

MEDICIÓN EFECTIVA DEL RIESGO OPERACIONAL

Santiago Carrillo Menéndez (*)

Alberto Suárez (*)

(*) Santiago Carrillo Menéndez pertenece al Departamento de Matemáticas de la Facultad de Ciencias de la Universidad Autónoma de Madrid. Por su parte, Alberto Suárez trabaja en el Departamento de Informática de la Escuela Politécnica Superior de la Universidad Autónoma de Madrid.

Medición efectiva del riesgo operacional

En este trabajo se estudian algunos de los problemas planteados hoy a los gestores de riesgo operacional en el diseño e implementación de modelos avanzados. Utilizando datos sintéticos se analizan los impactos del riesgo de modelo, las dificultades en la aplicación de la teoría de valores extremos para el ajuste de la distribución de severidad y, finalmente, la validez de aproximaciones numéricas al uso. Las simulaciones realizadas muestran que, con bases de datos de pérdidas operacionales de tamaño reducido, que es la situación en la que numerosas entidades se encuentran actualmente, es difícil distinguir entre modelos alternativos que proporcionan estimaciones del capital regulatorio muy distintas. Una aplicación directa de la teoría de valores extremos basada en el ajuste de distribuciones de Pareto a sucesos de pérdidas por encima de un umbral elevado tiende a producir estimaciones que son muy inestables y excesivamente elevadas, por lo que existen muchas dudas acerca de su relevancia económica. Finalmente, las aproximaciones numéricas propuestas en la literatura podrían no ser adecuadas en los rangos de valores de las pérdidas operacionales reales. Estas observaciones ponen de manifiesto la necesidad de adoptar una actitud cauta al elegir los modelos para la cuantificación del riesgo operacional y la importancia de contrastar los modelos alternativos con datos empíricos.

1 Introducción

El riesgo operacional de una entidad financiera tiene su origen en sucesos que no pueden ser adscritos a riesgo de mercado o de crédito. Son fuentes de riesgo operacional las pérdidas debidas a la inadecuación o a fallos de los procesos, el personal o los sistemas internos y las causadas por acontecimientos externos a la entidad. La importancia del riesgo operacional ha sido reconocida recientemente en el mundo financiero. A consecuencia de su relevancia, el Comité de Supervisión Bancaria de Basilea ha incluido en el documento «Convergencia internacional de medidas y normas de capital-Marco revisado», conocido como Basilea II, un apartado específico sobre riesgo operacional. En Basilea II se describen diferentes metodologías para el cálculo de una cifra de capital regulatorio, requerido para hacer frente a posibles pérdidas por este tipo de riesgo.

Las metodologías propuestas tienen un grado de complejidad creciente. En el enfoque básico (*Basic Indicator Approach*), el capital regulatorio es calculado simplemente como un porcentaje fijo de un indicador de exposición al riesgo operacional (los ingresos brutos) para la entidad en su conjunto. En el enfoque estándar (*Standardised Approach*) se aplican distintos porcentajes, fijados por el regulador, a este indicador desagregado para cada una de las líneas de negocio del banco. Finalmente, para el enfoque avanzado (*Advanced Measurement Approaches*, AMA), Basilea II establece unas directrices generales, pero permite e incentiva que los propios bancos construyan su propio modelo de medición y gestión del riesgo operacional. Se espera que la metodología desarrollada dentro del enfoque avanzado refleje de manera más detallada el perfil de riesgo específico de cada entidad. Esto permitirá una cuantificación más precisa del capital regulatorio, el cual debería ser notablemente inferior al calculado por los enfoques básico y estándar, y una mejor gestión del riesgo. La implementación de estos modelos avanzados para la medición del riesgo operacional es uno de los retos más importantes de la puesta en marcha de Basilea II, tanto por parte de las instituciones financieras como para el regulador.

La metodología basada en modelizar la distribución (agregada) de pérdidas (*Loss Distribution Approach* o LDA) es uno de los paradigmas más usados en el enfoque avanzado, después de haber sido usada con éxito en el entorno de las matemáticas actuariales desde hace dos

décadas. En esta metodología el capital regulatorio se calcula como el percentil de la distribución agregada de pérdidas, para un nivel de probabilidad del 99,9% y un horizonte temporal de un año. Dado que el método de cálculo es muy similar al del valor en riesgo (VaR), esta cifra se suele denominar VaR operacional o bien capital en riesgo (*Capital at Risk*, CaR) operacional. Otra medida útil para cuantificar la exposición al riesgo es el capital en riesgo condicional (*Conditional CaR*, CCaR). El CaR condicional se calcula como el valor medio de las pérdidas condicionado a que dichas pérdidas se encuentren por encima del CaR. Entre otras ventajas [por ejemplo, es una medida coherente de riesgo en el sentido de Artzner et al. (1999)], proporciona una estimación del valor medio de pérdidas extremas e infrecuentes.

Para realizar los ajustes a distribuciones se utilizan datos históricos de pérdidas operacionales recientes del propio banco (datos internos, recogidos durante un intervalo de cinco años, aunque inicialmente, y de manera transitoria, se permite que el período comprenda al menos tres años) complementados por datos externos, con el fin de reflejar adecuadamente el impacto de pérdidas extremas, poco frecuentes. También se debe integrar en el cálculo de capital información sobre factores de negocio y de control interno, así como un análisis de escenarios (por ejemplo, mediante el uso de simulaciones) y la incorporación de información cualitativa.

En la modelización de las pérdidas agregadas, se parte del supuesto de que las severidades son independientes entre sí, y de que dichas severidades son independientes de la frecuencia de los sucesos. Estas hipótesis permiten modelizar por separado la distribución de severidad y la distribución de frecuencias. Una elección posible consiste en utilizar una Poisson para modelizar la distribución de frecuencias, y una lognormal para la distribución de severidad.

La inexistencia de bases de datos de riesgo operacional con una profundidad suficiente, el reto que supone la recogida de los datos de pérdidas, así como cierta evidencia empírica de que la magnitud del capital regulatorio podría estar determinada únicamente por pérdidas extremas y muy poco frecuentes, han motivado la propuesta de una metodología alternativa. Esta metodología consiste en tomar en cuenta únicamente los datos por encima de un umbral elevado (por ejemplo, 10.000 euros) para el cálculo de la contribución por riesgo operacional al capital en riesgo. El fundamento teórico de este planteamiento alternativo, que consisten en modelizar los excesos por encima de un determinado nivel (Peaks Over Threshold), es la teoría de valores extremos: para un umbral suficientemente elevado y suponiendo que se ha alcanzado un régimen asintótico, las dos distribuciones que describen el proceso (severidad y frecuencia) son conocidas. La primera se aproxima a una distribución de Pareto generalizada y la segunda a una distribución de Poisson.

Sobre la adecuación de uno u otro modelo a los datos de pérdidas reales se ha generado un vivo debate, aún no zanjado. La cuestión es importante, ya que las estimaciones del capital regulatorio proporcionadas por las dos aproximaciones son a menudo muy dispares. El enfoque basado en la modelización de la distribución de la severidad mediante una distribución que no asigne tanto peso a las colas (lognormal, gaussiana inversa u otras) proporciona estimaciones robustas y estables del capital regulatorio. El enfoque basado en la teoría de valores extremos goza de cierta aceptación en entidades que han optado por modelos avanzados en riesgo operacional, es sencillo de implementar y permite que las instituciones financieras necesiten recabar información únicamente sobre las pérdidas elevadas. Sin embargo, las cifras de capital estimadas son inestables, excesivamente sensibles a la presencia o ausencia de eventos extremos aislados y, a menudo, muy elevadas. Todo esto dificulta su implementación y su interpretación como capital regulatorio.

En este artículo se hace un análisis cuantitativo de las posibles dificultades en la modelización de las pérdidas por riesgo operacional y de las consecuencias en las medidas de riesgo de la elección de modelos incorrectos o incompletos. Utilizando datos de pérdidas simulados, mostraremos el comportamiento de los estimadores de las distribuciones lognormal y Pareto cuando el tamaño de las muestras a partir de las cuales se realizan los ajustes varía, y el impacto de estas fluctuaciones de muestreo en los valores estimados para el capital regulatorio. En concreto se investigan los efectos de elegir umbrales de recogida excesivamente elevados, la posible confusión entre muestras empíricas de lognormal y Pareto, se analiza de manera crítica el uso de la teoría de valores extremos en el contexto de riesgo operacional y se muestra la gran inestabilidad de las estimaciones de medidas de riesgo en el caso de que la cola de la distribución de severidad sea de tipo Pareto con valores elevados del parámetro de forma. Finalmente, se estudia la validez de dos aproximaciones analíticas para el cálculo del capital económico.

El artículo tiene la siguiente estructura: En la sección 2 se describe, en el contexto de las directrices de Basilea II, el enfoque avanzado para la estimación del capital regulatorio por riesgo operacional basado en modelizar la distribución agregada de pérdidas. La sección 3 aborda el problema de las colas pesadas en las distribuciones de severidad de pérdidas y su posible modelización mediante la teoría de valores extremos. Las dificultades para una correcta modelización de dichas colas cuando las muestras a partir de las cuales se estiman los parámetros de los modelos son de tamaño reducido son el objeto de la sección 4. En la sección 5 se analiza la inestabilidad de los ajustes a distribuciones de Pareto, su dependencia con la elección del umbral de ajuste, el efecto de la incorporación de nuevos datos de pérdidas, y su sensibilidad de los ajustes a sucesos extremos. La sección 6 investiga la aplicabilidad de diversas aproximaciones analíticas para el cálculo de capital regulatorio. Finalmente, en la sección 7 se resumen las principales conclusiones de los estudios realizados. Material técnico sobre las distribuciones de Pareto, empíricas y a trozos Pareto/empírica ha sido recogido en el apéndice.

2 El enfoque distribución de pérdidas y Basilea II

Para la medición del riesgo operacional, el Comité de Basilea ha definido ocho líneas de negocio (administración de activos, banca comercial, banca minorista, intermediación minorista, finanzas corporativas, negociación y ventas, pagos y liquidación y servicios de agencia) y siete categorías de sucesos de riesgo (fraude interno; fraude externo; relaciones laborales y fallos de seguridad en el puesto de trabajo; clientes, productos y prácticas empresariales; daños a activos materiales, incidencias en el negocio y fallos en los sistemas y ejecución, entrega y gestión de procesos).

Cada una de las pérdidas por riesgo operacional debe asignarse a una de las celdas de esta matriz de riesgos y el capital en riesgo debe calcularse en este marco. Una primera aproximación a la cuantificación de la contribución del riesgo operacional al capital en riesgo total consiste en calcular el CaR para cada una de las 56 celdas de esta matriz y sumar dichos capitales. Este método parte de un supuesto pesimista, en el que los riesgos en distintas celdas están totalmente correlacionados, y no contempla la posibilidad de la reducción del capital gracias a efectos de diversificación. En el marco de Basilea II, también es posible tener en cuenta la estructura de dependencia entre estas pérdidas en distintas celdas para reflejar el impacto de dicha diversificación de riesgos.

La metodología LDA ha sido generalmente considerada como la más idónea dentro de los enfoques avanzados. Esta metodología requiere la modelización por separado de las dos variables aleatorias que describen las pérdidas: la frecuencia N y la severidad X . La distribución de pérdida agregada para una celda determinada tiene la forma:

$$S = \sum_{i=1}^N X_i$$

Es decir, la pérdida agregada es suma de un número aleatorio de valores de pérdidas también aleatorios. Supondremos que las severidades X_i son independientes entre sí y también independientes de N .

El objetivo es estimar las distribuciones que mejor describen estas dos variables aleatorias para cada una de las 56 celdas de la matriz de riesgos. Entre las distribuciones candidatas a modelizar la distribución de frecuencias, se encuentran la distribución de Poisson, la distribución binomial negativa y las compuestas por estas, es decir la distribución de sumas aleatorias de tales variables. Para la modelización de la severidad, existe una serie de familias paramétricas que permiten aproximarse a la tarea: lognormal, gamma, Pareto, valores extremos generalizados, Weibull, Burr y combinaciones que se pueden obtener mediante mixturas de estas o ajustes a trozos. Todas ellas presentan colas más pesadas que la exponencial con el fin de poder captar la leptocurtosis observada en las distribuciones empíricas de pérdidas por riesgo operacional.

El hecho de que el valor de N sea aleatorio aumenta la dispersión de la variable aleatoria que describe la pérdida agregada. Los momentos de la distribución de pérdidas agregadas son fáciles de calcular. Sin embargo, para el cálculo de los percentiles que determinan el VaR operacional o CaR por riesgo operacional no existen expresiones analíticas exactas hasta la fecha.

3 Consideraciones metodológicas

El problema crucial en la modelización del riesgo operacional es el comportamiento extremo de las colas de las distribuciones de severidad. En esta sección se examinan las cuestiones que es necesario abordar para realizar un tratamiento estadístico de los datos correcto. Se hace también una breve exposición de la teoría de valores extremos y su aplicación a la modelización del riesgo operacional.

3.1 EL PROBLEMA DE LAS COLAS PESADAS

Dado que la cifra de capital económico (CaR) corresponde a un percentil muy elevado (para un nivel de probabilidad del 99,9%), de la distribución agregada de pérdidas, una modelización de la cola de la distribución es crucial para que los valores de CaR estimados sean correctos, especialmente cuando las pérdidas tienen una frecuencia baja. Los experimentos que se presentan en este artículo ilustran cómo a partir de muestras de datos de pérdidas de tamaño reducido puede ser difícil distinguir entre modelos cuyo comportamiento asintótico (es decir, en la cola de la distribución) es muy dispar.

