métodos solo difiere en 0,00002741 lo cual corresponde a un porcentaje de menos del 1% (0,4466%) lo cual hace que el uso del uno o el otro parámetro para hacer cálculos sea indiferente. La disminución del parámetro de forma es también uno de los objetivos del control del riesgo operativo ya que este parámetro mientras más grande es indicativo de colas más pesadas y por lo tanto pérdidas económicas mayores por lo tanto el objetivo es disminuir este valor lo más posible.

Si una muestra tiene una gran curtosis tendera a tener un pico distintivo cerca a la media declina rápidamente y tendrá colas pesadas. Las distribuciones con

curtosis relativa positiva tienen colas pesadas y son llamadas leptocurticas. Los valores de la curtosis y la asimetría dan a entender que la distribución tiene colas pesadas, según la gráfica que aparece a continuación la distribución más acorde a estos parámetros es la distribución Beta General.

Figura 4.11. Ajuste de las pérdidas económicas en los valores extremos a la Beta

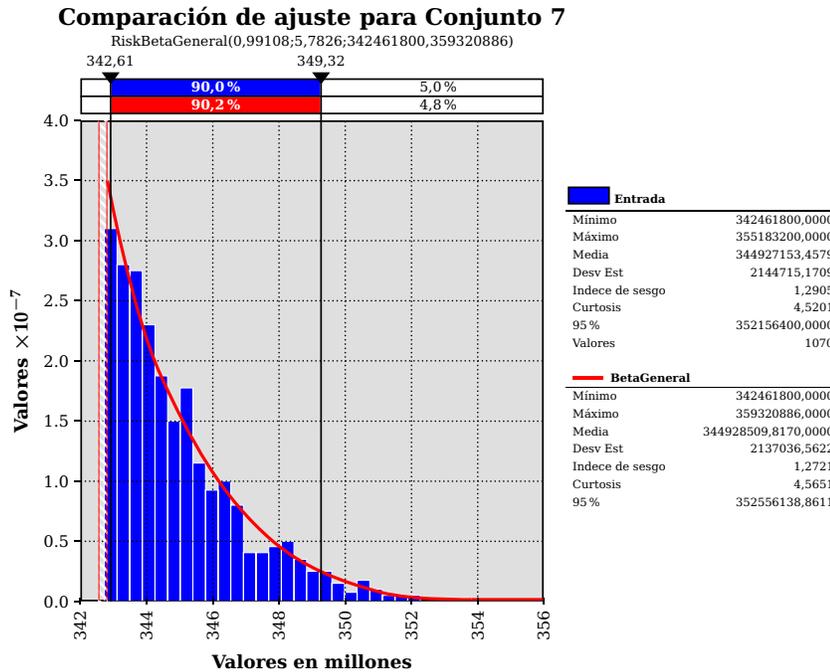


Tabla 4.15. Estimación de parámetros de la distribución de Pérdida de la cola derecha

	ξ	Distribución	Estadístico de K-S
Fraude Externo	0,00613629	Beta Generalizada	0,0142

Intervalo de Confianza para el Índice de cola

Tabla 4.16. Intervalos de Confianza para el parámetro de forma

	ξ		Límite inferior	Límite superior
Fraude Externo	Método I	Método II	0,0057894	0,0065274
	0,00613629	0,00610887	0,00576353	0,00649823

Se pueden usar el valor del parámetro de forma ξ obtenido por el método de Hill para calcular el valor del percentil (X_p) para diferentes valores de p que nos indicaría la probabilidad de que una pérdida tomará un valor inferior al valor $x = y + u$; su complementario $(1 - p)$ indica la probabilidad de que una pérdida

sea superior al valor $x = y + u$ condicionado a superar el umbral u . No se uso sino un solo valor $\xi = 0,00613629$ y no $\xi = 0,00610887$ porque deben dar muy parecidos. También se ha calculado el valor de Expected Shortfall, ES que con el VaR son las dos medidas más utilizadas para cuantificar el riesgo operativo asociado a las pérdidas debidas al evento de Fraude Externo. El valor del X_p estima el Var operacional.

