

## EVALUACIÓN DE LAS METODOLOGÍAS PARA MEDIR EL VALOR EN RIESGO

Clara I. González (\*)

Ricardo Gimeno (\*)

(\*) Clara I. González y Ricardo Gimeno pertenecen a la Dirección General del Servicio de Estudios del Banco de España. Los autores agradecen los comentarios recibidos de Vicente Salas.



## Evaluación de las metodologías para medir el valor en riesgo

El valor en riesgo (VaR) constituye una medida habitual en la medición del riesgo de mercado entre las entidades financieras, debido a su utilidad y fácil interpretación. Sin embargo, no existe una metodología comúnmente aceptada para su cálculo.

En el presente artículo se comparan los resultados obtenidos en la medición del VaR de la distribución estadística de los rendimientos de un conjunto de carteras de mercado, utilizando las metodologías más habituales en el cálculo del VaR (simulación histórica, GARCH, valores extremos). Los criterios utilizados para valorar las cuatro metodologías son la proporción de rendimientos que superan el valor del VaR, la cifra de VaR medio y la cuantía de todas las realizaciones de rendimientos superiores al VaR. En la aplicación del método de los valores extremos se propone un procedimiento automático y objetivo para separar los valores extremos y se supone que las colas de la distribución de probabilidad se distribuyen de acuerdo con una distribución de Pareto.

### 1 Introducción

El VaR se ha convertido en la medida estándar que los analistas financieros usan para cuantificar el «riesgo de mercado», es decir, la posibilidad de pérdidas por disminución en el valor de mercado de una determinada cartera de valores. El VaR se define como la máxima pérdida posible de una cartera debida a movimientos adversos de su valor, para una probabilidad dada de que ocurra en un período de tiempo determinado. La gran popularidad que ha adquirido este instrumento se debe, esencialmente, a su simplicidad conceptual, ya que el VaR reduce el riesgo asociado a una cartera a un solo número, indicando la pérdida esperada con un nivel de confianza determinado.

Las metodologías para el cálculo del VaR tienen múltiples aplicaciones, y son usadas tanto para la gestión interna de los riesgos a los que una entidad se ve expuesta como con objetivos legislativos y reguladores. Así, por ejemplo, el Banco de Pagos Internacionales de Basilea (BIS) establece la utilización del VaR para la incorporación del riesgo de mercado a los requerimientos de capital de las instituciones financieras.

Desde que el Comité de Basilea anunció en 1995 que el establecimiento de las reservas de capital de las instituciones financieras tenía que basarse en las metodologías de VaR hasta la actualidad, han ido surgiendo diversos estudios y análisis de la amplia variedad de metodologías que cabe aplicar.

Sin embargo, como han señalado muchos estudios, desde Beder (1995) hasta Kuester et al. (2006), los resultados de las estimaciones del VaR cambian de forma muy significativa según la metodología utilizada. Por este motivo se hace necesaria la evaluación de la capacidad de cobertura del riesgo de cada una de ellas.

Las diferentes metodologías para el cálculo del VaR se pueden clasificar en métodos paramétricos, no paramétricos y semiparamétricos<sup>1</sup>. En este artículo se procede a analizar y comparar cuatro de las metodologías más populares: dos paramétricas (normal y GARCH), una no paramétrica (simulación histórica) y una semiparamétrica (valores extremos). Se estima diariamente el VaR para 36 índices de mercados bursátiles de todo el mundo, en el período 1999-2003, lo que permite aproximar carteras con distintos perfiles de riesgo, y evaluar el grado de

1. Véase Engle y Manganelli (2004).

acierto en el acercamiento al VaR real de cada una de las cuatro medidas. Las metodologías aplicadas en este artículo son las mismas que propone analizar Longin (2000), si bien este último se limita a comparar las estimaciones resultantes de la simulación histórica frente a las otras tres, pero no llega a comprobar la efectividad de cualquiera de las cuatro en cuanto al objetivo final del VaR, que es la cobertura de riesgos.

A partir de estas estimaciones, procedemos a comparar las distintas metodologías en función de tres criterios: la proporción de rendimientos que no superan la cifra establecida por el VaR (lo deseable sería que se acercara al nivel de confianza establecido, sin ser nunca inferior a este); la cifra de VaR medio, que indicaría la provisión media que se tendría que utilizar para cubrirse, de manera que, cuanto menor sea la provisión media, menos recursos quedan asignados a cubrir la posible pérdida; y la cantidad de pérdidas no anticipadas, es decir, los rendimientos por encima del VaR establecido (lo deseable sería que el resultado fuera una cifra muy reducida). El método preferido será el que se acerque más a los valores objetivo establecidos para cada criterio, pues ello significará que se está capturando mejor el riesgo que realmente afecta a la cartera estudiada.

A pesar de la amplia literatura de trabajos que evalúan las distintas metodologías de VaR, resulta difícil encontrar trabajos completos, pues en su mayor parte o utilizan un número reducido de criterios para la comparación o dejan fuera alguna de las metodologías habituales. Así, por ejemplo, Hendricks (1996) utiliza hasta nueve criterios de comparación, entre los que se incluyen los tres que se utilizan en este artículo, con series temporales de tipos de cambio, si bien entre sus aproximaciones no incluye ni las metodologías basadas en los modelos GARCH ni los basados en valores extremos (EV). Christoffersen et al. (2001) sí incorporan los modelos GARCH, aunque no EV, pero la evaluación de las metodologías con datos del S&P 500 se reduce solo al ajuste respecto a la proporción de rendimientos que no superan la cifra establecida por el VaR, y tampoco establece preferencias con respecto a quedar por debajo mejor que por arriba. Por el contrario, Bali (2003) sí incluye EV, pero no modelos GARCH, mientras que la evaluación la hace con tipos de interés de deuda pública norteamericana, comparando la proporción de aciertos del VaR y la pérdida esperada en caso de que los rendimientos superen el VaR, y no estudiando los requerimientos de cobertura consecuencia de ese nivel de VaR.