En los experimentos se simulan T datos de pérdidas. A partir de esta muestra de tamaño finito, se estiman los parámetros de la distribución de pérdidas postulada. Se realizan simulaciones con muestras de distinto tamaño con el fin de investigar la dependencia de los estimadores de dichos parámetros con el tamaño de la muestra. Algunos experimentos corresponden a una única simulación. En otros se han realizado M simulaciones (habitualmente $M = 100$) con el fin de mostrar la diversidad de valores que se pueden obtener. Dada la gran variabilidad en los resultados, se ha optado por usar la mediana y el rango intercuartílico (entre corchetes en los cuadros) de las M simulaciones como estadísticos descriptivos de los resultados.

A lo largo del artículo hemos supuesto que la frecuencia de sucesos de pérdidas sigue una distribución de Poisson con distintos valores de λ . Suponiendo que los datos de pérdidas abarcan un período de 5 años, como se especifica en Basilea II, el valor del parámetro de la

distribución Poisson estimado a partir una muestra de tamaño T es la frecuencia anual media $\lambda = T/5$. Este es el valor que se ha utilizado en las simulaciones para calcular las medidas de riesgo.

Para la distribución de severidad se simulan datos con distribuciones leptocúrticas de dos tipos: En primer lugar, consideraremos distribuciones lognormales. En segundo lugar, distribuciones con cuerpo lognormal y colas de tipo Pareto. Para ciertas combinaciones de valores de parámetros, estas distribuciones son muy similares entre sí, excepto por comportamiento asintótico, y, para muestras de pequeño tamaño, pueden ser confundidas.

En todos los experimentos, las condiciones de las simulaciones y los parámetros de las distribuciones han sido elegidos de forma que correspondan a modelos para datos de pérdidas reales.

3.2 TEORÍA DE VALORES
EXTREMOS APLICADA AL RIESGO
OPERACIONAL

Sean X_1, \dots, X_n , n variables aleatorias, independientes e idénticamente distribuidas con función de distribución F . La teoría «clásica» de probabilidades (Teorema Central del Límite, Ley de los Grandes Números) se centra en el estudio de las sumas (o de las medias correspondientes):

$$S_n = \sum_{i=1}^n X_i$$

La distribución de esta suma de variables aleatorias independiente es normal (en el límite) cuando F tiene momento de orden 2 finito. Cuando ese no es el caso, existe un resultado asintótico y da lugar a las distribuciones α -estables ($0 < \alpha < 2$) en caso contrario. Los resultados mencionados tienen que ver con el comportamiento medio de las X_i .

La teoría de valores extremos¹ se centra en el estudio de los comportamientos extremos de variables aleatorias. Describe las variables aleatorias:

$$M_n = \max(X_1, \dots, X_n)$$

Concretamente, cuando x tiende a infinito, y para una clase muy amplia de variables aleatorias, se tiene que:

$$F_u(x) = P(X-u / X>u) \approx 1 - \left(1 + \xi \frac{x}{\beta}\right)^{-1/\xi} = G_{\xi, \beta}(x)$$

Es decir, los excesos por encima de u se distribuyen como una distribución de Pareto generalizada² (véase apéndice A). Varios autores han estudiado la aplicación de la teoría de valores extremos para la medición del capital por riesgo operacional³.

4 El riesgo de modelo

Como se ha señalado, para muestras reales, debido a la escasez de datos de cola, puede resultar difícil distinguir entre distribuciones de pérdidas similares para la mayor parte del rango de pérdidas, pero con distinto comportamiento asintótico. Con el fin de ilustrar esta situación, se han considerado dos casos de estudio para la distribución de severidad: el de una distribución lognormal con parámetros⁴ $\mu = 5$ y $\sigma = 2$ y el de una distribución con cola más

1. Véase Gumbel (1935) para el trabajo seminal y Embrechts et al. (1997) o Kotz et al. (2000) para una exposición enfocada a su uso en finanzas o gestión de riesgos. 2. Peak Over Threshold o POT en la literatura en inglés. 3. Véase, por ejemplo, Moscadelli (2004), para un primer estudio con los datos del Comité de Basilea, Mignola y Ugoccioni (2005) y Dutta y Perry (2006). 4. Estos parámetros son compatibles con situaciones reales en riesgo operacional (suponiendo que las pérdidas están expresadas en miles de euros) y similares a los usados por otros autores [véase Mignola y Ugoccioni (2005) o Moscadelli (2004)].

T	DISTRIBUCIÓN	MEDIA	DESVIACIÓN TÍPICA	COEFICIENTE DE ASIMETRÍA	CURTOSIS	MÁXIMO (X 10 ⁻³)
1.000	LN	1.056 [956, 1.174]	4.491 [3.504, 6.205]	13 [10, 18]	204 [120, 385]	90 [57, 146]
	LN + Pareto	1.070 [961, 1.217]	4.954 [3.587, 7.558]	14 [10, 20]	255 [139, 471]	104 [63, 188]
10.000	LN	1.085 [1.044, 1.130]	5.940 [5.079, 7.189]	26 [20, 38]	1.019 [586, 2.077]	297 [210, 456]
	LN + Pareto	1.148 [1.090, 1.242]	8.503 [6.406, 13.022]	41 [28, 63]	2.238 [1.075, 4.752]	544 [331, 1.060]
100.000	LN	1.096 [1.081, 1.112]	6.966 [6.274, 7.999]	51 [38, 75]	4.599 [2.549, 10.083]	873 [644, 1.275]
	LN + Pareto	1.186 [1.149, 1.232]	13.709 [10.403, 20.685]	118 [78, 183]	19.637 [911, 41.953]	2.590 [1.583, 5.067]

a. Mediana y rango intercuartílico para los diversos estadísticos calculados en las 1.000 simulaciones del experimento 1.

pesada obtenida a trozos a partir de esta misma lognormal para la modelización del cuerpo y de una distribución de Pareto, con parámetros⁵ $u_0 = 1.930$; $\beta_0 = 2.300$ y $\xi_0 = 0,7$, para modelizar el 10% de los datos de cola.

Para este primer experimento (*experimento 1*) se han simulado datos de pérdidas con las dos distribuciones consideradas. A partir de muestras de distinto tamaño ($T = 1.000$, 10.000 y 100.000) se estiman los valores de la media, desviación típica, coeficiente de asimetría y curtosis. Estos tamaños podrían corresponder a datos de pérdidas recogidos durante aproximadamente 5 años, cuya frecuencia anual fuera $\lambda = 200$, 2.000 y 20.000 , respectivamente. Los valores tabulados corresponden a $M = 1.000$ simulaciones. Como se puede constatar, exceptuando la media, los demás estadísticos son muy poco estables. Adicionalmente, los intervalos de estimación para dichos estadísticos a partir de muestras de pequeño tamaño presentan un gran solapamiento, excepto cuando la muestra es grande ($T = 10.000$). Todo ello debe llevarnos a extremar las precauciones a la hora de elegir los estadísticos que se vayan a usar para describir los datos y en la interpretación de sus valores, sobre todo en muestras de tamaño reducido.

4.1 EFECTOS ASOCIADOS A LA ELECCIÓN DEL UMBRAL DE AJUSTE

La elección de un umbral elevado puede tener un primer impacto inmediato en el ajuste de la distribución de severidad. Para ilustrar esta afirmación, se han generado 30.000 números aleatorios⁶ distribuidos según una distribución lognormal para diversos valores de σ y $\mu = 0$. El valor de μ no afecta a las conclusiones obtenidas, ya que es solo un factor de escala.

En el cuadro 2 se reflejan los ajustes por máxima verosimilitud de las distribuciones obtenidas por ajuste a datos censurados para umbrales de 6.000 y 10.000 euros. Para cada uno de los valores de σ se muestra la distribución que ajusta mejor, de acuerdo con los test estadísticos de Kolmogorov-Smirnov y de Anderson-Darling (en cada caso se elige aquel test que discrimine mejor entre las distintas alternativas).

Estos resultados ponen de manifiesto la dificultad de extrapolar la forma de la distribución de severidad completa a partir del conocimiento de su cola derecha.

En realidad, la elección del umbral también tiene un efecto sobre el ajuste de la distribución de frecuencia: los datos observados son solo una parte de los datos totales, por lo que la fre-

5. Se ha elegido un parámetro de forma (ξ) compatible con situaciones correspondientes a datos reales [véase Ferreras (2005)]. 6. La herramienta utilizada para realizar los cálculos ha sido Matlab, excepto en aquellas simulaciones en las que se indique expresamente.

σ	6.000	10.000
1,00	Weibull	Weibull
1,25	Pareto	Pareto
1,50	Lognormal	GEV
1,75	Lognormal	Lognormal
2,00	Weibull	Lognormal
2,25	Lognormal	Pareto
2,50	Lognormal	Lognormal

cuencia real es mayor que la frecuencia observada, ya que esta corresponde únicamente a pérdidas por encima del umbral. Por ejemplo, suponiendo un modelo de Poisson para la frecuencia, un umbral de recogida u , una frecuencia anual de pérdidas por encima de dicho umbral de λ_U y que F sea la función de distribución de la variable aleatoria X que representa la severidad⁷, el valor extrapolado para la frecuencia de pérdidas es:

$$\hat{\lambda} = \frac{\lambda_U}{1 - F(x)}$$

La combinación de ambos efectos en la distribución agregada de pérdidas puede tener consecuencias dramáticas en el cálculo del CaR definido como el percentil 99,9 de dicha distribución agregada o del CaR condicional, el valor medio de la pérdida una vez que se supera el CaR: una medida de riesgo que tiene la ventaja de ser coherente⁸.

Para ilustrar este aspecto, se ha realizado el *experimento 2*, utilizando datos de pérdidas simulados para un periodo de aproximadamente 5 años que siguen una distribución lognormal ($\mu = 5$, $\sigma = 2$). Se han utilizado distintos valores para el umbral de ajuste y se ha supuesto que la distribución de frecuencias es de tipo Poisson con un parámetro λ (frecuencia anual).

En el cuadro 3.A se presentan los valores estimados de los parámetros, tanto de la distribución de frecuencia como de severidad obtenidos usando diversos umbrales de ajuste. El número de puntos utilizados en el ajuste (el número de datos por encima del umbral u de ajuste) está indicado en la primera columna (N). La segunda columna del cuadro muestra la frecuencia anual de pérdidas por encima del umbral. Con el fin de obtener la frecuencia total de pérdidas, es necesario hacer un ajuste a la distribución de severidad de forma que se pueda estimar la probabilidad de pérdidas por debajo del umbral. Los parámetros de la distribución de pérdidas han sido obtenidos mediante un ajuste a una lognormal mediante una optimización por máxima verosimilitud con restricciones en los parámetros ($\mu \in [2, 8]$, $\sigma \in [1, 3]$) con el fin de evitar la generación de valores poco razonables del capital económico.

Para poder interpretar estos valores, en el cuadro 3.B se representan los valores del CaR (segunda columna) y CaR condicional (cuarta columna) para el valor real de λ , por lo que los errores tienen su origen exclusivamente en los errores de estimación de los parámetros de la distribución de severidad.

Las columnas 3 y 5 muestran el CaR y el CaR condicional utilizando el valor de $\hat{\lambda}$ extrapolado mediante la fórmula anterior, por lo que tienen en cuenta los errores de estimación de los pa-

7. Es decir, que $F(x) = P(X \leq x)$. 8. Véase Artzner et al. (1999).

T	λ_u	u	$\hat{\lambda}$	μ	σ
1.000	200	0	200	5,01 [4,97, 5,05]	2,00 [1,97, 2,03]
500	100	149 [143, 160]	195 [173, 235]	5,01 [4,66, 5,33]	1,98 [1,86, 2,06]
250	50	580 [551, 610]	184 [125, 311]	5,24 [4,16, 5,98]	1,97 [1,73, 2,29]
200	40	810 [764, 869]	210 [125, 385]	4,97 [3,58, 5,96]	2,00 [1,67, 2,37]
150	30	1.227 [1.106, 1.290]	191 [103, 561]	5,18 [3,26, 6,21]	1,97 [1,62, 2,34]
100	20	1.964 [1.765, 2.124]	170 [78, 505]	5,17 [3,39, 6,63]	1,93 [1,52, 2,31]

a. Parámetros del ajuste con 1.000 datos de pérdidas (5 años de pérdidas, $\lambda = 200$), generados a partir de una distribución lognormal ($\mu = 5$; $\sigma = 2$) para distintos valores del umbral de ajuste.

MEDIDAS DE RIESGO PARA LOS DIVERSOS AJUSTES DE UNA LOGNORMAL (a)

T	$CaR_i \times 10^{-3}$	$CaR_i \times 10^{-3}$	$CCaR_i \times 10^{-3}$	$CCaR_i \times 10^{-3}$
1.000	1.216 [1.125, 1.432]		1.886 [1.733, 2.247]	
500	1.218 [993, 1.559]	1.203 [943, 1.704]	1.855 [1.451, 2.560]	1.838 [1.377, 2.818]
250	1.150 [877, 1.462]	1.231 [778, 1.768]	1.728 [1.306, 2.535]	1.896 [1.103, 3.025]
200	1.121 [899, 1.421]	1.207 [779, 1.853]	1.786 [1.274, 2.391]	1.871 [1.088, 3.190]
150	1.147 [812, 1.417]	1.161 [662, 1.840]	1.655 [1.272, 2.353]	1.808 [905, 3.329]
100	1.166 [843, 1.537]	1.163 [674, 1.660]	1.631 [1.113, 2.362]	1.810 [910, 2.737]

a. Valores del CaR y del CaR condicional obtenidos a partir de los ajustes con 1.000 datos de pérdidas (aproximadamente 5 años de pérdidas, $\lambda = 200$) generados a partir de una distribución lognormal ($\mu = 5$, $\sigma = 2$) para distintos valores del umbral de ajuste. Los valores teóricos correspondientes (escalados por un factor 1.000) son 1.251 y 1.945.

rámetros de la distribución de severidad y de frecuencia. El valor teórico del CaR es de 1.251.000, mientras que el del CaR condicional es de 1.945.000.