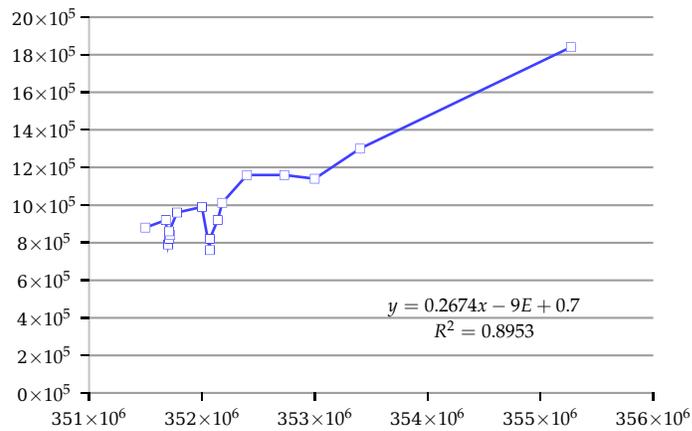
Tabla 4.17. Medidas OpVaR y Expected Shortfall, ES_α para cuantificar el riesgo

P	$1 - p$	$OpVaR_\alpha$	Beta General	ES_α
0,10	0,90	333.275.322	349.460.700	337.454.144
0,15	0,85	333.512.753	349.587.300	337.693.041
0,20	0,80	333.764.673	349.664.000	337.946.517
0,25	0,75	334.032.960	349.776.500	338.216.460
0,30	0,70	334.319.881	349.872.400	338.505.153
0,35	0,65	334.628.210	349.908.900	338.815.386
0,40	0,60	334.961.388	350.037.100	339.150.621
0,45	0,55	335.323.759	350.134.000	339.515.229
0,50	0,50	335.720.914	350.320.600	339.914.835
0,55	0,45	336.160.218	350.488.000	340.356.852
0,60	0,40	336.651.655	350.778.700	340.851.323
0,65	0,35	337.209.230	351.267.200	341.412.341
0,70	0,30	337.853.471	351.418.600	342.060.559
0,75	0,25	338.616.234	351.610.800	342.828.032
0,80	0,20	339.550.942	351.711.700	343.768.511
0,85	0,15	340.757.881	352.050.000	344.982.902
0,90	0,10	342.462.587	352.156.400	346.698.133
0,95	0,05	345.386.643	352.965.400	349.640.242
0,99	0,01	352.224.247	352.965.400	356.520.064

Un segundo procedimiento: Gráfica de exceso medio

Un primer aspecto a resolver es la selección del umbral óptimo a partir del cual la distribución de Pareto generalizada pueda ser ajustada a los excesos sobre dicho valor.

Para resolver este problema nos valemos de la siguiente gráfica: para representar esta gráfica la función de exceso medio empírica se tomo como variable dependiente y como variable independiente los valores del umbral o prioridad $u = X_{k+1}$.

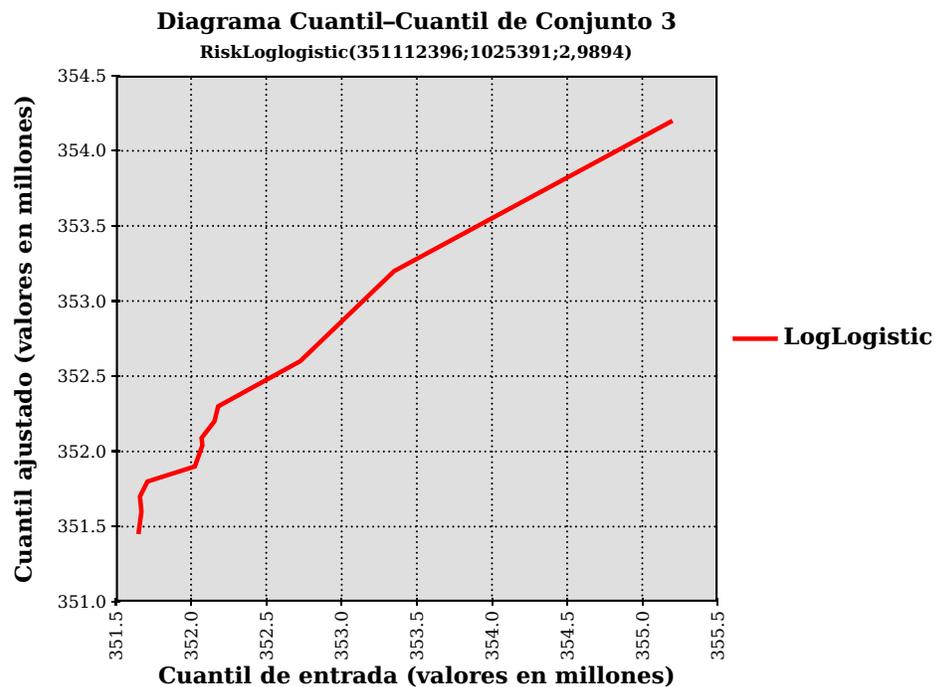
Figura 4.12. Función de Distribución Empírica versus Función de Distribución teórica

En la gráfica se observa que a partir de $351.610.800$ el gráfico del exceso medio empieza a ser creciente indicativo de cola gruesa a partir de dicho punto.