Entre los que sí utilizan EV, como son Engle y Manganelli (2004) o Bali (2003), se inclinan por la aproximación paramétrica a la estimación de la cola de la distribución, pues esta puede ser automatizada, mientras que el enfoque semiparamétrico requiere la intervención del analista con cada nueva estimación de los parámetros [Danielsson y de Vries (1997)], lo que hace impracticable su uso en un proceso de evaluación de metodologías de VaR, a pesar de que, tal y como señalan Jansen y de Vries (1991), el método semiparamétrico logra estimadores más eficientes que los paramétricos estimados por máxima verosimilitud. Este es el motivo por el que en el presente artículo se propone un método automático que permite el uso de una estimación semiparamétrica de EV.

El resto del artículo se estructura como sigue: en la sección 2 se analizan las distribuciones estadísticas de los rendimientos financieros; a continuación, en la sección 3 se presentan las distintas medidas de VaR que se van a utilizar en el trabajo, así como una metodología alternativa al cálculo de los valores extremos; en el apartado 4 se comparan los resultados obtenidos con cada tipo de VaR, mientras que en la sección 5 se presentan las conclusiones.

## **2 Distribución estadística de los rendimientos financieros**

El análisis del riesgo de un activo financiero requiere el estudio de la distribución estadística que siguen sus rendimientos, ya que sin ella no podemos asignar probabilidades a los pagos futuros derivados de su posesión.

ÍNDICE	TODA LA MUESTRA		SIN VALORES EXTREMOS		% OBS. ELIMINADAS
	JARQUE BERA	P-VALUE	JARQUE BERA	P-VALUE	
AEX ÁMSTERDAM	276,20	0,00	0,06	0,97	5,44
ATENAS ASE	647,31	0,00	0,85	0,66	3,91
BANGKOK	7113,85	0,00	0,22	0,89	2,20
BCN MID-50	1.722.657,70	0,00	0,04	0,98	5,86
BOVESPA	64,28	0,00	0,93	0,63	3,39
BRUSELAS BEL	582,78	0,00	1,55	0,46	9,01
CAC40 PARÍS	86,87	0,00	0,03	0,98	2,91
COPENHAGUE	79,18	0,00	0,21	0,90	2,15
DAX XETRA	1.483,62	0,00	0,47	0,79	3,30
DJ COMPOSITE	448,86	0,00	0,02	0,99	4,02
EL CAIRO	507,61	0,00	0,11	0,95	9,05
EUROTOP 100	94,05	0,00	0,07	0,96	3,57
FTSE-100	97,32	0,00	0,16	0,92	2,86
HANG SENG	7.225.670,60	0,00	0,06	0,97	2,99
HELSINKI HEX	8.790.202,60	0,00	0,76	0,68	2,92
IBEX-35	16,21	0,00	0,09	0,96	1,50
JOHANESBURGO	675,11	0,00	0,01	1,00	2,26
KUALA LUMP.	1.089,02	0,00	3,16	0,21	5,22
LIMA IG	878,56	0,00	0,06	0,97	4,69
MERVAL	1.213,11	0,00	0,21	0,90	6,32
MÉXICO IPC	186,27	0,00	0,79	0,67	3,49
MIB 30	166,34	0,00	0,11	0,95	2,30
MIBTEL	209,77	0,00	0,07	0,96	2,50
MOSCÚ RST	209,56	0,00	0,29	0,87	3,44
NASDAQ 100	143,27	0,00	0,04	0,98	1,84
NIKKEI	4.752.273,70	0,00	0,49	0,78	1,27
S&P 500	158,28	0,00	0,00	1,00	2,94
SANTIAGO IGP	405,76	0,00	1,13	0,57	3,85
SIDNEY A. OR.	2.033,68	0,00	0,07	0,97	1,74
SINGAPUR ST	645,94	0,00	0,49	0,78	2,17
SLOUGH	26.572,59	0,00	0,35	0,84	4,02
SWISS MARKET	9.251.170,90	0,00	0,04	0,98	4,74
TEL-AVIV 100	752,20	0,00	0,89	0,64	2,87
TSE-WEIGHTED	6.645.968,90	0,00	2,40	0,30	2,65
VIENA ATX	202,04	0,00	0,21	0,90	3,77
YAKARTA	1.020,40	0,00	0,22	0,90	2,83

Así, por ejemplo, podríamos asumir que los rendimientos financieros se comportan según una distribución normal con parámetros estacionarios a lo largo del tiempo. Este supuesto tiene como principal ventaja que facilita la estimación de modelos econométricos y la construcción de estadísticos, al tiempo que se puede caracterizar el comportamiento del activo a través del vector de medias y la matriz de varianzas-covarianzas. Además, si tenemos en cuenta que la suma de normales es otra normal, cualquier cartera construida con activos financieros también tendrá una distribución normal y son numerosos los modelos teóricos que asumen esta hipótesis en sus planteamientos (por ejemplo, el modelo de valoración de opciones de Black-Scholes).