4.2 PARETO FRENTE A LOGNORMAL

En las etapas iniciales de Basilea II algunos expertos propugnaron el uso del binomio lognormal/Poisson para la medición de la contribución al capital en riesgo por riesgo operacional. Esta propuesta intentaba replicar la práctica habitual en la estimación del VaR normal. Hasta la fecha, este planteamiento no se ha impuesto.

Más recientemente, numerosos expertos han propuesto que la distribución correcta para modelizar la severidad de los datos de pérdidas por encima de un umbral suficientemente alto es una distribución de Pareto y que el modelo lognormal no debería ser utilizado. En el caso de encontrarnos en el régimen asintótico, esta observación sería correcta y estaría justificada por la teoría de valores extremos. Sin embargo, la cuestión de si el rango de valores observado en pérdidas por riesgo operacional corresponde a un régimen asintótico en el que la teoría de valores extremos es aplicable no está resuelta de manera definitiva.

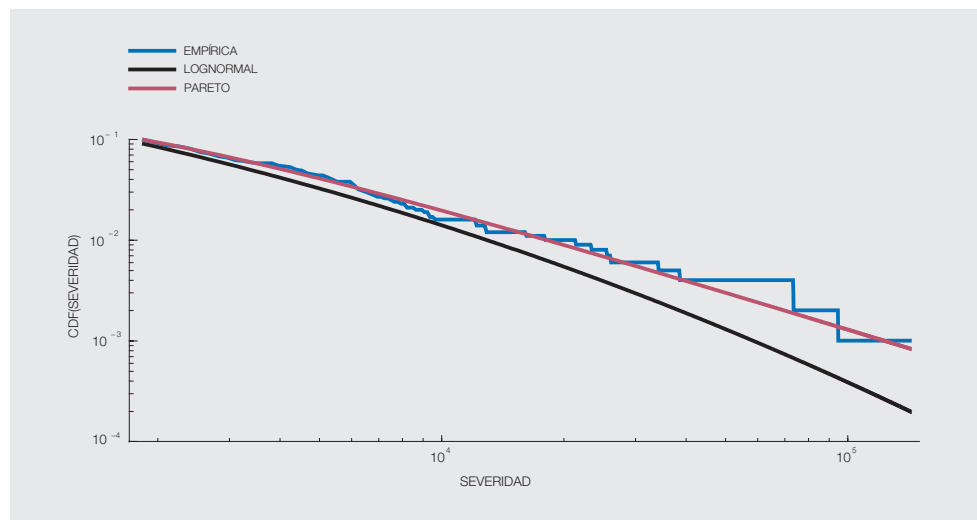
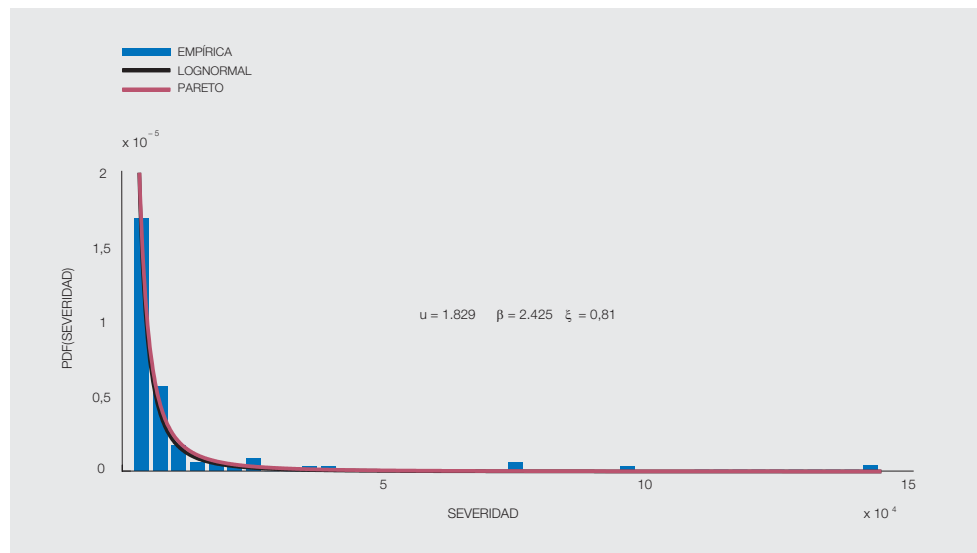
Sin entrar en la cuestión de si el modelo correcto es el que utiliza una distribución lognormal para todo el rango de severidades, o el que modeliza la cola de la distribución de severidad mediante una Pareto, el objetivo de esta sección es poner de manifiesto la importancia de determinar correctamente el modelo para la distribución de severidad de pérdidas. Hay dos casos opuestos. En el primero, los datos de severidad tienen un comportamiento Pareto, pero

T	T _{cola}	U	β	ξ
250	25	1.887 [1.653, 2.115]	2.319 [1.748, 3.115]	0,60 [0,43, 0,80]
500	50	1.987 [1.804, 2.195]	2.477 [2.069, 2.790]	0,65 [0,47, 0,81]
1.000	100	1.963 [1.843, 2.082]	2.242 [2.004, 2.577]	0,72 [0,61, 0,84]
10.000	1.000	1.927 [1.880, 1.975]	2.331 [2.253, 2.406]	0,70 [0,67, 0,73]
100.000	10.000	1.928 [1.915, 1.944]	2.323 [2.284, 2.352]	0,70 [0,69, 0,71]

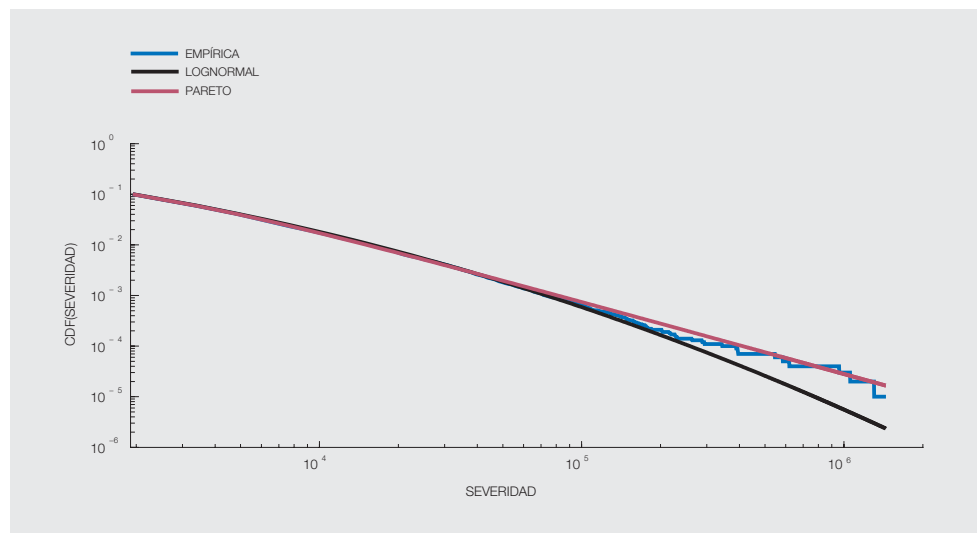
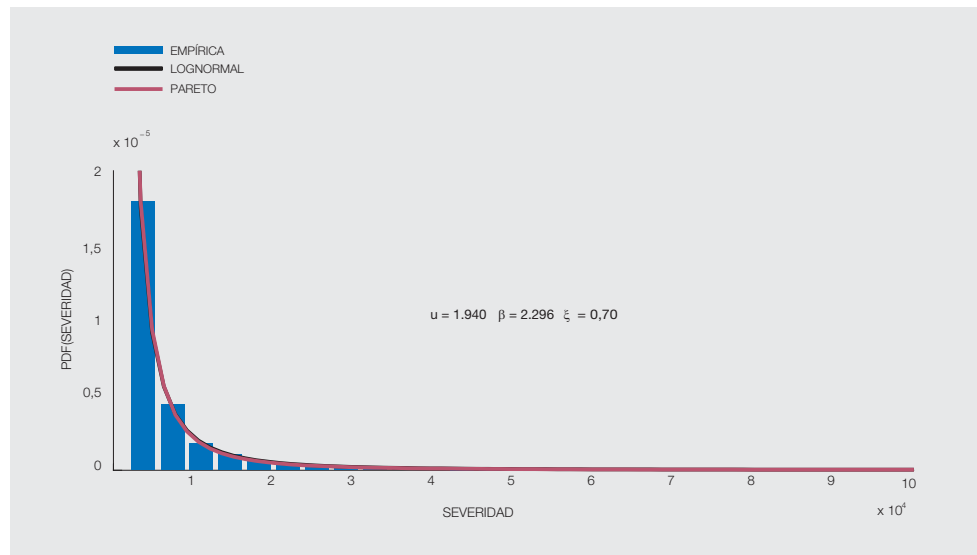
a. Ajuste a una distribución de Pareto de datos generados a partir de una distribución lognormal con parámetros $\mu = 5$, $\sigma = 2$. Los valores tabulados corresponden a la mediana y al espaciado intercuartil para $M = 100$ simulaciones con $p = 0,1$.

LOGNORMAL FRENTE A PARETO PARA UNA MUESTRA CON 1.000 DATOS (a)

GRÁFICO 1



a. Comparación entre los ajustes a una distribución lognormal y una distribución de Pareto ($p = 0,1$) a partir de $T = 10^3$ datos de pérdidas (aproximadamente 5 años de datos con $\lambda = 200$) con una distribución lognormal de parámetros $\mu = 5$ y $\sigma = 2$.



a. Comparación entre los ajustes a una distribución lognormal y una distribución de Pareto ($p = 0,1$) a partir de $T = 10^5$ datos de pérdidas (aproximadamente 5 años de datos con $\lambda = 20.000$) con una distribución lognormal de parámetros $\mu = 5$ y $\sigma = 2$.

han sido modelizados mediante una distribución lognormal. Este error de modelo lleva a subestimar las medidas de riesgo. En el caso opuesto, es decir, si los datos tienen un comportamiento lognormal, pero son modelizados mediante una distribución de Pareto, las medidas de riesgo son sobreestimadas.

El objetivo del *experimento 3* es poner de manifiesto la dificultad de determinar a partir de una muestra de tamaño reducido si los datos tienen un comportamiento lognormal o Pareto en la cola. A partir de datos de severidad generados mediante una distribución lognormal de parámetros $\mu = 5$ y $\sigma = 2$, se han realizado un primer ajuste suponiendo un modelo lognormal para todo el rango de pérdidas y un segundo ajuste a una distribución de Pareto para la cola, definida como conteniendo del 10% de los datos. En el cuadro 4 se muestran los resultados de dicho ajuste utilizando un número muestras de datos de pérdidas de distinto tamaño.

p	u	β	ξ	$CaR_x \times 10^{-3}$	$CCaR_x \times 10^{-3}$
0,25	566 [548, 609]	988 [939, 1.089]	0,80 [0,71, 0,86]	7.160 [3.238, 12.328]	31.165 [10.340, 67.661]
0,20	804 [769, 856]	1.247 [1.142, 1.363]	0,77 [0,67, 0,84]	5.269 [2.623, 10.651]	21.036 [7.570, 53.648]
0,15	1.189 [1.144, 1.255]	1.615 [1.469, 1.829]	0,70 [0,62, 0,81]	3.523 [1.867, 8.582]	11.339 [4.367, 41.566]
0,10	1.953 [1.810, 2.051]	2.360 [2.032, 2.625]	0,67 [0,54, 0,77]	2.771 [1.328, 6.081]	7.995 [2.561, 25.605]
0,05	4.031 [3.604, 4.434]	4.017 [3.342, 4.807]	0,59 [0,46, 0,78]	2.023 [784, 5.926]	4.835 [1.254, 21.806]
0,03	6.319 [5.930, 7.017]	5.414 [4.312, 7.090]	0,53 [0,38, 0,72]	1.333 [687, 3.868]	2.558 [978, 11.226]

a. Resultados obtenidos en $M = 100$ simulaciones. En cada simulación se han generado $T = 1.000$ datos a partir de una distribución de severidad lognormal $\mu = 5$, $\sigma = 2$ (experimento 3). Las columnas segunda a cuarta muestran los parámetros obtenidos en el ajuste a una distribución de Pareto para la cola (10% de los datos). Las columnas quinta y sexta presentan las medidas de riesgo (CaR y CaR condicional, escalados por un factor de 1.000) calculadas utilizando una distribución de Poisson ($\lambda = 200$) para las frecuencias. En el caso de que se realicen ajustes a una distribución lognormal, los valores calculados son 1.202 ([1.095, 1.374]) para el CaR y 1.863 ([1.676, 2.152]) para el CaR condicional. El valor teórico es 1.251 para el CaR y 1.945 para el CaR condicional.

p	u	β	ξ	$CaR_x \times 10^{-3}$	$CCaR_x \times 10^{-3}$
0,25	573 [560, 584]	978 [943, 1.004]	0,81 [0,79, 0,84]	51.364 [38.911, 75.292]	215.248 [150.680, 352.693]
0,20	800 [779, 815]	1.230 [1.181, 1.271]	0,78 [0,75, 0,81]	37.268 [27.625, 50.405]	139.140 [95.318, 207.758]
0,15	1.178 [1.156, 1.205]	1.615 [1.565, 1.690]	0,74 [0,71, 0,78]	25.379 [19.028, 35.893]	84.275 [56.385, 134.049]
0,10	1.931 [1.872, 1.976]	2.308 [2.221, 2.417]	0,70 [0,66, 0,73]	18.363 [12.744, 25.441]	53.172 [32.302, 82.072]
0,05	3.982 [3.887, 4.079]	4.041 [3.837, 4.296]	0,64 [0,59, 0,68]	12.026 [7.676, 16.951]	29.084 [16.355, 47.923]
0,03	6.368 [6.190, 6.561]	5.891 [5.392, 6.355]	0,60 [0,52, 0,67]	8.920 [5.714, 15.077]	19.485 [9.582, 39.857]

a. Resultados obtenidos en $M = 100$ simulaciones. En cada simulación se han generado $T = 10.000$ datos a partir de una distribución de severidad lognormal $\mu = 5$, $\sigma = 2$ (experimento 3). Las columnas segunda a cuarta muestran los parámetros obtenidos en el ajuste a una distribución de Pareto para la cola (10% de los datos). Las columnas quinta y sexta presentan las medidas de riesgo (CaR y CaR condicional, escalados por un factor de 1.000) calculadas utilizando una distribución de Poisson ($\lambda = 2.000$) para las frecuencias. En el caso de que se realicen ajustes a una distribución lognormal, los valores calculados son 4.878 ([4.668, 5.018]) para el CaR y 6.429 ([6.117, 6.630]) para el CaR condicional. El valor teórico es 4.912 para el CaR y 6.464 para el CaR condicional.