Por lo tanto se tomará como umbral este valor y a partir del cual se consigue una muestra de 14 observaciones.

La pendiente de la recta es igual a $0,2674$ y se sabe que teóricamente también es igual a $\frac{\xi}{1-\xi}$ por lo tanto igualando estos dos valores y despejando el valor de ξ se obtiene el valor de $\xi = 0,36500137313$.

Para contrastar esta afirmación y confirmar la existencia de cola gruesa en el problema que se está resolviendo se presenta un gráfico de cuantil-cuantil los cuantiles empíricos frente a los cuantiles teóricos de la distribución a la cual se ajustan los valores que están por encima del umbral.

Figura 4.13. Cuantil empírico versus cuantil teórico

La curva es lineal lo cual sugiere un buen ajuste a una Distribución Beta generalizada que cuando tiene dos parámetros es una la Distribuciones Generalizada Pareto.

Descripción de la muestra a modelar

Los estadísticos de la muestra que exceden el umbral seleccionado son:

Tabla 4.18. Valores muestrales que están por encima del umbral

Tamaño de la muestra : 14		Varianza: $9,08877 \times 10^{11}$
Media : 337.242.010		Desviación estándar: 953350,1626
Mínimo: 351.610.800		Asimetría : 2,15
Máximo: 355.183.200		Curtosis: 5,37
Percentil (Q)	Valor para el percentil (XQ)	Número de valores por encima de XQ(K)
Q = 25 %	$X_{0,25} = 351.788.025$	K = 4
Q = 50 %	$X_{0,50} = 352.088.300$	K = 7
Q = 75 %	$X_{0,75} = 352.635.675$	K = 10
Q = 90 %	$X_{0,90} = 353.229.860$	K = 12

El método de estimación de los parámetros fueron los estimadores de Drees-Pickands.

Los parámetros estimados para el modelo, optimizando el ajuste con un umbral de 351.610.800 y 14 excesos, son los siguientes:

El valor del parámetro de escala β se calcula con la siguiente ecuación:

$$\beta = \sigma + \xi(u - \mu) \quad (4.5)$$

$$\mu = 337.242.010 \quad ; \quad \sigma = 4.177.520,628 \quad ; \quad \xi = 0,36500137 \quad ; \quad u = 351.610.800$$

$$\beta = 4.177.520.628 + \xi(351.610.800 - 337.242.010)$$

$$\beta = 9.422.148,77$$

La función de Distribución Generalizada de Pareto se ha estimado para que ajuste los valores extremos que exceden el umbral o la prioridad u y no los excesos, ya que en caso de ajustarse los excesos sobre $u(X - u)$, el parámetro de localización, μ , sería 0 en lugar de 337.242.010.

Comprobación de la bondad del ajuste

A continuación se presenta un gráfico de cuantil-cuantil (Q-Q Plot), que representa los cuantiles empíricos (valores muestrales) frente a los cuantiles teóricos obtenidos a través de la distribución estimada:

Grafico cuantil-cuantil

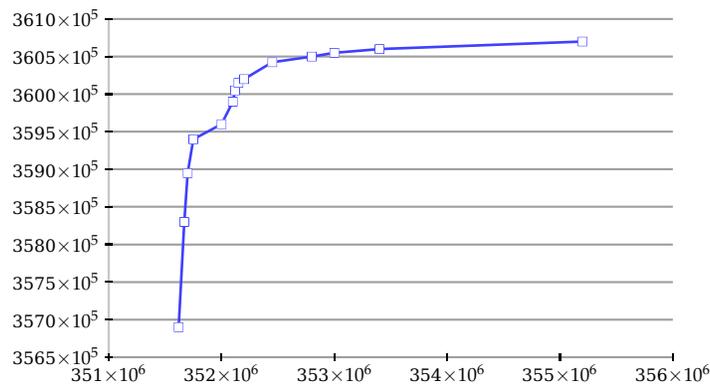
$$\begin{aligned} G_{\xi,\beta,\mu}(p) &= u + \frac{\beta}{\xi} \left\{ \left[\frac{n}{N_u}(1-p) \right]^{-\xi} - 1 \right\} \\ &= 351.610.800 + \frac{9.422.148,177}{0,36500137} \left\{ \left[\frac{10698}{14}(1-p) \right]^{-0,36500137} - 1 \right\} = x_p \\ x_p &= 351.610.800 + 25.814.009 \left\{ [764,142857(1-p)]^{-0,36500137} - 1 \right\} \end{aligned}$$