Sin embargo, las observaciones empíricas de las distribuciones de los rendimientos contradicen sistemáticamente la hipótesis de normalidad. Ya en los trabajos de Mandelbrot (1963) y Fama (1965) se señalan las características típicas de estas variables, que se pueden resumir en las siguientes: en primer lugar, los movimientos bruscos suelen ser más pronunciados en el caso de las bajadas de precios que en el de las subidas (asimetría a la izquierda); los mercados tienden a ser más sensibles a las malas noticias que a las buenas; además, se comprueba que la volatilidad de los activos no es constante, sino que se suceden momentos de fuertes cambios de precios seguidos de otros más estables, lo que lleva a la existencia de más realizaciones en los extremos de la distribución de las esperadas si siguieran una distribución normal (colas anchas); y, por último, en torno al valor medio se produce concentración de probabilidad, es decir, la distribución es más apuntada y estrecha, presentando más probabilidad que la esperada en los movimientos muy pequeños. En definitiva, nos encontramos ante una distribución con mayor masa de probabilidad en los movimientos menores y en los más extremos, siendo menores los cambios de precios que podríamos denominar «intermedios».

Si se analiza la distribución de rendimientos de un valor bursátil, se puede comprobar que, efectivamente, nos encontramos con distribuciones leptocúrticas, asimétricas y con colas anchas. Así, por ejemplo, si se analizan los rendimientos de 36 índices bursátiles mundiales entre los años 1999 y 2003<sup>2</sup>, se puede comprobar cómo sus distribuciones no son normales. Los valores esperados de asimetría negativa y alta curtosis son indicios de que las características anteriormente señaladas están presentes en todos ellos.

A través del estadístico de normalidad de Jarque-Bera<sup>3</sup> se pueden contrastar las diferencias de los coeficientes de curtosis y asimetría de estos índices respecto de las cifras que cabría esperar en una variable que fuera normal. En el cuadro 1 se observa cómo en todos los casos se rechaza la hipótesis de que procedan de una distribución normal. Los índices que destacan por sus elevadas cifras de curtosis (BCN MID-50, HANG SENG, HELSINKI HEX, NIKKEI, SWISS MARKET y TSE-WEIGHTED) también son los que tienen un estadístico más elevado.

Dadas las características sistemáticamente presentes en rendimientos financieros de distinta procedencia, es evidente que el cálculo de las probabilidades a partir de sus distribuciones no es directo, lo que ha llevado a la aparición de distintas metodologías de cálculo del VaR.

### **3 Las medidas de valor en riesgo**

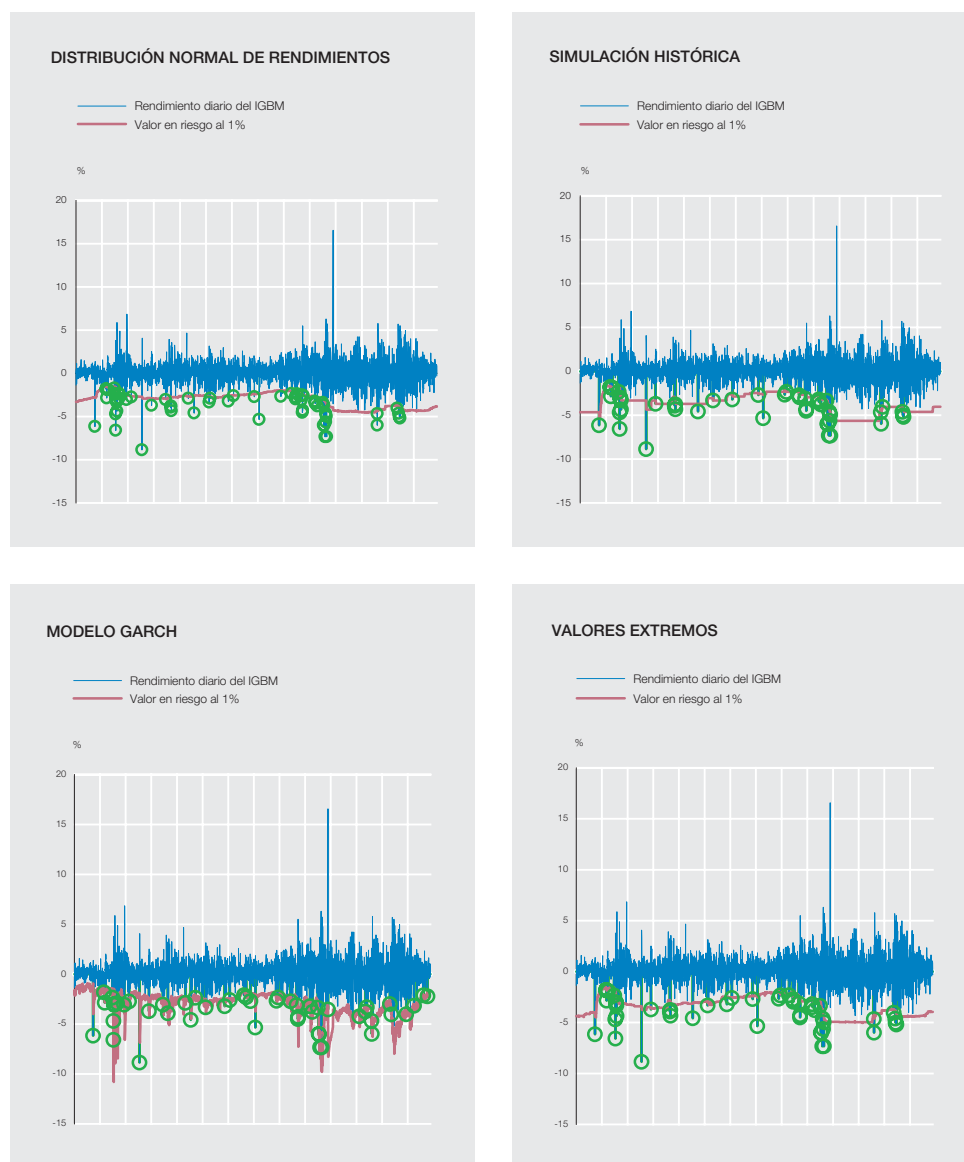
El concepto de VaR es bastante intuitivo, pero no existe una forma única generalmente aceptada para su cálculo, requiriendo además el uso de herramientas estadísticas que pueden ser conceptualmente complejas. La disparidad de métodos radica en la necesidad de encontrar un modelo adecuado para los rendimientos de una cartera. El éxito de estas metodologías dependerá de la capacidad que tengan para recoger las características típicas que habitualmente se encuentran en las series financieras históricas mencionadas en el apartado anterior. La multitud de metodologías propuestas por la literatura para calcular el valor en riesgo pueden agruparse en tres: no paramétricas, paramétricas y semiparamétricas. Como referencia para medir el grado de éxito de cada una de las aproximaciones consideradas utilizaremos también una medida del VaR que asume que los rendimientos tienen una distribución de probabilidad normal, y que, por tanto, no tiene en cuenta ninguna de las tres propiedades señaladas.