Los gráficos 1 y 2 muestran los resultados de dichos ajustes en dos casos particulares en los que se utilizaron $T = 1.000$ y 100.000 datos de pérdidas para el ajuste, respectivamente.

Con el fin de cuantificar los efectos de un error de modelo consistente en ajustar una distribución de Pareto a datos con comportamiento lognormal en las medidas de riesgo, se ha realizado el siguiente experimento: suponiendo que se han recogido aproximadamente 5 años de datos de pérdidas, se han generado $T = 1.000$ datos de severidad para una distribución de severidad lognormal de parámetros $\mu = 5$ y $\sigma = 2$. Utilizando estos datos simulados se realizan ajustes por máxima verosimilitud a una distribución de Pareto para distintos posibles inicios de la zona de cola derecha, definida como conteniendo distintas proporciones p de los mismos. Para cada uno de los valores del umbral a partir del cual se considera que los datos siguen una distribución de Pareto se han calculado las distintas medidas de riesgo (CaR, CaR condicional). En el cuadro 5 se resumen los resultados de estos ajustes y los valores correspondientes para el CaR y CaR condicional, en el supuesto de una distribución de Poisson ($\lambda = 200$) para la frecuencia.

Los valores teóricos son, respectivamente, 1.251.000 y 1.945.000. Es decir, que el error de modelo inducido por el uso de un ajuste Pareto lleva a sobreestimar en varias veces el capital económico «real».

Para confirmar que el origen de estas desviaciones reside realmente en el error de modelo y no en el uso de un número relativamente pequeño de datos, se ha repetido el experimento usando 10.000 datos (aproximadamente 5 años de datos con $\lambda = 2.000$) con los resultados que se pueden observar en el cuadro 6. En este caso los valores teóricos del CaR y CaR condicional son, respectivamente, 4.912.000 y 6.464.000.

4.3 USO INCORRECTO DEL RÉGIMEN ASINTÓTICO

El *experimento 4* tiene como fin estimar el error que se puede estar cometiendo cuando no se dispone de la información relativa a la parte de la distribución de severidad situada por debajo del umbral de modelización.

Para ello, se han generado datos según una distribución lognormal ($\mu = 5$, $\sigma = 2$). Se han definido varios posibles valores para el umbral u de ajuste, fijando la probabilidad p de que la severidad de la pérdida sea mayor que el umbral.

La cola derecha de la distribución es modelizada mediante una distribución lognormal. Para el cuerpo se han considerado dos casos extremos: las pérdidas por debajo del umbral tienen todas severidad 0 (caso 1) o severidad igual al valor del umbral (caso 2). El caso de referencia consiste en utilizar los datos empíricos generados por la distribución empírica sin manipulación alguna (caso 0). En los tres casos se ha mantenido constante la masa de probabilidad por debajo del umbral.

A partir de estas distribuciones modificadas, y para distintos valores de λ , se calculan las medidas de riesgo operacional:

- Los valores CaR_0 y CCaR_0 corresponden al CaR y CaR condicional obtenido utilizando la distribución empírica para pérdidas por debajo del umbral de cola.
- Los valores CaR_- y CCaR_- corresponden al CaR y CaR condicional para el caso 1.
- Los valores CaR_+ y CCaR_+ corresponden al CaR y CaR condicional para el caso 2.

Los resultados de este experimento están recogidos en el cuadro 7. Obsérvese que, con el fin de analizar la aportación del cuerpo de la distribución a la hora de calcular el capital, se han usado los «verdaderos» valores ($\mu = 5$, $\sigma = 2$) para generar los datos de cola y obviar así las posibles fluctuaciones debidas a la escasez de datos para un ajuste robusto de la misma.

Esta precaución, junto con el hecho de haber utilizado la misma masa de probabilidad para el cuerpo de la distribución en los tres casos, permite centrar el análisis en el impacto de la modelización de dicho cuerpo (o de la falta de información relativa al mismo) en la estimación final del CaR.

Como se puede observar, para valores bajos de λ y umbrales no demasiado elevados, las fluctuaciones del CaR o CCaR debidas a la falta de información relativa al comportamiento de las pérdidas por debajo del umbral no son significativas. En estas situaciones, el capital está realmente determinado por un pequeño número de eventos de alta severidad y bajo impacto.

	p	u	CaR ₀ × 10 ⁻³	CaR ₋ × 10 ⁻³	CaR ₊ × 10 ⁻³	$\frac{CaR_+ - CaR_-}{CaR_0} \times 100$	CCaR ₀ × 10 ⁻³	CCaR ₋ × 10 ⁻³	CCaR ₊ × 10 ⁻³	$\frac{CaR_+ - CaR_-}{CaR_0} \times 100$
λ = 200	1,00	0	1.252				1.945			
	0,50	149	1.253	1.249	1.263	1,14	1.947	1.943	1.958	0,76
	0,25	567	1.247	1.227	1.313	6,85	1.941	1.920	2.005	4,35
	0,20	797	1.249	1.223	1.349	10,09	1.944	1.917	2.044	6,49
	0,15	1.174	1.245	1.212	1.414	16,17	1.938	1.907	2.107	10,32
	0,10	1.946	1.256	1.204	1.557	28,05	1.952	1.901	2.252	18,01
λ = 2.000	1,00	0	4.903				6.461			
	0,50	149	4.909	4.858	5.010	3,10	6.463	6.410	6.564	2,37
	0,25	567	4.883	4.691	5.530	17,19	6.438	6.241	7.081	13,05
	0,20	797	4.896	4.624	5.897	25,98	6.453	6.178	7.450	19,71
	0,15	1.174	4.893	4.516	6.524	41,05	6.449	6.071	8.075	31,08
	0,10	1.946	4.911	4.399	7.903	71,36	6.474	5.959	9.464	54,14
λ = 20.000	1,00	0	28.620				31.872			
	0,50	149	28.655	28.126	29.651	5,32	31.915	31.393	32.901	4,73
	0,25	567	28.479	26.461	34.973	29,89	31.742	29.728	38.231	26,79
	0,20	797	28.567	25.853	38.660	44,83	31.832	29.120	41.914	40,19
	0,15	1.174	28.524	24.870	44.843	70,02	31.776	28.132	48.096	62,82
	0,10	1.946	28.727	23.519	58.630	122,22	32.009	26.803	61.899	109,64

a. Influencia del cuerpo de la distribución empírica en el cálculo de las medidas de riesgo, CaR y CaR condicional (experimento 4).
p representa la probabilidad de la cola derecha y u el umbral correspondiente.

Sin embargo, a medida que λ se incrementa, el cuerpo de la distribución tiene mayor influencia en las estimaciones de las medidas de riesgo, por lo que es necesario utilizar umbrales inferiores para definir la cola de manera que se obtengan medidas de capital satisfactorias. Con 2.000 eventos de pérdida, usando el 10% de los datos para definir la cola y suponiendo que no existe ninguna otra fuente de error en el modelo, las fluctuaciones en la estimación son del 70%.

Esto tiene que ver con el hecho de que, en estos casos, también se llega a percentiles elevados de la función de distribución agregada mediante la suma de muchos de los eventos del cuerpo de la distribución de severidad.

A nuestro juicio, estos resultados ponen en cuestión la idea según la cual basta fijarse en los eventos de alta severidad y bajo impacto para poder hacer un cálculo correcto del capital regulatorio, especialmente para celdas con una frecuencia elevada de eventos de pérdidas.

5 Inestabilidad de los ajustes a distribuciones de Pareto

En la sección anterior se ha puesto de manifiesto cómo el uso de la teoría de valores extremos, fuera del régimen asintótico, puede producir estimaciones erróneas del capital económico. No es esta la única dificultad en el uso de modelos basados en dicha teoría. Por ejemplo, no es infrecuente que, al realizar ajustes de datos empíricos de severidad de pérdidas operacionales a una distribución de Pareto, el parámetro de forma (ξ) estimado sea mayor que 0,5, lo cual implica la inexistencia de los momentos de orden superior o igual a 2 o, incluso, mayor

que 1, situación en la cual ni siquiera existe el momento de orden uno⁹, lo cual implica que la pérdida inesperada es infinita, pero que también lo es la pérdida esperada.

El debate acerca del uso de este tipo de metodología en riesgo operacional ha estado presente en la literatura especializada desde hace algún tiempo¹⁰. Nuestro objetivo en esta sección no es tanto reproducir los argumentos manejados por otros autores como señalar algunos aspectos críticos relacionados con el uso de estas metodologías.

5.1 SENSIBILIDAD DE LOS PARÁMETROS AL UMBRAL DE AJUSTE

Las estimaciones de los valores de los parámetros de una distribución de Pareto obtenidas a partir de un ajuste con un número reducido de datos son muy inestables. En concreto, las fluctuaciones de los valores del parámetro de cola obtenidos por máxima verosimilitud o mediante otros procedimientos estándar son muy elevadas. Esta limitación introduce una gran incertidumbre en el cálculo de percentiles elevados de la distribución agregada de pérdidas y, por lo tanto, en las medidas de riesgo, que son el objetivo final de nuestro análisis. La elección de umbrales de cola distintos para realizar el ajuste a una distribución tipo Pareto puede conducir a estimar valores muy distintos para dichas medidas de riesgo.

Para ilustrar la magnitud de estas fluctuaciones se ha realizado el *experimento 5*, en el que se ha partido de una distribución de severidad cuyo cuerpo está modelizado por una distribución lognormal de parámetros $\mu = 5$ y $\sigma = 2$ y cuya cola derecha (el 10% superior) sigue una distribución de Pareto con parámetros $u^0 = 1.930$; $\beta^0 = 2.300$ y $\xi^0 = 0,7$ (idéntica a la utilizada en el experimento 1).

Para diversos valores de p , se ha ajustado una distribución de Pareto a los datos correspondientes a la cola derecha, definida por la proporción p de datos que contiene. Para cada uno de estos valores se han calculado las distintas medidas de riesgo (CaR, CaR condicional), suponiendo $\lambda = 200$. En el cuadro 8 se recogen los parámetros de la distribución de Pareto obtenida a partir de ajustes realizados para diferentes valores de p (lo cual equivale a tomar diferentes umbrales) a partir de una muestra de tamaño 1.000 de la lognormal señalada, y los valores correspondientes de CaR y CaR condicional suponiendo una frecuencia anual de pérdidas de $\lambda = 200$.

Conviene destacar la gran sensibilidad de las medidas de riesgo calculadas al umbral a partir del cual se define la cola (para realizar el ajuste por una distribución de Pareto), pese a la información gráfica suministrada por los gráficos 1 y 2 relativa a la bondad del ajuste de la cola de la lognormal por una Pareto. Los gráficos 3 y 4 evidencian dicha sensibilidad. Incluso en el rango 5%-10%, valores habituales en teoría de valores extremos para definir la cola, el rango intercuartílico supone un margen de más del 90% de variabilidad del capital en el valor del capital en riesgo.

Es bien sabido que, cuanto mayor sea el número de datos utilizado para realizar el ajuste a una distribución Pareto, mejores serán las estimaciones tanto de los parámetros de la distribución como de las medidas de riesgo. Pero ¿cómo de grande ha de ser la muestra para una estimación fiable de los parámetros?