Con $p = \frac{r}{n+1}$

r : Posición del valor al cual se le va a calcular la función de distribución acumulada.

n : Tamaño de la muestra

p : Percentil

Figura 4.14. Función de Distribución Empírica versus Función de Distribución teórica

Inferencia y Predicción

Tabla 4.19. Algunos posibles resultados obtenidos a partir de la función de Distribución y función de supervivencia teórica

X	$G(x)$	$1 - G(x)$
360.000.000	0,9994	0,0006
355.000.000	0,9991	0,0009
352.000.000	0,9987	0,0013
352.500.000	0,9988	0,0012

Por ejemplo, habiendo superado una pérdidas económica 351610800, la probabilidad de que la pérdida sea inferior \$ 360.000.000 es de 99,94 % y por lo tanto la probabilidad de que sea superior a \$ 360.000.000 es de 0,06 %. Igual interpretación se pueden hacer con los demás valores.

Tabla 4.20. Pérdidas para varios percentiles y frecuencia de pérdida

$P(\%)$	Frecuencia de pérdida $\frac{1}{1-p}$	Pérdida (X_p) en miles de pesos
90,0	10	331.099.455
95,0	20	332.626.002
99,0	100	338.085.196
99,9	1000	354.273.966

Por ejemplo se espera con una probabilidad del 99,9 %, una vez que una pérdida supere el umbral de \$ 351.610.800 el valor de ella será inferior a \$ 354.273.966. Por el contrario, con una probabilidad del 0,1 % (esto es 1 de cada 1000 eventos superan el umbral) la pérdida será superior a \$ 354.273.966.

Estimación de las pérdidas por encima del percentil 99,9 % por los métodos Valor

en riesgo y EVT

Se puede calcular el valor esperado de las pérdidas en la cola con la siguiente ecuación:

$$ES_{\alpha} = \frac{\hat{x}_{\alpha}}{1 - \hat{\xi}} + \frac{\bar{\beta} - \hat{\xi}(u)}{1 - \hat{\xi}} \quad (4.6)$$

Tabla 4.21. Comparación de las pérdidas esperadas en la distribución de la cola

	ζ	LDA	POT	EL	LDA-EL	POT-EL
Fraude	0,00613629	352.556.139	355.189.911	344.928.510	7.627.629	10.261.401
Externo	0,36500137	357.136.480	341.622.818	352.354.042	4.782.438	10.731.224

El modelo LDA incluye las pérdidas esperadas (EL) y las pérdidas inesperadas (UL) para un nivel de confianza del 99,5 % para un valor del parámetro de forma $\zeta = 0,36500137$ las pérdidas para este modelo son mayores que para el modelo POT en \$ 15.513.662. Cuando $\zeta = 0,00613629$ sucede todo lo contrario las pérdidas según el modelo POT son mayores que las del modelo LDA y la diferencia es de \$ 2.633.772.

Tabla 4.22. Estimación de las pérdidas esperadas en la cola

	EL	ES
Fraude	344.928.510	359.504.038
Externo	352.354.042	350.719.721

CONCLUSIONES

En el caso de estudio el fraude externo es el evento que implica mayores pérdidas el 51,02 % del total de pérdidas en segundo lugar las Fallas Tecnológicas con el 12,02 % y en tercer lugar Daños a activos fijos 10,04 % para un gran total de estos tres eventos del 73,08 % del total de las pérdidas.

La alternativa de modelación del método de simulación de Montecarlo aplicado al método de Pérdidas Agregadas muestra que la carga de capital con un millón de iteraciones para el evento Fraude Externo para una probabilidad del 99.9 % es de \$ 342.190.000, es decir hay una probabilidad del 99.9 % de que la carga de capital no sea excedida por \$ 342.190.000. Y las pérdidas agregadas esperadas para un mes son de \$ 336.690.000. Valores a los cuales convergen estas dos variables.

Cuando se aplica el método Picos sobre un umbral se observa que mientras menor sea el parámetro de escala menor es la carga de capital que se debe reservar. Se presentan dos casos dependiendo del valor del parámetro de forma (ξ), si $\xi = 0,00613629$ el requerimiento de capital es de \$ 355.189.911 y si $\xi = 0,36500137$ es de \$ 341.622.818.

Los valores menores de carga de capital cuando se mide el riesgo a través del VaR se están logrando con el método Pérdidas Agregadas cuando el parámetro de forma es $\xi = 0,36500137$ y cuando el parámetro de forma es $\xi = 0,00613629$ el método más eficiente es el Picos sobre un umbral.