#### **3.1 MODELOS NO PARAMÉTRICOS**

Las características de los rendimientos de una cartera anteriormente señaladas pueden ser tratadas, en una primera aproximación, sin hacer ninguna asunción sobre la distribución de

---

2. Véase cuadro 2. 3. Véase Jarque y Bera (1980).



esta variable, sino utilizando la técnica que se conoce como simulación histórica. Esta metodología se basa en el uso de una ventana deslizante de datos históricos de los rendimientos observados en la cartera que se está estudiando, generalmente entre seis meses y dos años, dependiendo del nivel de significación que se quiera utilizar.

Una vez se dispone de una muestra de rendimientos históricos, estos se ordenan y se comprueba en esa ventana temporal cuál es el valor que determina el nivel de confianza deseado por el VaR, de tal manera que se repetirá sucesivamente moviendo la ventana utilizada para la estimación.

La principal ventaja de este método es que replica la distribución empírica de los rendimientos de la cartera, incluyendo la leptocurtosis y asimetría típica de ellos. No obstante, existe una serie de problemas con los que nos vamos a encontrar: el primero es que la presencia de un valor extremo negativo hoy va a provocar no solo la aparición inmediata de un incremento en el VaR, sino la predicción de un descenso en el momento en que ese valor salga de la muestra (gráfico 1); además, dado que se asume que la distribución es constante en toda la venta-

na temporal, en los momentos en los que se produzcan transiciones entre épocas de baja y de alta volatilidad estas tardarán un tiempo en reflejarse en la ventana, por lo que mientras se producirán sesgos en la estimación del VaR<sup>4</sup>.

### 3.2 MODELOS PARAMÉTRICOS

Un segundo grupo de metodologías para el cálculo del VaR son las que incorporan las dinámicas observadas en la volatilidad de los rendimientos de las carteras. Para ello requieren la proposición de algún tipo de modelo econométrico que explique la evolución de las varianzas. Ese es el caso de los llamados modelos GARCH, donde la volatilidad de la cartera en un período está en función de la observada en períodos anteriores, en lo que se podría considerar una media móvil de los cuadrados de los rendimientos (aproximaciones de las volatilidades) ponderando más los que están más próximos en el tiempo.

La necesidad de parametrización de la evolución de la varianza es el principal problema de esta metodología, pues el VaR estimado queda expuesto a errores en la especificación, tanto de la ecuación de la varianza como de la distribución que siguen los rendimientos.

Como puede observarse en el gráfico 1, la metodología GARCH es la que requiere unos cambios más frecuentes y bruscos en la cuantía del VaR, lo cual exige una gestión más activa para la entidad y una mayor vigilancia para el regulador.

### 3.3 MODELOS SEMIPARAMÉTRICOS

Las metodologías basadas en los valores extremos no intentan estimar un modelo para toda la muestra de rendimientos, sino solo para la parte relevante en el cálculo del VaR, que es la de las pérdidas severas en la cartera. La principal virtud de este tipo de modelización radica en que aúna las ventajas de los métodos no paramétricos, pues recoge las características empíricamente observadas sobre la distribución de los rendimientos de los activos, pero con la ventaja añadida de no ser tan sensible a la aparición de valores extremos esporádicos, y a su vez no plantea los problemas de especificación de los modelos paramétricos. Sus principales inconvenientes se encuentran en la ausencia de dinámicas que recojan la autocorrelación de la volatilidad y la necesidad de una definición adecuada de lo que va a ser considerado como «valor extremo».

Entre los modelos que pretenden modelizar los valores extremos que se presentan a partir de un determinado valor denominado «umbral», es decir, la pérdida que sobrepasa un determinado valor fijado, destaca el llamado *Peaks Over the Threshold (POT)*, que puede ser utilizado para estimar el exceso de la distribución respecto de un determinado valor  $a$ , para así estimar la forma de la cola de la distribución original. Dentro de los modelos POT podemos encontrar dos tipos de modelos, uno de ellos es la familia de modelos semiparamétricos construidos alrededor del estimador de Hill<sup>5</sup>, y el otro son los modelos paramétricos basados en la distribución generalizada de Pareto (GDP)<sup>6</sup>.

La distribución más utilizada para las colas es la de Pareto, a partir del estimador de Hill<sup>7</sup>, que separa los rendimientos extremos de los no extremos. El principal reto de esta metodología radica en encontrar el valor correspondiente al umbral ( $a$ ) y al índice de cola ( $\alpha$ ) que nos permita modelizar los valores extremos. Si el umbral es conocido, entonces  $\alpha$  puede ser estimado por máxima verosimilitud a través del estimador de Hill<sup>8</sup>. Sin embargo, el valor del umbral es desconocido, siendo la elección de este punto de corte lo más relevante de todo el proceso. Para encontrarlo, el método habitual consiste en calcular y dibujar el estimador de Hill para diferentes valores del umbral<sup>9</sup>, seleccionando aquel para el cual el estimador es estable. Pero

4. Véase Engle y Manganelli (2004). 5. Véanse Beirlant et al. (1996) y Danielsson et al. (1998). 6. Véase Embrechts et al. (1999). 7. Véase Hill (1975). 8. Véanse Beirlant et al. (1996) y Danielsson et al. (1998). 9. Véase Dress et al. (2000).



esta elección no siempre es evidente y supone un cierto grado de subjetividad por parte del investigador que podría llevar a errores en las estimaciones. Debido a su arbitrariedad, en este artículo se propone un método alternativo de estimación del umbral que permita su elección de forma totalmente objetiva y automatizada<sup>10</sup>.