En el cuadro 9 se muestran las estimaciones por máxima verosimilitud de los parámetros de una distribución de Pareto obtenidos a partir de una muestra de tamaño 10.000 y los valores correspondientes de CaR y CaR condicional suponiendo una frecuencia anual de pérdidas de

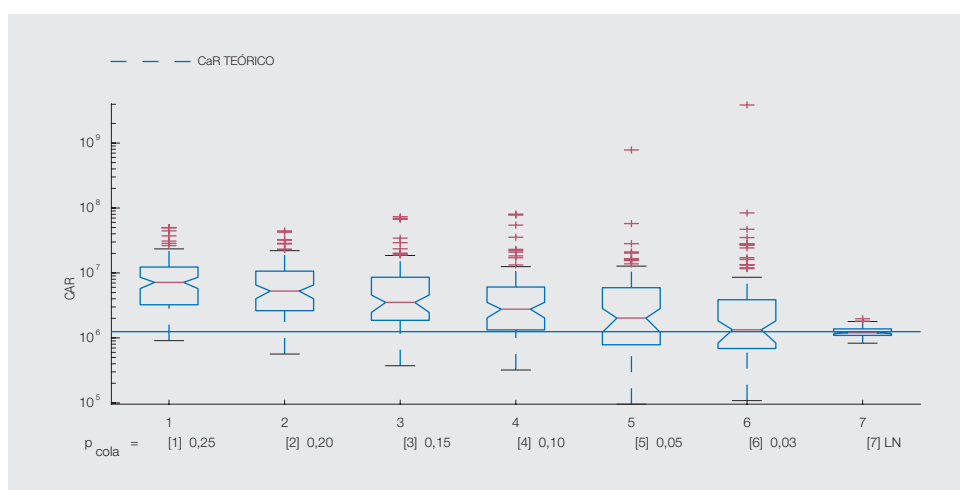
9. Véase, por ejemplo, Ferreras (2005) o Nešlehová et al. (2006). 10. Véanse Medova et al. (2001) y Moscadelli (2004); también, Embrechts et al. (2003), Nešlehová et al. (2006) y Mignola et al. (2006) para las dificultades a la hora de aplicar estos modelos.

p	u	β	ξ	$CaR_x \times 10^{-3}$	$CCaR_x \times 10^{-3}$
0,25	578 [530, 620]	982 [918, 1.061]	0,79 [0,72 , 0,87]	6.323 [3.649, 12.857]	27.010 [12.051, 73.065]
0,20	809 [751, 853]	1.251 [1.165, 1.345]	0,75 [0,67 , 0,84]	4.667 [2.599, 9.526]	17.162 [7.439, 48.392]
0,15	1.174 [1.129, 1.244]	1.680 [1.533, 1.815]	0,72 [0,64 , 0,78]	3.835 [2.050, 6.847]	12.849 [5.309, 27.423]
0,10	1.924 [1.912, 1.937]	2.334 [2.070, 2.656]	0,66 [0,54 , 0,77]	2.795 [1.058, 6.432]	7.992 [2.106, 24.817]
0,05	4.001 [3.741, 4.293]	3.915 [3.098, 4.477]	0,71 [0,53 , 0,84]	3.599 [1.239, 8.242]	11.299 [2.435, 41.304]
0,03	6.371 [5.706, 6.980]	5.621 [4.212, 7.326]	0,60 [0,39 , 0,83]	2.136 [654, 8.965]	5.216 [966, 44.126]

a. Resultados obtenidos en $M = 100$ simulaciones. En cada simulación se han generado $T = 1.000$ datos a partir de una distribución de severidad con cuerpo lognormal ($\mu = 5, \sigma = 2$) y cola Pareto (10% de los datos, $u = 1.930, \beta = 2.300, \xi = 0,7$). Las columnas segunda a cuarta muestran los parámetros obtenidos en el ajuste a una distribución de Pareto para la cola (10% de los datos). Las columnas quinta y sexta presentan las medidas de riesgo (CaR y CaR condicional, escalados por un factor de 1.000) calculadas utilizando una distribución de Poisson ($\lambda = 200$) para las frecuencias. En el caso de que se realicen ajustes a una distribución lognormal, los valores calculados son 1.238 ([1.100, 1.375]) para el CaR y 1.920 ([1.687, 2.163]) para el CaR condicional. Los valores teóricos son 3.604 y 11.401, para el CaR y el CaR condicional, respectivamente.

UMBRAL DE COLA Y CAPITAL EN RIESGO (a)

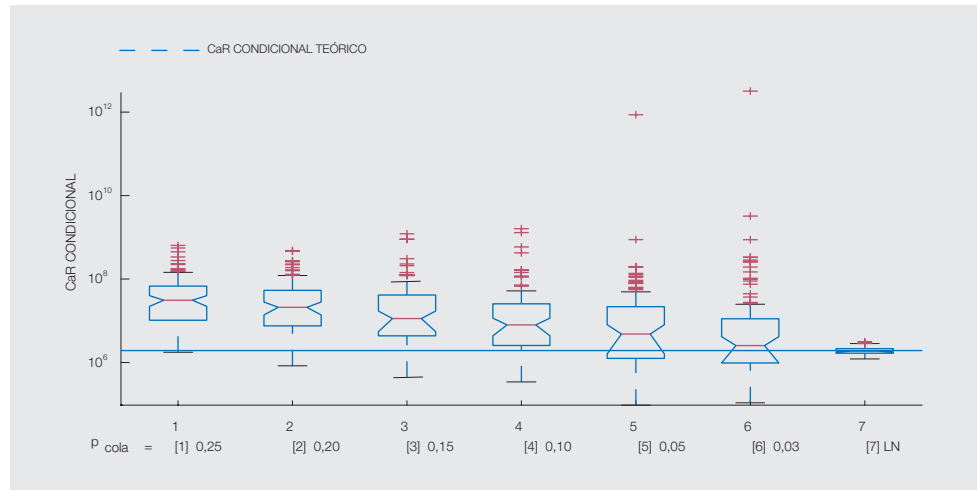
GRÁFICO 3



a. CaR en función de la elección del comienzo de la cola para datos de pérdidas simulados con una distribución de frecuencia tipo Poisson ($\lambda = 200$) y de severidad lognormal ($\mu = 5$ y $\sigma = 2$). Los datos de severidad han sido modelizados mediante una distribución de Pareto para distintos valores de la probabilidad de cola. El ajuste de la derecha se corresponde a un ajuste lognormal.

$\lambda = 2.000$. La mediana de las estimaciones en 100 experimentos para $p \geq 0,10$ se aproxima a los valores correctos para las medidas de riesgo (19.308.000 para el CaR, 58.385.000 para el CaR condicional), pero la dispersión de los valores obtenidos, estimada por el rango intercuartílico, indica que las fluctuaciones son aún muy elevadas y probablemente inaceptables para que las medidas de riesgo sean realmente útiles.

Para los valores de los parámetros de la distribución de severidad utilizados, únicamente cuando la muestra es de tamaño 100.000 (es decir, cuando se tienen al menos 10.000 observaciones a partir de un umbral de aproximadamente 1.930) se pueden tener estimaciones estables de las medidas de riesgo. Dicho de otra manera, para poder extrapolar la función agregada hasta percentiles tan altos como los que se requiere en Basilea II a partir de un ajuste Pareto a la distribución de severidad es preciso disponer de un gran número de datos por encima del umbral (al menos 10.000 datos en nuestro caso).



a. CaR condicional en función de la elección del comienzo de la cola para datos de pérdidas simulados con una distribución de frecuencia tipo Poisson ($\lambda = 200$) y de severidad lognormal ($\mu = 5$ y $\sigma = 2$). Los datos de severidad han sido modelizados mediante una distribución de Pareto para distintos valores de la probabilidad de cola. El ajuste de la derecha se corresponde a un ajuste lognormal.

AJUSTES A UNA PARETO DE DATOS CON COLA PARETO PARA DIFERENTES UMBRALES (II) (a)

CUADRO 9

p	u	β	ξ	$CaR_3 \times 10^{-3}$	$CCaR_3 \times 10^{-3}$
0,25	574 [562, 586]	987 [960, 1.009]	0,80 [0,79 , 0,83]	47.860 [40.119, 64.948]	195.584 [155.258, 288.098]
0,20	800 [784, 814]	1.242 [1.219, 1.262]	0,77 [0,75 , 0,79]	35.106 [26.459, 43.433]	128.875 [89.690, 171.711]
0,15	1.177 [1.160, 1.203]	1.619 [1.592, 1.669]	0,74 [0,71 , 0,76]	25.978 [17.884, 33.178]	86.332 [53.793, 118.091]
0,10	1.930 [1.928, 1.930]	2.314 [2.251, 2.384]	0,70 [0,65 , 0,74]	18.209 [12.317, 25.117]	53.059 [31.172, 82.851]
0,05	3.990 [3.920, 4.033]	3.728 [3.482, 3.975]	0,69 [0,63 , 0,76]	17.064 [10.070, 29.958]	49.569 [23.483, 105.616]
0,03	6.266 [6.122, 6.403]	5.393 [4.949, 5.767]	0,70 [0,64 , 0,76]	18.334 [11.071, 31.096]	53.600 [26.119, 106.176]

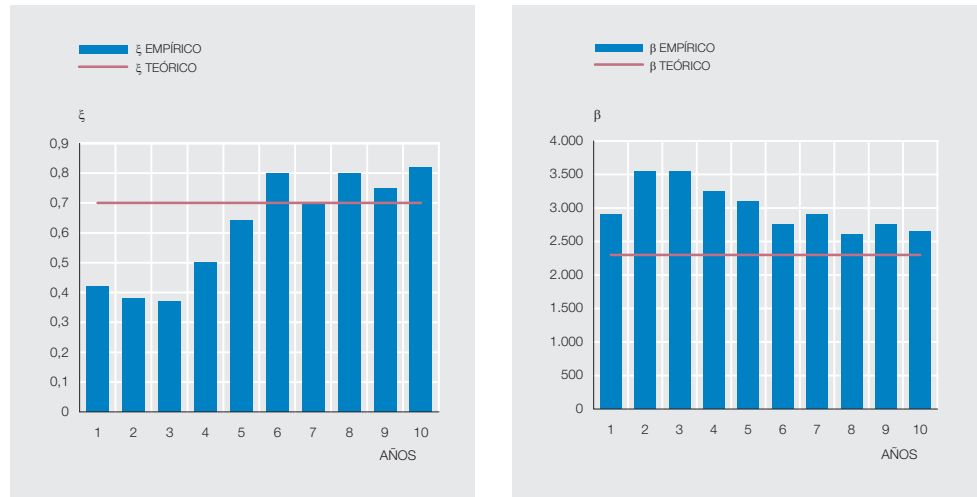
a. Resultados obtenidos en $M = 100$ simulaciones. En cada simulación se han generado $T = 10.000$ datos a partir de una distribución de severidad con cuerpo lognormal ($\mu = 5$, $\sigma = 2$) y cola Pareto (10% de los datos, $u = 1.930$, $\beta = 2.300$, $\xi = 0,7$). Las columnas segunda a cuarta muestran los parámetros obtenidos en el ajuste a una distribución de Pareto para la cola (10% de los datos). Las columnas quinta y sexta presentan las medidas de riesgo (CaR y CaR condicional, escalados por un factor de 1.000) calculadas utilizando una distribución de Poisson ($\lambda = 2.000$) para las frecuencias. En el caso de que se realicen ajustes a una distribución lognormal, los valores calculados son 4.874 ([4.715, 5.042]) para el CaR y 6.419 ([6.174, 6.682]) para el CaR condicional. Los valores teóricos son 19.308 y 58.385, para el CaR y el CaR condicional, respectivamente.

5.2 EFECTO DE LA INCORPORACIÓN DE NUEVOS DATOS DE PÉRDIDAS

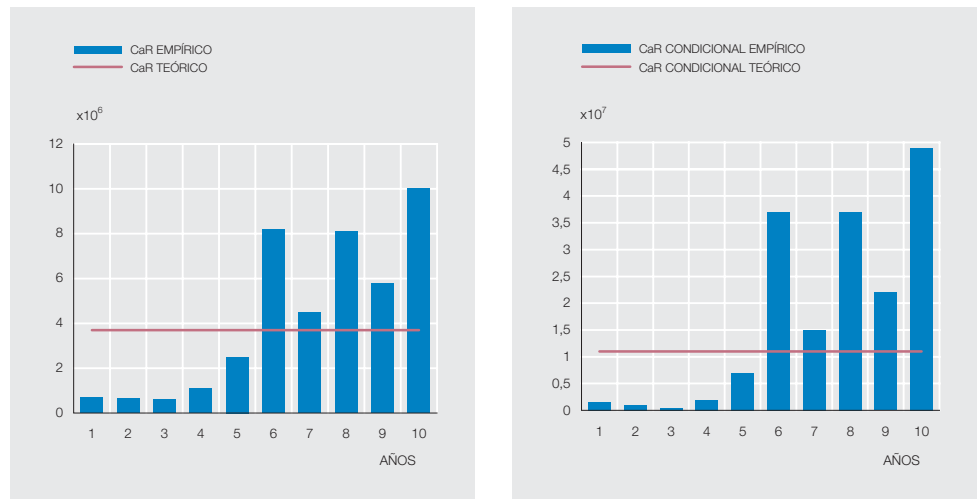
Otro elemento crítico a la hora de usar la distribución de Pareto para el ajuste de las colas de las distribuciones de severidad es que da lugar a ajustes poco robustos, cuyos resultados pueden fluctuar mucho en el tiempo a medida que se van incorporando nuevos eventos a la base de datos de pérdidas por riesgo operacional.

Para ilustrar este hecho, se ha diseñado el *experimento 6*, consistente en trazar la evolución del ajuste del índice de cola cuando la serie temporal (sintética) usada ha sido generada mediante una distribución cuyo cuerpo ha sido generado con una distribución lognormal ($\mu = 5$, $\sigma = 2$) y la cola derecha (el 10% de la derecha, correspondiente a un umbral $u = 1.930$) con una distribución Pareto ($\xi = 0,7$).