El valor del Expected Shortfall es \$ 350.504.038 cuando el parámetro de forma es $\xi = 0,00613629$ y de \$ 350.319.721 cuando $\xi = 0,36500137$

Una sugerencia que se hace a partir del presente estudio es que se calcule y estandarice el parámetro de forma para diferentes empresas según sus objetivos y tamaño ya que este parámetro es el más sensible para la obtención del capital

económico mínimo requerido que las instituciones financieras utilizan internamente para manejar su propio riesgo.

Se recomienda que se adquiriera la cultura de beneficiarse del cálculo del parámetro de forma para controlar el logro de la disminución del riesgo en las colas ya que haciéndolo cada vez más pequeño la empresa esta menos propensa a tener eventos de baja frecuencia y alta severidad y por lo tanto menos propensa a quiebras y grandes pérdidas de dinero cuando se usa el VaR.

Una recomendación que se hace para próximos trabajos es que se modele el VaR y el ES con metodologías que modelen la dependencia de la varianza condicional es decir, que se implementen modelos como RiskMetrics, ARMA-GARCH y ARMA-GARCH-EVT hasta este trabajo se supuso que las pérdidas económicas debidas al riesgo operativo son una variable aleatoria independiente e idénticamente distribuida acá se ha tenido en cuenta la Teoría del Valor Extremo sin modelar dependencia existente en el primer y segundo momento de la serie condicional de la serie. Y los modelos que se mencionan antes modelan los dos tipos de dependencia.

ANEXO A

FUNCIÓN DE DISTRIBUCIÓN ACUMULADA PARA LA
PÉRDIDA POR FRAUDE EXTERNO CON SEVERIDAD
NORMAL

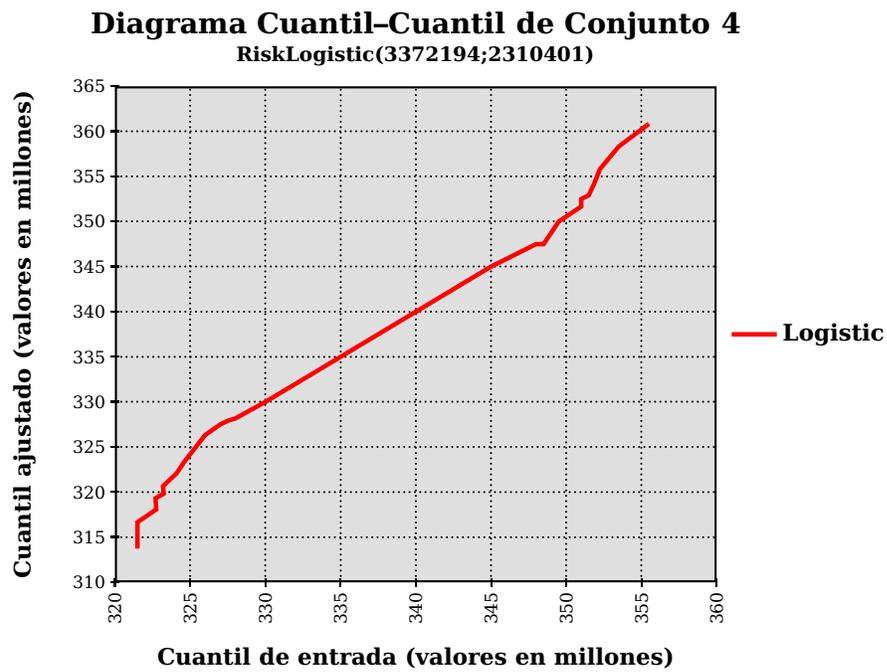
Figura A.1. Función de distribución Acumulada para las pérdidas de Fraude Externo