La aproximación que presentamos se basa en la hipótesis de que los valores extremos de la distribución vienen generados por la llegada de noticias que alejan el comportamiento del mercado de lo que sería normal. Esas reacciones anómalas del mercado, mayores en las bajadas y más frecuentes en determinados momentos de tiempo, serían suficientes para explicar las características típicas de las distribuciones de rendimientos financieros.

Se comprueba cómo, al separar los rendimientos extremos de los que serían pertenecientes a un comportamiento estable del mercado, estos últimos tienen una distribución con los coeficientes de asimetría y curtosis que cabría esperar de una normal<sup>11</sup>. Para ello se siguen los siguientes pasos:

- 1 Obtención de la muestra ordenada de las observaciones de los rendimientos en valor absoluto,  $y_1 \geq y_2 \geq \dots \geq y_T$ , siendo  $y_i = |r_i| \quad \forall i = 1, \dots, T$ . Una cuantía más alta implica un movimiento fuerte en la cotización, no siendo relevante, en este caso, si se trata de una subida o una bajada de precio.
- 2 Contratación de la normalidad (Jarque-Bera o similar), para el subconjunto  $\{r_j\} \quad \forall j = s, \dots, T$ .
- 3 Si se rechazase la hipótesis de normalidad, se elimina el rendimiento más alto en valor absoluto ( $y_1$ ) de la muestra y se vuelve a empezar.

Este proceso se repite hasta que la muestra reducida presente el máximo de la probabilidad asociada al estadístico del contraste<sup>12</sup>. En esta fase se han utilizado tres contrastes de normalidad: el contraste de Bondad del Ajuste, el contraste Kolmogorov-Smirnov y el contraste Jarque-Bera<sup>13</sup>. Aunque con los tres contrastes se obtienen resultados similares, utilizaremos este último para la estimación posterior del VaR, porque tiene en cuenta en su cálculo los coeficientes de curtosis y asimetría.

Al ir ajustando la distribución y descender el número de observaciones, la desviación típica de la nueva muestra será menor que la de la serie original. Si se disminuye el tamaño de la muestra, la distribución cambia de forma y se vuelve más apuntada, al concentrarse la probabilidad en torno a la media de la nueva distribución normal. Los valores de asimetría y curtosis de los rendimientos que se mantienen dentro de la muestra se reducen al disminuir los valores anómalos contenidos en la distribución central. Como se puede apreciar en el cuadro 2, las cifras de asimetría son ahora próximas a cero y las de curtosis próximas a tres. En consecuencia, el contraste de normalidad Jarque-Bera obtiene probabilidades asociadas al estadístico próximas a uno, por lo que podemos aceptar la normalidad en las observaciones que han quedado en el centro de la distribución. Al mismo tiempo, a través del porcentaje de observaciones eliminadas con este procedimiento podemos conocer el umbral correspondiente en cada caso, a partir del cual los valores extremos se comportan como una distribución de Pareto.

---

10. Véase Dress et al. (2000) para una exposición de métodos alternativos. 11. Véase cuadro 2. 12. Véase cuadro 1. 13. Véase Jarque y Bera (1980).

ÍNDICE	TODA LA MUESTRA			SIN VALORES EXTREMOS		
	DESV. TÍPICA (%)	ASIMETRÍA	CURTOSIS	DESV. TÍPICA (%)	ASIMETRÍA	CURTOSIS
AEX ÁMSTERDAM	1,90	0,11	5,59	1,44	0,01	3,04
ATENAS ASE	1,60	-0,04	7,01	1,26	0,06	3,09
BANGKOK	1,75	-0,58	16,38	1,39	0,03	3,04
BCN MID-50	28,51	-0,08	208,09	0,60	-0,01	2,98
BOVESPA	2,03	-0,01	4,27	1,74	-0,04	2,87
BRUSELAS BEL	1,52	0,43	6,95	1,01	-0,11	3,02
CAC40 PARÍS	1,78	0,05	4,45	1,55	0,01	3,01
COPENHAGUE	1,37	-0,01	4,41	1,22	0,04	3,02
DAX XETRA	2,14	-0,18	8,98	1,73	-0,05	3,05
DJ COMPOSITE	1,27	-0,25	6,31	1,04	0,00	2,99
EL CAIRO	0,66	0,49	7,02	0,43	-0,02	3,05
EUROTOP 100	1,62	0,03	4,53	1,37	0,02	2,99
FTSE-100	1,41	-0,08	4,55	1,22	0,03	2,98
HANG SENG	6,24	0,12	426,92	1,34	0,01	3,04
HELSINKI HEX	16,22	0,10	465,56	2,42	0,05	2,92
IBEX-35	1,64	0,09	3,61	1,53	0,02	2,97
JOHANESBURGO	1,25	-0,25	7,07	1,07	0,00	2,99
KUALA LUMP.	1,10	-0,44	8,17	0,80	0,12	2,84
LIMA IG	0,83	0,06	7,76	0,63	-0,02	3,02
MERVAL	2,81	0,76	8,35	1,92	0,03	3,04
MÉXICO IPC	1,56	0,08	5,15	1,31	-0,07	3,02
MIB 30	1,54	-0,12	5,00	1,37	0,03	3,00
MIBTEL	1,39	-0,22	5,21	1,22	0,02	2,99
MOSCÚ RST	2,42	-0,44	5,13	2,01	-0,03	2,94
NASDAQ 100	2,98	0,25	4,82	2,69	0,00	3,04
NIKKEI	6,23	-0,40	367,81	1,62	0,03	3,11
S&P 500	1,39	0,01	4,97	1,20	0,00	3,02
SANTIAGO IGP	0,61	0,35	6,12	0,49	0,08	3,07
SIDNEY A. OR.	0,77	-0,64	10,21	0,65	-0,02	2,98
SINGAPUR ST	1,34	-0,48	6,90	1,16	0,05	2,96
SLOUGH	2,76	-0,52	30,54	1,90	-0,05	2,99
SWISS MARKET	10,57	-0,87	477,53	1,12	0,01	2,98
TEL-AVIV 100	1,65	0,22	7,75	1,36	0,07	3,10
TSE-WEIGHTED	10,94	0,39	424,22	1,72	0,13	3,00
VIENA ATX	0,90	-0,28	5,17	0,73	0,03	2,96
YAKARTA	1,45	-0,76	7,85	1,22	-0,01	3,08