Los gráficos 5 a 7 presentan la evolución de los parámetros de la distribución de Pareto ajustada y de las medidas de riesgo a lo largo de diez años para diferentes valores del número



a. Evolución de los parámetros de ajustes de Pareto (ξ , β) para datos de pérdidas representando 10 años con una frecuencia anual de $\lambda = 200$ (experimento 6).

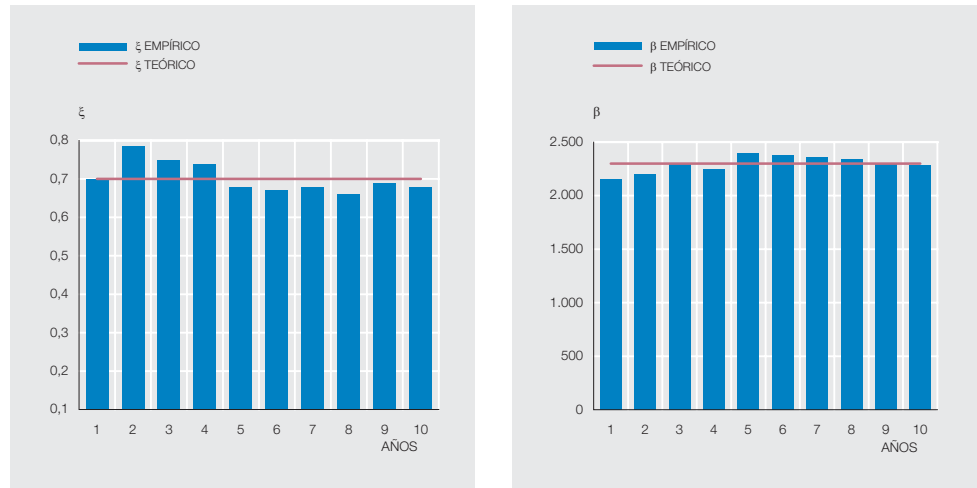


a. Evolución de las medidas de riesgo (CaR, CaR condicional), correspondientes al gráfico 5.A, para datos de pérdidas representando 10 años con una frecuencia anual de $\lambda = 200$ (experimento 6).

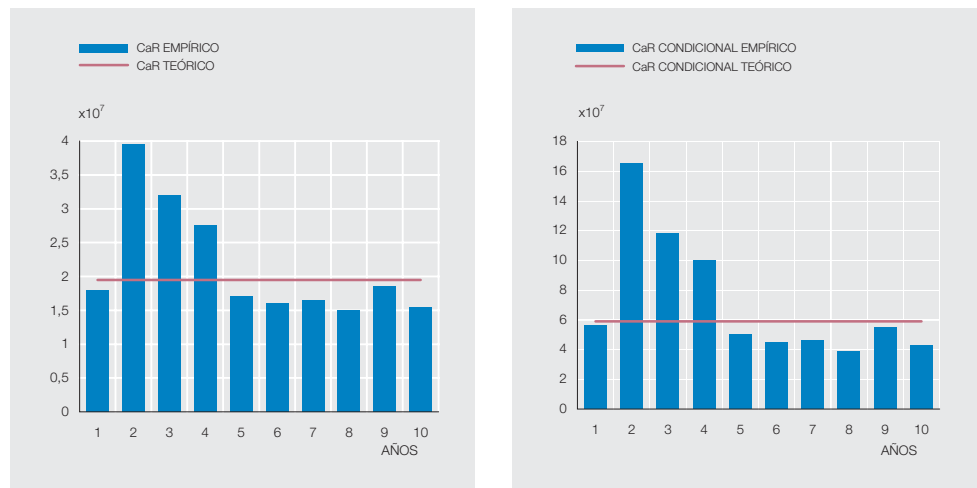
medio de eventos anuales (λ). En los gráficos 5.A y 5.B, se muestra una posible senda de las variaciones anuales de los parámetros determinantes (ξ y β) de la Pareto así como las fluctuaciones correspondientes del CaR y CaR condicional.

Los grupos de gráficos 6.A-6.B y 7.A-7.B recogen el mismo análisis con frecuencias anuales de $\lambda = 2.000$ y $\lambda = 20.000$, respectivamente.

Los resultados de estos experimentos permiten formular algunas conclusiones. En primer lugar, conviene destacar las grandes fluctuaciones, tanto del CaR como del CaR condicional, especialmente cuando el número de eventos es pequeño. Incluso para valores de λ de 2.000 y 20.000, aunque las variaciones en los parámetros estimados parezcan relativamente pequeñas, estas pequeñas fluctuaciones se traducen en importantes variaciones en términos de capital.



a. Evolución de los parámetros de ajustes de Pareto (ξ , β) para datos de pérdidas representando 10 años con una frecuencia anual de $\lambda = 2.000$ (experimento 6).

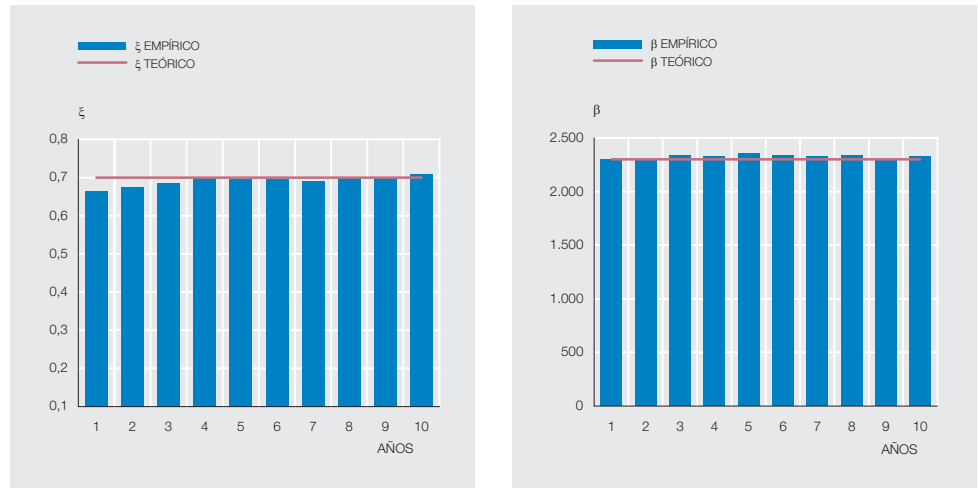


a. Evolución de las medidas de riesgo (CaR, CaR condicional), correspondientes al gráfico 6.A, para datos de pérdidas representando 10 años con una frecuencia anual de $\lambda = 2.000$ (experimento 6).

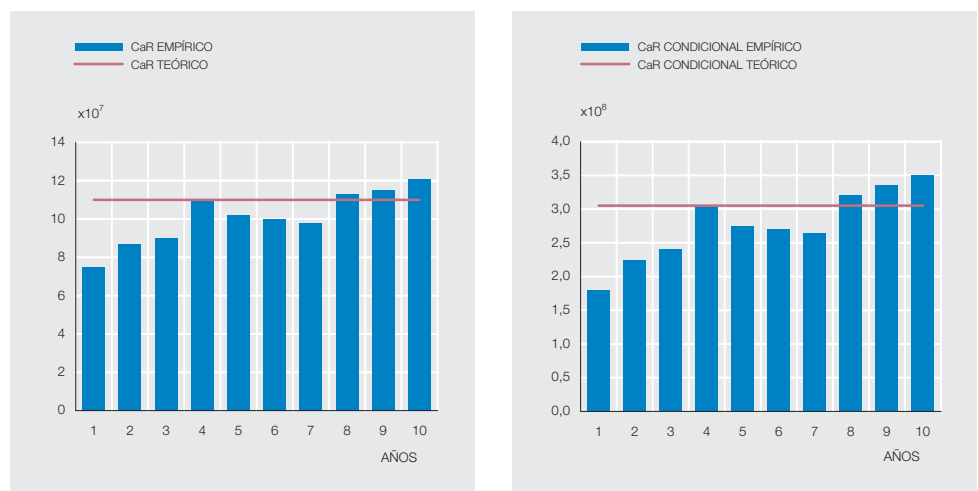
Otro hecho notable, contrario a una idea comúnmente aceptada, es que las estimaciones del capital tan pronto sobreestiman el valor teórico objetivo como lo subestiman. Incluso con valores grandes de λ , se pueden estar infravalorando las necesidades de capital. Este hecho es aún más llamativo para valores menores de λ (véase gráfico 5.B).

Los modelos de Pareto plantean un problema de implantación más complejo que otros, debido a la escasa robustez de la estimación de sus parámetros, cosa que no pasa con otros modelos, como el lognormal (véanse gráficos 8.A y 8.B).

Para ilustrar la situación con la que podrían encontrarse quienes desde una hipotética entidad fuesen registrando unos treinta sucesos superiores a los 10.000 euros al trimestre (hipótesis pesimista en nuestro entorno), distribuidos a partir de una Pareto ($\xi = 0,6$), se ha realizado una



a. Evolución de los parámetros de ajustes de Pareto (ξ , β) para datos de pérdidas representando 10 años con una frecuencia anual de $\lambda = 20.000$ (experimento 6).



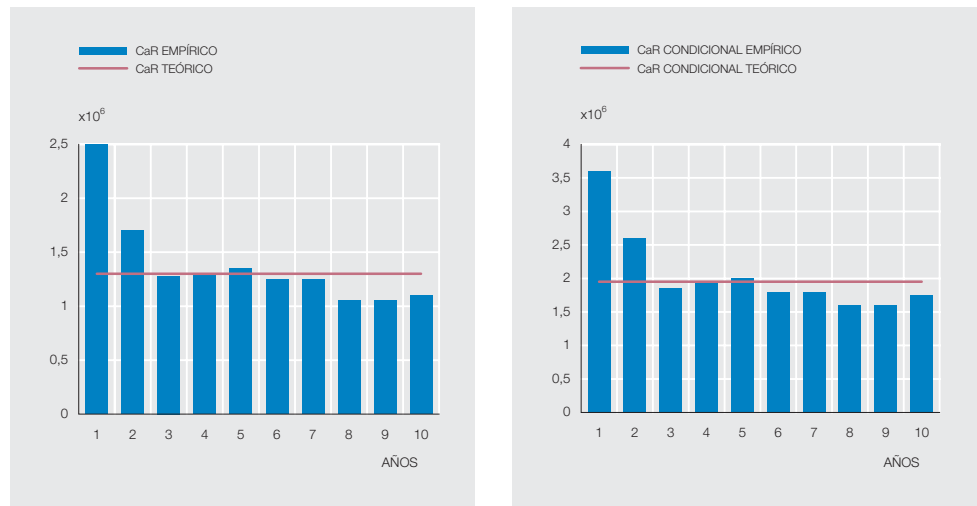
a. Evolución de las medidas de riesgo (CaR, CaR condicional), correspondientes al gráfico 7.A, para datos de pérdidas representando 10 años con una frecuencia anual de $\lambda = 20.000$ (experimento 6).

variante del experimento 6 con el fin de representar la posible dinámica trimestral del capital económico calculado (véase gráfico 9).

Este gráfico recoge una posible senda. En el eje horizontal se representan los trimestres transcurridos. Se observarán las enormes fluctuaciones del capital calculado incluso después de cuatro o cinco años. Sin embargo, ante la evidencia empírica de datos ajustados a tales distribuciones¹¹, conviene tomar medidas que ayuden a estabilizar el cómputo del capital económico y a darle sentido.

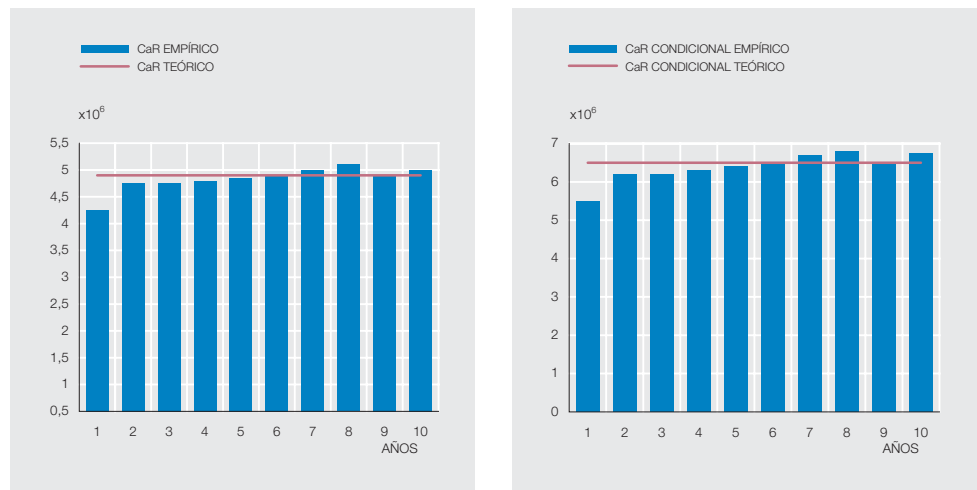
11. Véase Dekoker et al. (2005).

LOGNORMAL (I) (a)



a. Evolución de las medidas de riesgo para datos de pérdidas generados a lo largo de 10 años con una frecuencia anual de $\lambda = 200$ a partir de una distribución de severidad lognormal ($\mu = 5$, $\sigma = 2$).