ANEXO B

GRÁFICO Q-Q PARA LA PÉRDIDA PARA FRAUDE EXTERNO CON SEVERIDAD NORMAL

Figura B.1. Gráfico Q-Q para la pérdida para fraude externo con severidad normal



ANEXO C

VaR PARA DIFERENTES PERCENTILES Y VALOR DEL ESTADÍSTICO DE PRUEBA SMIRNOV-KOLMOGOROV

Tabla C.1

Percentil	Estadístico Smirnov-Kolmogorov	VaR	Distribución a la cual mejor se ajusta
90,0	0,0142	349.321.600	BETA GENERAL
90,5	0,0163	349.438.300	BETA GENERAL
91,0	0,0159	349.460.700	BETA GENERAL
91,5	0,0152	349.587.300	BETA GENERAL
92,0	0,0166	349.640.597	BETA GENERAL
92,5	0,0170	349.761.219	BETA GENERAL
93,0	0,0180	349.876.824	BETA GENERAL
93,5	0,0209	350.015.699	BETA GENERAL
94,0	0,0214	350.133.081	BETA GENERAL
94,5	0,0326	350.323.047	BETA GENERAL
95,0	0,0285	350.372.647	BETA GENERAL
95,5	0,0277	350.574.540	BETA GENERAL
96,0	0,0286	350.832.689	BETA GENERAL
96,5	0,0311	351.058.164	BETA GENERAL
97,0	0,0367	351.342.423	BETA GENERAL
97,5	0,0443	351.528.079	BETA GENERAL
98,0	0,0622	351.864.389	BETA GENERAL
98,5	0,0471	352.213.400	LOGNORMAL
99,0	0,0396	352.535.861	BETA GENERAL
99,5	0,0758	353.341.678	BETA GENERAL

ANEXO D

CÓDIGO DE MATLAB SIMULACIÓN DE MONTECARLO (BINOMIAL NEGATIVA-NORMAL)

```
% LECTURA DE LOS DATOS DE ENTRADA
Clear all;
Format short;
Clc;
% PARAMETROS DE LA BINOMIAL NEGATIVA
R=1
P=0.0544
% PARAMETROS DE LA NORMAL
X=19400638.2;
S=12794784.4
% NUMERO DE DATOS A PROCESAR
N=1000000
sumt=0;    k=1;
for i=1:N
    if rem(i,1000)==0
        i
    end
    NBNeg=nbinnrnd(R,P);
    sump=0;
    for j=1:NBNeg
        perdidai = normrnd(X,S);
        sump=sump+perdidai;
    end
    sumt=sumt+sump;
    PROM(1,i)=sumt/i;
    if ((i <= 10000) & (rem(i,1000) == 0))
        DD1(1,k)=i;
        DD4(1,k)=PROM(1,i);
    end
end
```

```
    DD5(1,k)=prctile(PROM,99.9);
    Datos=[DD1' DD4' DD5'];
    save Datos;
    save PROM;
    k=k+1
else
    if ((i > 10000) & (i <= 100000) & (rem(i,10000) == 0))
        DD1(1,k)=i;
        DD4(1,k)=PROM(1,i);
        DD5(1,k)=prctile(PROM,99.9);
        Datos=[DD1' DD4' DD5'];
        save Datos;
        save PROM;
        k=k+1
    else
        if ((i > 100000) & (i <= 1000000) & (rem(i,100000) == 0))
            DD1(1,k)=i;
            DD4(1,k)=PROM(1,i);
            DD5(1,k)=prctile(PROM,99.9);
            Datos=[DD1' DD4' DD5'];
            save Dato;
            save PROM;
            k=k+1
        end
    end
end
end
end
figure
plot(DD1,DD4,'-xb')
set(gca,'FontSize',12);
xlabel('Numero de Iteraciones','FontSize',12);
ylabel('Perdida Promedio','FontSize',12);
X(1,1)=DD1(1,1);
X(1,2)=DD1(1,k-1);
Y(1,1)=DD5(1,k-1);
Y(1,2)=DD5(1,k-1);
figure
plot(DD1,DD5,'-xb')
set(gca,'FontSize',12);
xlabel('Numero de Iteraciones','FontSize',12);
```

```
ylabel('OpVaR(99.9%)', 'FontSize', 12);  
format bank;  
[DD1; DD4; DD5]'
```


BIBLIOGRAFÍA

- AKKIZIDIS, I. y BOUCHEREAU, V. (2006), *Guide to Optimal Operational Risk & Basel II*, Boca Raton , FL , Auerbach Publications.
- BALKEMA, L. de haan, (1996), *Residuan lifetime at great age*, *Ann Probad* 2. (1974), p. 792–804.
- BEIRLANT, J. TEUGELS, J. & VINCKIER, P., *Practical analysis of extreme*.
- BERLANT J. y Otros (2005), *Statistics. Theory and Applications*, Editorial John Wiley & Sons, Ltd. Inglaterra, pp. 490.
- BOCKER, K. (2006), *Closed-form approximation form OpVaR when high frequency losses are parameterized by a generalized Pareto distribution (GPD)*.
- BOCKER K., HYPOVEREINSBANK M. (2006), *A closed-form approximation for OpVaR when high frequency losses are parameterized by a generalized Pareto distribution (gpd)*. January 2006.
- BUHLMANN, H. (1970), *Mathematical Methods in risk theory*. Springer-Verlag.
- CARRILLO, S. (2006), *Riesgo Operacional: Medición y Control*. Jornadas Técnicas de Basilea II, UNIA, Sevilla, Septiembre.
- CHERNOBAI SVENTLOZAR, A., RACHEV, T. (2006), *Applying Robust Methods to Operational Risk Modelind*.
- CHRISTOPHER M. L., (2001), *Measuring and Managing Operational Risk in Financial Institutions*, Singapore, John Wiley & Sons.
- COMITE DE SUPERVISION BANCARIA BASILEA, *Operational Risk*, Basilea, Mayo de 2001