De esta forma, se ha aplicado un método cuya ventaja, frente a la representación gráfica del estimador de Hill, reside en el cálculo automático y objetivo del umbral de la distribución de Pareto, lo que permitirá la posterior aplicación de la metodología semiparamétrica de VaR asociada a la teoría de los valores extremos.

#### 4 Evaluación de la capacidad de cobertura de las medidas de valor en riesgo

Para la evaluación de las metodologías VaR mencionadas, partimos de los valores obtenidos diariamente para 36 índices bursátiles mundiales con las cuatro metodologías descritas para el período 1999-2003. Para ello, comenzaremos con una ventana de 500 observaciones, que utilizaremos para estimar la pérdida máxima esperada para el día siguiente

con un nivel de confianza dado<sup>14</sup>. A continuación, comprobamos si la predicción ha sido correcta o no. Volvemos entonces a realizar un nuevo cálculo, pero moviendo un día el intervalo de tiempo utilizado, lo que repetimos sucesivamente hasta cubrir el total de la muestra.

Una vez obtenidas las estimaciones, procedemos a realizar una comparación de los resultados conseguidos con cada método, lo que permite elaborar una ordenación de preferencias entre metodologías de VaR. Con este fin utilizamos tres criterios de evaluación: la tasa de fallos en la predicción de la pérdida máxima esperada, valor medio de cobertura requerido y la suma de las pérdidas mayores que las esperadas.

#### 4.1 ANÁLISIS DE LA TASA DE FALLOS DEL VaR

Tal y como se definió en la introducción, el VaR determina la pérdida máxima esperada para una cartera en un plazo establecido. Esta medida no es exacta, pues está asociada a un nivel de confianza, lo cual implica que existe una determinada probabilidad de sufrir una pérdida superior. Una buena metodología de cálculo del VaR debería obtener unos aciertos de predicción en el largo plazo similares al nivel de confianza establecido a priori.

Por tanto, el primer criterio que vamos a utilizar para evaluar las metodologías propuestas será la tasa de fallos en la predicción de la pérdida máxima. Esta medida consiste en obtener el porcentaje de rendimientos (negativos) que han sido superiores a la cota que se obtenía por el VaR a lo largo de la muestra. Un valor inferior al marcado implicaría que las medidas que se hubieran tomado para cubrir esa eventualidad extrema no habrían sido suficientes. Si el número de ocasiones en que esto sucede es demasiado elevado, se tendría un serio problema, pero también si la frecuencia fuera excesivamente baja sería una señal de que el ajuste no es bueno y de que el método es mejorable.

El criterio de la tasa de fallos debe llevar a descartar los métodos que supongan una sistemática infravaloración del VaR, pues la exposición al riesgo resultante será superior a la estimada. Por otro lado, de entre aquellos cuyo porcentaje de errores sea inferior al esperado, será preferible el que se acerque más a ese nivel, pues no es eficiente tener que cubrirse de riesgos inexistentes.

Si consideramos individualmente cada uno de los 36 índices bursátiles utilizados en nuestro análisis, no resulta posible indicar preferencias entre ninguno de los métodos de cálculo del VaR de forma sistemática para todos ellos, por lo que será necesario hacer las comparaciones respecto a los valores medios obtenidos para el total de las carteras de la muestra. Esto mismo es lo que haremos en los otros dos criterios.

En el cuadro 3 se presentan las tasas medias de fallos para el VaR a un día con un nivel de confianza del 99% promediada para los 36 índices bursátiles desde 1999 hasta 2003. Como se puede observar, el método que asume una distribución de probabilidad normal de los rendimientos financieros se queda por encima de lo esperado para el nivel de confianza del 99%, mientras que entre los demás métodos, que sí logran mantener las tasas de fallos por debajo del 1%, el peor de ellos sería el de simulación histórica, siendo los modelos paramétricos o semiparamétricos los que conseguirían resultados más próximos al objetivo. En el cuadro 4 se encuentran los contrastes de igualdad de medias de estos valores. En ellos se verifica que es el método normal el que consigue unos valores sistemáticamente peores que los demás.

---

14. Utilizaremos como referencia el VaR para un nivel de confianza del 99%, que se aproxima a los requerimientos del BIS.