LOGNORMAL (II) (a)



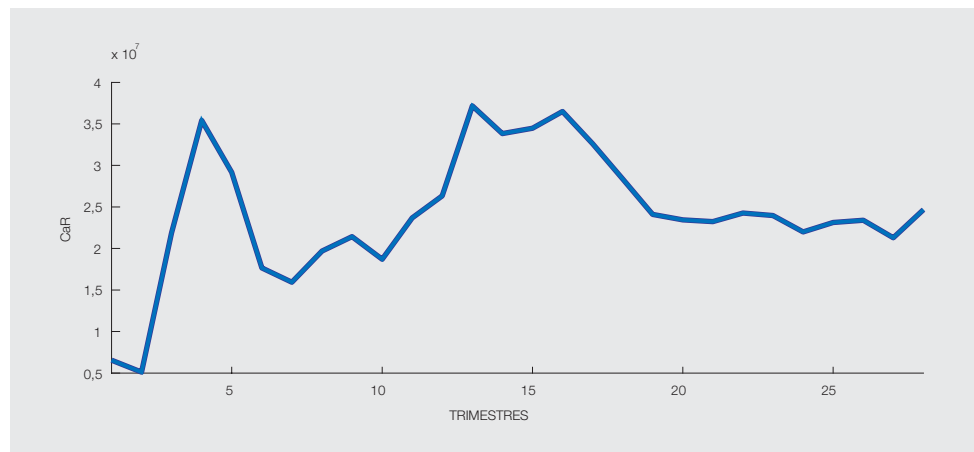
a. Evolución de las medidas de riesgo para datos de pérdidas generados a lo largo de 10 años con una frecuencia anual de $\lambda = 2.000$ a partir de una distribución de severidad lognormal ($\mu = 5$, $\sigma = 2$).

Puede contribuir a este fin el usar períodos largos (un ciclo económico) para la estimación de los mismos. Otras aproximaciones al problema pueden pasar por usar procedimientos tipo *bootstrapping* o estimaciones bayesianas del parámetro de cola para las distribuciones de Pareto.

5.3 SENSIBILIDAD DE LOS AJUSTES A SUCESOS EXTREMOS

Hemos señalado la sensibilidad de los ajustes Pareto a nuevos datos y las posibles fluctuaciones resultantes en la estimación de capital. En realidad, para este tipo de distribuciones, el parámetro de cola, y por consiguiente las medidas de riesgo, son muy sensibles a la presencia o ausencia de un único dato extremo de pérdidas. Los ajustes lognormales no presentan un comportamiento tan inestable cuando se incluye o se elimina un dato extremo.

PARETO (a)



a. Una posible senda de capital económico ajustando una distribución de Pareto a datos Pareto ($\xi = 0,6$) cuando se dispone de 30 pérdidas superiores a 10.000 euros al cuatrimestre.

Para ilustrar esta observación, se ha diseñado el *experimento 7* con el siguiente protocolo: Se genera una muestra de T valores de pérdidas con una distribución lognormal para el cuerpo y Pareto para la cola con los mismos parámetros que en el experimento 6.

Hemos considerado dos casos: el de la serie así generada (caso 2) y el de la serie obtenida eliminando de la primera el valor máximo simulado (caso 1). En ambos casos hemos supuesto que la distribución de frecuencias sigue una distribución de tipo Poisson.

Asimismo, en los dos casos se ha realizado un ajuste por máxima verosimilitud a una distribución de Pareto para la cola y se han calculado los valores de CaR y CaR condicional, utilizando la distribución empírica para describir el cuerpo de la distribución.

En el cuadro 10 se recogen los resultados de este experimento para una muestra de datos de pérdidas de tamaño $T = 1.000$, suponiendo que las frecuencias siguen una distribución de Poisson con $\lambda = 200$.

Los valores tabulados corresponden a la mediana y al rango intercuartílico para 100 simulaciones. Como se puede constatar, el impacto en términos de capital supone una fluctuación del 100%, lo que lleva a una estimación del capital equivalente al tercio del valor teórico (real).

Incluso con 10.000 datos ($\lambda \approx 2.000$), la supresión del dato mayor de la muestra lleva a una fluctuación media próxima al 20%. A título de comparación, en el caso de la lognormal ($\mu = 5$, $\sigma = 2$), estas fluctuaciones serían del 5% ($\lambda \approx 200$) y de menos del 1% ($\lambda \approx 2.000$).

En resumen, en el caso de que los datos de pérdidas por riesgo operacional muestren un comportamiento Pareto en la cola de la distribución con valores elevados del parámetro de forma ($\xi \geq 0,6$), las estimaciones empíricas de las medidas de riesgo son muy variables. En concreto, su valor fluctúa dependiendo de la presencia o ausencia de valores extremos. Esta desmesurada sensibilidad a la presencia o ausencia de dichos valores extremos en las bases de datos significa que el cálculo de un percentil elevado de la distribución de pérdidas agregada debería ser complementado con otros análisis cuantitativos con el fin de obtener medidas de riesgo estables y con sentido económico.

	u	β	ξ	$CaR_x \times 10^{-3}$	$CCaR_x \times 10^{-3}$
Teórico	1.930	2.300	0,7	3.604	11.401
Caso 1	1.900 [1.876, 1.914]	2.352 [2.086, 2.726]	0,55 [0,45, 0,70]	1.144 [661, 3.167]	2.282 [1.030, 9.817]
Caso 2	1.928 [1.913, 1.934]	2.303 [2.022, 2.619]	0,66 [0,55, 0,81]	2.492 [1.261, 7.695]	7.097 [2.627, 35.071]
Variación	32 [19, 50]	-65 [-107, -39]	0,1 [0,08, 0,13]	1.335 [538, 4.537]	4.678 [1.354, 24.933]
Variación (%)	1,67 [0,97, 2,69]	-2,78 [-4,48, -1,72]	17,92 [13,03, 26,96]	104,01 [70,36, 145,72]	179,78 [110,35, 283,22]

a. Variación de los parámetros de ajuste y de las medidas de riesgo para una muestra de tamaño $T = 1.000$ valores simulados a partir de una distribución de severidad con cuerpo lognormal ($\mu = 5$, $\sigma = 2$) y cola Pareto ($\rho = 0,1$; $u_0 = 1.930$; $\beta_0 = 2.300$; $\xi_0 = 0,7$). Se ha supuesto que la frecuencia sigue una Poisson con $\lambda = 200$ (experimento 7).

6 Aproximaciones analíticas

En un entorno actuarial, algoritmos como el de Panjer¹² o el uso de la transformada de Fourier (*Fast Fourier Transform*) han permitido un tratamiento eficiente con los percentiles al uso de los modelos LDA. El principal problema, desde un punto de vista computacional, con las condiciones de Basilea II, radica en el cálculo de un percentil muy elevado, algo intensivo en recursos de cálculo. Algunos de los enfoques analizados en este trabajo (el uso de Paretos en un modelo POT, por ejemplo) tienen entre sus objetivos el simplificar estos procedimientos. La búsqueda de aproximaciones analíticas válidas en ciertos regímenes es una línea de investigación activa y de gran interés.

6.1 LA FÓRMULA BÖCKER-KLÜPPELBERG

Recientemente, Böcker y Klüppelberg (2006) han propuesto una aproximación analítica basada en la teoría de valores extremos para el cálculo del capital. De acuerdo con la recomendación de Basilea de recoger las propiedades de colas pesadas de las distribuciones de severidad, estos autores consideran el caso de las distribuciones subexponenciales: una familia muy amplia que contiene las lognormales y todas las distribuciones con colas más pesadas; en particular las de Pareto.

Con la notación de la sección 2, se supondrá que X es una variable aleatoria subexponencial, cuya función de distribución es F_X , y que la distribución de frecuencias es una distribución de Poisson de parámetro λ . Si F_S es la función de distribución de la función agregada de pérdidas, el CaR correspondiente al nivel π satisface:

$$CaR(S, \pi) = F_S^{(-1)}(\pi) \approx F_X^{(-1)}\left(1 - \frac{1 - \pi}{\lambda}\right)$$

Se trata de una fórmula fácil de entender y de aplicar y que se puede generalizar a otras funciones de distribución de frecuencias, sustituyendo λ por $E[N]$ (el primer momento de N).

6.2 APROXIMACIÓN BASADA EN LA FÓRMULA DE RÉNYI

Consideremos el caso de la variable aleatoria de pérdidas agregadas:

$$S = \sum_{i=1}^N X_i$$

donde μ es la media de X y N una variable aleatoria con valores enteros, geométrica de parámetro p :

$$P(N = n) = p(1 - p)^{n-1}$$

12. Véase, por ejemplo, Klugman et al. (2005).

p	r	E[N]	EVT	Rényi	MC (\pm error _{MC})
0,5	10	10	252	50	255 (\pm 37)
	100	100	751	293	902 (\pm 109)
	1.000	1.000	1.996	2.414	3.144 (\pm 205)
	10.000	10.000	4.870	22.617	16.017 (\pm 607)
0,05	0,5263	10	252	121	292 (\pm 42)
	5,2632	100	751	334	888 (\pm 108)
	52,6316	1.000	1.996	1.709	3.218 (\pm 301)
	526,3158	10.000	4.870	13.161	16.284 (\pm 396)
0,005	0,0503	10	252	601	668 (\pm 60)
	0,5025	100	751	1.189	1.417 (\pm 128)
	5,0251	1.000	1.996	3.254	3.871 (\pm 191)
	50,2513	10.000	4.870	16.456	17.890 (\pm 375)

a. Comparación de las aproximaciones Böcker-Klüppelberg (EVT) y Rényi del CaR en el caso de valores de pérdidas simulados a partir de una distribución de severidad lognormal ($\mu = 5$, $\sigma = 2$). La sexta columna recoge el valor por simulación de Montecarlo. Los valores de CaR están multiplicados por un factor de 10^{-3} .

p	r	E[N]	EVT	Rényi	MC (\pm error _{MC})
0,5	1	1	72	15	73 (\pm 4)
0,05	1	19	347	152	396 (\pm 16)
0,005	1	199	1.017	1.515	1.835 (\pm 50)
0,0005	1	1.999	2.632	15.151	15.355 (\pm 225)

a. Comparación de las aproximaciones Böcker-Klüppelberg y Rényi en el caso de valores de pérdidas simulados a partir de una distribución de severidad lognormal ($\mu = 5$, $\sigma = 2$) para distintos valores de p manteniéndose r constante. Los valores de CaR están multiplicados por un factor de 10^{-3} .

El teorema de Rényi establece que

$$P(S_N \leq \frac{X}{p}) \rightarrow 1 - e^{-x/\mu}$$

cuando $p \rightarrow 0$. Este teorema es uno de los fundamentos de la teoría matemática de la fiabilidad. Teniendo en cuenta que una variable aleatoria binomial negativa (r, p) se puede expresar como una suma de r variables aleatorias geométricas de parámetro p , el resultado anterior puede extenderse al caso en el que la variable N sigue una distribución binomial negativa:

$$P(S_N \leq \frac{X}{p}) \rightarrow 1 - \Gamma_{r,1/\mu}(x) \Rightarrow \text{CaR}(S, \pi) = \frac{1}{p} \Gamma_{r,1/\mu}^{(-1)}(\pi)$$

En ambos casos (Böcker-Klüppelberg y Rényi) se obtiene una fórmula asintótica para estimar el capital económico. Los cuadros 11.A, 11.B y 11.C permiten comparar las estimaciones analíticas así obtenidas con las que da la simulación de Montecarlo.

p	r	E[N]	EVT	Rényi	MC (\pm error _{MC})
p = 0,5	10	10	6	11	10 (\pm 0)
	100	100	11	65	42 (\pm 1)
	1.000	1.000	17	539	293 (\pm 2)
	10.000	10.000	27	5.046	2.595 (\pm 4)
p = 0,05	0,5263	10	6	27	27 (\pm 2)
	5,2632	100	11	75	74 (\pm 2)
	52,6316	1.000	17	381	370 (\pm 5)
	526,3158	10.000	27	2.937	2.812 (\pm 12)
p = 0,005	0,0503	10	6	134	135 (\pm 10)
	0,5025	100	11	265	264 (\pm 18)
	5,0251	1.000	17	726	725 (\pm 18)
	50,2513	10.000	27	3.672	3.659 (\pm 35)

a. Comparación de las aproximaciones Böcker-Klüppelberg y Rényi en el caso de valores de pérdidas simulados a partir de una distribución de severidad lognormal ($\mu = 5$, $\sigma = 1$). Los valores de CaR están multiplicados por un factor de 10^{-3} .

Como se puede apreciar, la fórmula Böcker-Klüppelberg funciona mejor cuando hay un número pequeño de sucesos y el percentil de la distribución de pérdidas agregada está realmente descrito por la suma de un pequeño número de pérdidas. Es decir, cuando es adecuado aplicar la teoría de valores extremos para la función de severidad.

Por el contrario, cuando hay un mayor número de pérdidas registradas y se pueden alcanzar pérdidas agregadas grandes sumando muchas pérdidas pequeñas, la fórmula de Rényi parece dar un mejor resultado.

El problema consiste en que las situaciones del mundo real combinan ambas casuísticas y nos plantean la necesidad de seguir trabajando en estos temas.

7 Conclusiones

Varias son las conclusiones que se pueden extraer de los experimentos llevados a cabo. Por una parte, conviene destacar que estamos al principio del proceso que ha de llevar a una medición efectiva del riesgo operacional, a la dotación de capital correspondiente y, sobre todo, a su gestión. Por ello, muchas de las propuestas realizadas deben ser contrastadas con datos empíricos.