- COMITÉ DE SUPERVISION BANCARIA DE BASILEA. *Supervisory Guidance on Operational Risk Advanced Measurement Approaches for Regulatory Capital*, Basilea, Julio de 2003.
- CONNELL, P., DAVIES, M., (2006), *Safety First–Scenario Analysis under Basel II*. April p. 18.
- CONVERGENCIA INTERNACIONAL DE MEDIDAS Y NORMAS DE CAPITAL, Basilea 2006
- CROUHY, M. y Otros. (2001) *Risk Management*. Ed. McGraw Hill. pp. 709
- CRUZ, M. (2002), *Modeling, Measuring and Hedging Operational Risk*, John Wiley & Sons LTD, Reprinted August 2005.
- DA COSTA , N. (2004), *Operational Risk with excel an VBA*, John Wiley & Sons, New Jersey, p. 203–207.
- DE FONTNOUELLE P., JORDAN J., & ROSENGREN E., (2005), *Implication of Alternative Operational Risk Modeling Techniques*, February 2005.
- DE HAAN L. y FERREIRA A. (2006), *Extreme Value Theory, An Introduction*, Editorial Springer, Estados Unidos de América, pp. 417.
- DE LARA HARO A. (2004), *Medición y Control de los Riesgos Financieros*, 3a ed. Editorial Limusa, Grupo Noriega Editores, México, pp. 219.
- DUTTA, K. And PERRY, J. (2006), *A Tale of Tails: An Empirical Analysis of Loss Distribution Models for Estimating Operational Risk Capital*, FRB of Boston Working Paper No. 06–13.
- EMBRECHTS P., KLUPPELBERG C. & MIKOSCH T. (1997), *Modeling Extremal Events for insurance and Finance*, Springer-verlang.
- FERIA J. y JIMENEZ E. J. (2005), *El OpVAR como medida del riesgo Operacional*, Documento de trabajo, 2005.
- FISHER, R. A. & TIPPET, L. H. C. (1928), *Proc*, Cambridge Phil, Soc.
- FOUNTNOUELLE, P., JORDAN, J., ROSENGREN, E., (2005). *Implicacion of Alternative Operational Risk Modeling techniques*, pág. 46.
- FRACHOT, A., RONCALLI T., SALOMON, E., (2004), *The Correlation Problem in Operational Risk*.

- FRANCO L. C. (2010), *Análisis y comparación de alternativas para cuantificar el riesgo operacional*, Universidad Eafit, pp. 69.
- FRANCO, L. C., MURILLO J. G. et al. (2006), *Riesgo Operacional: reto actual de las entidades financieras*, Revistas Ingenierías Universidad de Medellín.
- GARCIA, A., (2006), *Una aportación al análisis de solvencia: La teoría del valor extremo*, Tesis Doctoral Alcalá de Henares, pág. 43-238.
- GONZALEZ M. (2004), *Análisis del Nuevo Acuerdo de Capitales de Basilea (BIS-II): Pyme-risk, Country-risk y Operational-risk*, pág. 68.
- HOSKING J. R. M. (1985), *Algorithm AS 215: Maximun-Likelihood Estimation of the Parameters of the Generalized Extreme-Value Distribution*, Royal Statistical Society, Vol. 34, No. 3, pág. 301-310.
- HOSKING, J. R. M. y WALLIS J.R. (1987), *Parameter and Quantile Estimation for the Generalized Pareto Distribution*, American Society for Quality, Vol. 29, No. 3, pág. 339-349.
- HOSKING, J. R. M., WALLIS J. R., & WOOD E.F., *Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighed Moments*, American Society for Quality, Vo. 27, No. 3, pág. 251-261.
- JIMENEZ M. A., *El tratamiento del riesgo operacional en Basilea II*, Dirección General de Supervisión del Banco de España.
- JOBST A. A. (2007), *Operational Risk-the sting is still in the tail but the Poisson Depends on the Dose*. Journal of Operatioanal Risk Vol. 2, No. 2.
- JOHNSON, N. L. y Otros. *Continuos Univariate Distributions*, Vol 1. 2ª edición. Editorial John Wiley & Sons, Inc. New York. pp. 756.
- JOHNSON, Normal L y Otros. *Continuos Univariate Distributions*, Vol 2. 2ª edición. Editorial John Wiley & Sons, Inc. New York. pp. 756.
- JORIAN, P. (2002), *Valor en Riesgo. Nuevo paradigma para control de riesgos con derivados*, Editorial Limusa, México 2002, pp. 357.
- LEIPPOLD M., VANINI P. (2003), *The Quantification of Operational Risk*, pág. 38.