	MÉTODO DE ESTIMACIÓN			
	DISTRIBUCIÓN NORMAL	SIMULACIÓN HISTÓRICA	MODELO GARCH	VALORES EXTREMOS
<b>Valores medios</b>				
Tasa de fallos (%)	1,17	0,58	0,82	0,77
Valor medio VaR	-0,076	-0,044	-0,063	-0,043
Excesos sobre el VaR	-0,292	-0,275	-0,272	-0,282
<b>Valores medianos</b>				
Tasa de fallos (%)	0,82	0,43	0,81	0,48
Valor medio VaR	-0,038	-0,044	-0,036	-0,041
Excesos sobre el VaR	-0,034	-0,020	-0,020	-0,023
<b>Desviación típica</b>				
Tasa de fallos (%)	1,24	0,53	0,49	0,71
Valor medio VaR	0,112	0,016	0,082	0,016
Excesos sobre el VaR	0,864	0,903	0,877	0,902

Sin embargo, si lo que consideramos no es la media de la tasa de fallos, sino la mediana, se observa cómo los valores bajan, e incluso la distribución normal consigue valores aceptables (por debajo del 1%). Esto es debido a la alta variabilidad que presenta este estimador según el mercado que se analice, de tal manera que la capacidad de acierto se reduce en gran medida cuando nos encontremos con carteras muy volátiles.

#### 4.2 ANÁLISIS DEL VaR MEDIO

La utilidad principal del VaR radica en el establecimiento de los requisitos necesarios para hacer frente a las pérdidas máximas esperadas. En cualquier caso, estas necesidades de cobertura suponen un coste tanto para la entidad como para la economía en general, pues se están inmovilizando recursos, con la consiguiente aparición de costes de oportunidad.

Por tanto, el segundo criterio de evaluación consistirá en elegir aquella metodología que, logrando los objetivos marcados por el nivel de confianza, lo haga con unos valores de VaR lo más reducidos posible. Es importante que la metodología esté funcionando, es decir, que obtenga una tasa de fallos dentro de los valores esperados, pues, en caso contrario, no estaríamos cumpliendo el objetivo de abaratar los costes de cobertura, sino simplemente deteriorando la calidad de esta.

Para la estimación de los costes de cobertura utilizaremos la media de las cotas marcadas por el VaR para cada día de la muestra. De nuevo, y como sucedía en el apartado anterior, dada la variabilidad de los resultados obtenidos en cada uno de los índices, no existe una medida que sea superior a las otras para todos ellos, por lo que las comparaciones se realizarán sobre las medias de los valores obtenidos en cada mercado.

En el cuadro 3 se recogen los VaR medios obtenidos con las cuatro aproximaciones para el nivel de confianza del 99%<sup>15</sup>. Siguiendo este criterio, las mejores metodologías encon-

<sup>15</sup>. A diferencia de lo que sucedía con la tasa de fallos, estos resultados no se ven afectados por cambios en el nivel de confianza utilizado.

	MÉTODO DE ESTIMACIÓN		
	SIMULACIÓN HISTÓRICA	MODELO GARCH	VALORES EXTREMOS
<b>Tasa de fallos</b>			
Distribución normal	2,389**	2,252**	1,679*
Simulación histórica		0,280	0,937
Modelo GARCH			0,715
<b>Valor medio VaR</b>			
Distribución normal	0,041	0,785	0,393
Simulación histórica		1,018	0,520
Modelo GARCH			0,583
<b>Excesos sobre el VaR</b>			
Distribución normal	2,584**	2,376**	1,649*
Simulación histórica		1,992*	1,338
Modelo GARCH			0,287

(\*) Indica que se rechaza la hipótesis de igualdad de medias a un nivel de significación del 10%.

(\*\*) Indica que se rechaza la hipótesis de igualdad de medias a un nivel de significación del 5%.

tradas son las que utilizan valores extremos y simulación histórica, pues exigen unas cantidades menores para la cobertura del riesgo de mercado. Por el contrario, el método GARCH exige unos costes superiores hasta en un 47% respecto a los de otras metodologías. Sin embargo, considerando la cartera mediana, los valores serían muy similares entre ellos; de hecho, los contrastes de igualdad de medias no rechazan esta hipótesis en ningún caso.

Las diferencias entre ambas magnitudes se deben a la alta variabilidad que el VaR medio tiene frente a la cartera considerada. De hecho, la desviación típica entre las carteras utilizadas es cinco veces superior en el caso del modelo GARCH frente a valores extremos o simulación histórica. Esto es debido a la forma misma del primero, que predice aumentos de la volatilidad en el período que sigue a la aparición de un cambio brusco en el valor de la cartera, circunstancia que no siempre se materializa en la práctica; por ello, el modelo paramétrico es menos adecuado en aquellas carteras que presentan una alta variabilidad en los rendimientos. En estas carteras, los métodos semiparamétricos o no paramétricos exigen una menor provisión, con igual grado de cobertura.

#### 4.3 ANÁLISIS DE LA SUMA DE PÉRDIDAS MAYORES QUE EL VaR

A la vista de la discusión del criterio anterior, podría aducirse que, si bien es cierto que se puede lograr similar tasa de fallos con una menor cifra media del VaR, el grado de cobertura obtenido no sería el mismo. En aquellos casos en los que el rendimiento fuera inferior al máximo esperado, la pérdida sufrida sería entonces superior cuanto menor fuera el VaR fijado en ese momento, lo que seguiría haciendo preferibles, en términos de seguridad, aquellas metodologías que exigieran mayor protección.

Por ello, el último criterio que vamos a considerar para evaluar las distintas aproximaciones será uno que indique cuáles son las pérdidas que quedan sin cubrir en cada una de ellas. La forma en la que lo vamos a cuantificar va a ser a través de la suma de pérdidas que excedan los valores proporcionados por el VaR.