Uno de los riesgos más importantes en este proceso es el riesgo de modelo, especialmente en esta primera fase, cuando la escasez de datos puede hacer difícil la obtención de las evidencias estadísticas necesarias para distinguir entre modelos que proporcionan estimaciones muy distintas de las medidas de riesgo. Una de las dificultades mayores en la cuantificación del riesgo operacional es la necesidad de calcular un percentil muy elevado (el 99,9%) de la distribución agregada de pérdidas a partir del ajuste de las distribuciones de severidad y frecuencia. El problema es que puede haber distintas distribuciones que proporcionan un ajuste satisfactorio para los datos de severidad observados pero que tienen comportamientos asintóticos (que luego llevan a cálculos de capital) muy diferentes. Con los datos disponibles y mientras no se disponga de muchos más datos de las colas de las distribuciones es difícil la extrapolación al percentil 99,9%. En esa línea, la posibilidad de disponer de datos externos, debidamente escalados, puede suponer una ayuda muy importante en la medición eficiente del riesgo operacional.

A pesar de su aceptación por parte de algunos analistas de riesgo y de algunas entidades¹³, un simple ajuste de distribuciones de Pareto a eventos superiores a un umbral determinado no parece ser una solución viable. Por una parte, los ajustes a este tipo de distribuciones son muy poco robustos. En concreto, presentan una enorme variabilidad, tanto en la estimación de parámetros como en el cálculo de las medidas de riesgo (capital económico y CaR condicional). Dichas estimaciones son muy sensibles a la incorporación de nuevos datos (incluso aunque la distribución real subyacente sea de tipo Pareto), por lo que existen muchas dudas sobre la relevancia económica de los resultados que se obtienen. Adicionalmente, con pocos datos no se tiene la seguridad de garantizar la cobertura de las necesidades mínimas de capital: el ajuste Pareto puede llevar a una infravaloración del capital económico en algún caso.

Por otra parte, uno de los principales defectos de una aplicación directa de la teoría de valores extremos a la cuantificación del riesgo operacional radica en la posibilidad de que las severidades de las pérdidas por riesgo operacional no correspondan a un régimen asintótico en el que la aproximación de Pareto sea aceptable. Adicionalmente, estos modelos desprecian las contribuciones de las pérdidas que se encuentran por debajo del umbral de cola. Aun siendo de menor cuantía, dichas pérdidas podrían tener una contribución importante cuando son agregadas, especialmente para frecuencias elevadas. Finalmente, las aproximaciones analíticas disponibles son insuficientes. En particular, no está estimado el término de error de las mismas.

Una de las explicaciones de la gran variabilidad e inestabilidad de los resultados obtenidos es la falta de robustez de los ajustes basados en la teoría de valores extremos. A menudo, al modelizar con distribuciones de Pareto las colas de los datos de pérdidas (incluso en el caso de distribuciones subexponenciales, pero con colas menos pesadas, como pueda ser una lognormal), se obtienen valores del parámetro de forma (*shape parameter*) superiores a 0,5, lo que implica que el momento de orden 2 es infinito. Trabajando con datos reales, se pueden obtener valores incluso superiores a 1, en cuyo caso, la propia media (es decir, la pérdida esperada) es infinita.

Estas observaciones no implican que la teoría de valores extremos no sea útil en la cuantificación del riesgo operacional. Es posible que una aplicación más cuidadosa de dicha teoría sea una herramienta adecuada para la estimación de medidas de riesgo operacional. En concreto, puede que tenga sentido realizar un ajuste de tipo Pareto a la distribución de pérdidas agregadas, después de haber realizado una simulación Montecarlo, con el fin de obtener una mejor descripción de la cola y una estimación más robusta de las medidas de riesgo.

De los experimentos realizados con datos simulados cabe inferir que es conveniente adoptar una actitud muy cauta ante el uso de las distribuciones de Pareto. Las medidas de riesgo obtenidas deben tener, ante todo, sentido económico. Para parámetros de forma elevados, próximos a 1, se obtienen cifras de capital regulatorio que pueden superar con mucho el propio capital de la entidad. Dado que, con datos reales, se obtienen a menudo parámetros de forma elevados, es necesario preguntarse si tiene sentido modelizar la severidad con este tipo de distribuciones o si, por el contrario, no se debería optar por versiones truncadas de las mismas. Esta opción, aún no explorada en la literatura, podría resolver muchos de los problemas señalados.

A nuestro juicio, es posible llevar a cabo un cálculo efectivo del capital económico, pero es necesario evitar hacer simplificaciones excesivas. La evidencia obtenida a partir de datos

13. Véase, por ejemplo, Dekoker et al. (2005).

reales apunta a que no solo los eventos extremos son importantes para la cuantificación del riesgo operacional, y que es necesario incluir en el cálculo el cuerpo de la distribución, al menos en celdas con frecuencias elevadas. Puede que sea necesario utilizar otras familias de distribuciones (por ejemplo, distribuciones de Pareto truncadas) para tener en cuenta tanto la leptocurtosis de la distribución de severidad de pérdidas como el hecho de que las pérdidas de una entidad no pueden ser infinitas.

BIBLIOGRAFÍA

- ARTZNER, P., F. DELBAEN y J.-M. EBER (1999). «Coherent measures of risk», *Mathematical Finance*, 9 (3), pp. 203-228.
- BÖCKER, K., y C. KLÜPPELBERG (2006). «Operational VaR: a Closed-Form Approximation», *Asia Risk Magazine*.
- CHAVEZ-DEMOULIN, V., P. EMBRECHTS y J. NEŠLEHOVÁ (2006). «Quantitative models for operational risk: Extremes, dependence and aggregation», *Journal of Banking and Finance*, 30, pp. 2635-2658.
- CUMMINS, J. D., C. M. LEWIS y R. WEI (2006). «The market impact of operational loss events for US banks and insurers», *Journal of Banking and Finance*, 30, pp. 2605-2634.
- DE FONTNOUVELLE, P., V. DE JESUS-RUEFF, E. S. ROSENGREN y J. S. JORDAN (2003). *Using loss data to quantify Operational Risk*, Working Paper, Federal Reserve Bank of Boston.
[<http://www.algorithmics.com/solutions/opvantage/docs/UsingLossData.pdf>]
- DE FONTNOUVELLE, P., E. S. ROSENGREN y J. S. JORDAN (2004). *Implications of Alternative Operational Risk Modeling Techniques*, Working Paper for NBER Project, SSRN.
[http://papers.ssrn.com/sol3/papers.cfm?abstract_id=556823]
- DEKOKER, R., y J. NEWBERRY (2005). *AMA Implementation at Citigroup, Where We Are and Outstanding Questions*, Working Paper, Federal Reserve Bank of Boston
[http://www.bos.frb.org/bankinfo/conevent/oprisk2005/dekoker_newberry.pdf]
- DUTTA, K., y J. PERRY (2006). *A Tale of Tails: An Empirical Analysis of Loss Distribution Models for Estimating Operational Risk Capital*, Working Paper, Federal Reserve Bank of Boston.
- EMBRECHTS, P., H. FURRER y R. KAUFMANN (2003). «Quantifying regulatory capital for operational risk, Derivatives Use», *Trading & Regulation*, 9(3), pp. 217-233.
- EMBRECHTS, P., C. KLÜPPELBERG y T. MIKOSCH (1997). *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag.
- FERRERAS, A. (2005). *Riesgo operacional: algunas consideraciones críticas relativas al uso de modelos avanzados*, V Jornada de Riesgos Financieros RiskLab.
- FRACHOT, A., O. MOUDOULAUD y T. RONCALLI (2003). *Loss Distribution Approach in Practice*, en M. Ong (ed.), *The Basel Handbook: A Guide for Financial Practitioners*, Risk Books.
- GUMBEL, E. J. (1935). *Les valeurs extrêmes des distributions statistiques*, Ann. L'Inst. Henri Poincaré 4, pp.115-158.
- KALASHNIKOV, V. (1997). *Geometric Sums: Bounds for Rare Events with Applications*, Kluwer Academic Publishers.
- KLUGMAN, S. A., H. H. PANJER y G. E. WILLMOT (2005). *Loss Models, From Data to Decisión*, 2.ª edición, John Wiley & Sons.
- KOTZ, S., y S. NADARAJAH (2000). *Extreme Value Distributions, Theory and Applications*, Imperial College Press.
- MEDOVA, E. A., y M. N. KYRIACOU (2001). *Extremes in operational risk management*, Working Paper, Centre for Financial Research, Judge Institute of Management, University of Cambridge.
- MIGNOLA, G., y R. UGOCCIONI (2005). *Tests of Extreme-Value Theory Applied to Operacional Risk Data*, en E. Davis (ed.), *Operacional Risk*, Risk Books.
- MOSCADELLI, M. (2004). *The Modelling of Operational Risk: the Experience with the Analysis of the Data Collected by the Basel Committee*, Banca d'Italia, Temi di Discussione del Servizio Studi (517).
- NA, H. S. (2004). *Analysing and scaling operational risk loss data*.
- NA, H. S., L. COUTO MIRANDA, J. VAN DEN BERG y M. LEIPOLDT (2005). *Data scaling for operacional risk modelliing*, ERIM Report Series Research in Management.
- NEŠLEHOVÁ, J., P. EMBRECHTS y V. CHAVEZ-DEMOULIN (2006). «Infinite-mean models and the LDA for operational risk», *Journal of Operational Risk*, 1.
- NIETO GIMÉNEZ-MONTESINOS, M. A. (2005) *Estabilidad Financiera*, 8, pp. 163-185.
- RÉNYI, A. (1956). *Characterization of Poisson Processes*, Magyar Tud. Akad. Mat. Kutatu Int. Kozl., 1, pp. 519-527 (en húngaro).

APÉNDICE

A. La distribución de Pareto generalizada

Sea X una variable aleatoria, que toma valores mayores que $u > 0$. Esta variable aleatoria está distribuida según una Pareto generalizada de parámetros u , $\beta \geq 0$, $\xi > 0$ si su densidad de probabilidad viene dada por:

$$g_{u,\beta,\xi}(x) = \frac{1}{\beta} \left(1 + \frac{\xi}{\beta} (x - u) \right)^{-1-1/\xi} \times \theta_{\{x-u\}}$$

siendo $\theta(z)$ la función que vale 1 para $z > 0$ y 0 para $z \leq 0$. Su función de distribución viene entonces dada por:

$$G_{u,\beta,\xi}(x) = P(X \leq x) = \left[1 - \left(1 + \frac{\xi}{\beta}(x-u) \right)^{-1/\xi} \right] \times \theta_{\{x-u\}}$$

y su inversa, determinante para el cálculo de los percentiles, es:

$$G_{u,\beta,\xi}^{-1}(\pi) = u + \frac{\beta}{\xi} \left((1-\pi)^{-\xi/(1+\xi)} - 1 \right)$$

Los ajustes de este tipo de distribuciones han de reajustarse en función de los umbrales utilizados. Supongamos que el ajuste a una tal distribución se hace con datos obtenidos a partir de un umbral $u' > u$. Sea $G_{u',\beta',\xi'}$ la función de distribución correspondiente. Es natural exigir para ello que la probabilidad de la cola por encima de u' sea independiente de la Pareto elegida:

$$1 - G_{u',\beta',\xi'}(x) = \frac{1 - G_{u,\beta,\xi}(x)}{1 - G_{u,\beta,\xi}(u')}, \quad \forall x > u'$$

lo cual solo se cumple si:

$$\xi' = \xi, \text{ y } \beta' = \beta - \xi(u' - u)$$

B. La distribución empírica

Consideremos una muestra aleatoria $X = (X_1, \dots, X_n)$; definiremos la distribución empírica asociada a X por su función de masa:

$$f_{\text{EMP}}(x; X) = \frac{1}{n} \sum_{i=1}^n \delta(x - X_i)$$

siendo δ la función que toma el valor 1 en 0 y 0 fuera.

Las funciones de distribución y distribución inversa vienen dadas por:

$$F_{\text{EMP}}(x; X) = \frac{1}{n} \sum_{i=1}^n \theta(x - X_i), \quad F_{\text{EMP}}^{-1}(p; X) = \text{prctile}(X, p)$$

siendo $\text{prctile}(X, p)$ la función percentil empírico asociada a la muestra X .

C. La distribución Pareto/empírica

Como hemos señalado, esta distribución se obtiene usando la empírica para el cuerpo de la distribución (hasta el umbral u que deja una masa de probabilidad igual a p a su derecha). Su densidad de probabilidad viene dada por:

$$f_{\text{EG}; u,\beta,\xi,p}(x; X) = \frac{1-p}{\sum_{i=1}^n \theta(u - X_i)} \sum_{i=1}^n \delta(x - X_i) \theta(u - X_i) + p \times g_{u,\beta,\xi}(x) \theta(x - u)$$

La función de distribución correspondiente y su inversa vienen dadas por

$$F_{\text{EG}; u,\beta,\xi,p}(x; X) = \frac{1-p}{\sum_{i=1}^n \theta(u - X_i)} \sum_{i=1}^n \theta(x - X_i) \theta(u - X_i) + [(1-p) + p \times G_{u,\beta,\xi}(x)] \theta(x - u)$$

$$F_{\text{EG}; u,\beta,\xi,p}^{(-1)}(\pi) = \theta(p - \pi) \text{prctile}(X(X < u), \pi(1-p)) + \theta(\pi - p) G_{u,\beta,\xi}^{-1} \left(\frac{\pi - p}{1-p} \right)$$