- McNEIL A., FERY R. (2005), *Quantitative Risk Management*, New Jersey, Princeton University Press, pág. 264–322.
- McNEIL, A. J. y Otros. (2005), *Quantitative Risk Management, Concepts, Techniques and Tools*, Ed. Princeton University Press. E.E.U.U. pp. 538.
- MELO VELANDIA L. F. y BECERRA CAMARGO O. R. *Medidas de riesgo, características y técnicas de medición: una aplicación del VaR y el ES a la tasa interbancaria de Colombia*.
- MIGNOLA G., UGOCCIONI R. (2002), *Test of Extreme Value Theory applied to Operational Risk data*, Operational Risk Management.
- MORA A. *Un estudio comparativo de algunos estimadores del Índice de valor extremo*, Documento borrador para propósitos de discusión, Colegio de Estudios Superiores de Administración.
- MORENO DIAZ, S. I. y VASCO FERNANDEZ C. P. (2007), *Diseño de una Metodología de Gestión del Riesgo Operacional Aplicado al Fondo Mutuo de Inversiones ABC*. Universidad EAFIT, Medellín, pp. 133.
- MOSCADELLI M. (2004), *The modeling of operational risk: experience with the analysis of the data collected by the Basel Committee*, pág. 26–53.
- MURILLO GÓMEZ J. G. (2009), *Cuantificación de las Pérdidas Económicas por riesgo operacional asociados a eventos extremos*, Tesis de Grado.
- PICKANDS J. (1975), *Statistical Inference Using Extreme Order Statistics*, *Annals of Statistics*, 3, pág. 119–131.
- REISS R. D. y THOMAS M. (2001), *Statistical Analysis of Extreme Values with Applications to insurance, Finance, Hydrology and other Fields*, Editorial Birkhauser Verlag, 2ª Edición, pp. 443.
- REISS R. D., & THOMAS J. A. (2001), *Statistical Analysis of Extreme Values*. Basel, Birkhauser.
- RESNIK S. I. (1987), *Extreme Values, Regular Variation, and Point Processes*, Editorial Springer-Verlag, Estados Unidos de América, pp. 320.
- RESNIK S. y STARICA C. (1995), *Consistency of Hill's Estimator for Dependent Data*. *Journal of Applied probability*, Vol. 32, No. 1, 139–167.

- SHEVCHENKO, P. (2005), *Operational Risk Modelling and Quantification, Invited talk for the fourth International Symposium on Business and Industrial Statistics (ISBIS4)*, Palm Cove, Australia.
- SHEVCHENKO, P. V. And DONNELLY J. (2005), *Validation of the Operational Risk LDA Model for Capital Allocation and AMA Accreditation under Basel II*, CMIS Confidential report prepared for Basel II programme ANZ bank, CMIS report number 05, pp. 132.
- SMITH RICHARD L., *Maximum Likelihood Estimation in a Class of Nonregular Cases*. Biometric Trust, Vol. 72, No. 1, pág. 67-90.
- SOLER RAMOS, J. (1999), *Gestión de Riesgos Financieros. Un enfoque práctico para países Latinoamericanos*, Ed. Banco Interamericano de Desarrollo, Washington, pp. 443.
- THE FOREIGN EXCHANGE COMMITTEE (2003), *management of Risk Operational in foreign Exchange*, March 2003.
- TRIPP M. H., BREDLEY H. I., DEVITT R., OVERTON G. L., PRYOR L. M., & SHAW R. A. (2004), *Quantifying Operational Risk in General Insurance Companies*, pp. 96.
- VELANDIA, O. (2007), *Comportamiento asintótico del VaR como media de valor extreme*, pág. 2-21.
- VENEGAS MARTÍNEZ F. (2008), *Riesgos Financieros y Económicos Productos Derivados y decisiones económicas bajo incertidumbre*, Edición 2ª, Editorial Cenegage Learning, México, pp. 1139.