En el cuadro 3 se observa que las pérdidas por encima del VaR son similares en todos los métodos, lo que implicaría que, independientemente de las exigencias de provisiones planteadas con cada metodología, no existen diferencias apreciables en el grado de cobertura de riesgo alcanzada por los modelos GARCH y valores extremos.

Estos resultados son consecuencia de que los fallos en la predicción del VaR en el modelo GARCH son debidos a la aparición de caídas bruscas en el valor de las carteras en épocas en las que se estaban haciendo predicciones de volatilidad baja. Como resultado, la cobertura en estos casos será inferior a la de los otros métodos, por lo que, a pesar de los mayores requisitos de provisiones de este método, no se obtiene una protección superior.

## 5 Conclusiones

En este artículo se propone una nueva metodología para la estimación semiparamétrica de los valores extremos en el cálculo del VaR. Con el objetivo de evaluar la capacidad predictiva de esta metodología, se ha realizado una comparación con las metodologías más habituales para el cálculo del VaR, aplicándolas a los rendimientos diarios de 36 índices bursátiles mundiales durante el período 1999-2003.

Dicha evaluación se ha realizado a través de tres criterios, que tienen en cuenta el grado de ajuste del VaR, los costes derivados de las provisiones que se han de realizar y la protección del riesgo proporcionada por cada uno de ellos.

A la vista de los análisis realizados según cada uno de los tres criterios, se puede afirmar que el método de valores extremos propuesto, que plantea un método automático y objetivo para la obtención del umbral que permite separar los valores extremos, obtiene unos grados de ajuste y de protección del riesgo al menos tan buenos como los que se derivan del uso de otras metodologías, pero utilizando unas provisiones inferiores. Estaría, además, especialmente indicado cuando las carteras objeto de análisis tengan una alta volatilidad.

En el caso del VaR basado en los modelos GARCH, logra similares resultados a los valores extremos en cuanto a ajuste y protección del riesgo, pero con el coste de incrementar notablemente las provisiones necesarias. Además, presenta el inconveniente añadido de la alta variabilidad en su cuantía, con lo que complica la gestión y agrega un factor de presión extra sobre los gastos derivados de la recomposición de la cartera y de la vigilancia del riesgo.

Por último, el VaR basado en la simulación histórica de los rendimientos de la cartera, que es de entre las tres metodologías consideradas el más sencillo de calcular, consigue unos niveles similares de protección de riesgo y de provisiones requeridas, pero con un peor ajuste en cuanto a la tasa de fallos. Esto podría interpretarse como un menor número de errores de predicción de la pérdida máxima, pero con la contrapartida de que, en caso de que la caída supere la cuota esperada, la cobertura en esos momentos sería inferior a la que proporcionan otros métodos.

## BIBLIOGRAFÍA

- BALI, T. G. (2003). «An Extreme Value Approach to Estimating Volatility and Value at Risk», *Journal of Business*, vol. 76, pp. 83-108.
- BEDER, T. S. (1995). «VAR: Seductive but Dangerous», *Financial Analysts Journal*, vol. septiembre-octubre, pp. 12-24.
- BEIRLANT, J., J. TEUGELS y P. VYNCKIER (1996). *Practical analysis of extreme values*, Leuven University Press, Lovaina.
- CHRISTOFFERSEN, P., J. HAHN y A. INOUE (2001). «Testing and Comparing Value-at-Risk Measures», *Journal of Empirical Finance*, vol. 8, pp. 325-342.
- COMITÉ DE SUPERVISIÓN BANCARIA DE BASILEA (2003). «El Nuevo Acuerdo de Capital de Basilea», Banco de Pagos Internacionales, documento de consulta.

- DANIELSSON, J., y C. G. DE VRIES (1997). *Value-at-Risk and Extreme Returns*, Working Paper, London School of Economics, Londres.
- DANIELSSON, J., P. HARTMANN y C. G. DE VRIES (1998). «The cost of conservatism: Extreme returns, value-at-risk, and the basle multiplication factor», *Risk*, vol. 11, pp. 101-103.
- DREES, H., L. DE HAAN y S. RESNICK (2000). «How to make a hill plot», *The Annals of Statistics*, vol. 28, pp. 254-274.
- EMBRECHTS, P., S. RESNICK y G. SAMORODNITSKY (1999). «Extreme value theory as a risk management tool», *North American Actuarial Journal*, vol. 3, pp. 30-41.
- ENGLE, R. F., y S. MANGANELLI (2004). «CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles», *Journal of Business & Economic Statistics*, vol. 22, pp. 367-381.
- FAMA, E. (1965). «The behavior of stock market prices», *Journal of Business*, vol. 38, pp. 34-105.
- HENDRICKS, D. (1996). «Evaluation of Value-at-Risk Models Using Historical Data», *Federal Reserve Bank of New York Economic Policy Review*, vol. 2, pp. 39-69.
- HILL, B. (1975). «A simple general approach to inference about the tail of a distribution», *The Annals of Statistics*, vol. 3, pp. 1163-1174.
- JANSEN, D. W., y C. G. DE VRIES (1991). «On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspectiva», *The Review of Economics and Statistics*, vol. 73, pp.18-24.
- JARQUE, C. M., y A. K. BERA (1980). «Efficient tests for normality, homoscedasticity and serial independence of regression residuals», *Economics Letters*, vol. 6, pp. 255-259.
- KUESTER, K., S. MITTNIK y M. PAOLELLA (2006). «Value-at-Risk Prediction: A Comparison of Alternative Strategies», *Journal of Financial Econometrics*, vol. 4, n.º 1, pp. 53-89.
- LONGIN, F. (2000). «From value at risk to stress testing: The extreme value approach», *Journal of Banking & Finance*, vol. 24, pp. 1097-1130.
- MANDELBROT (1963). «The variation of certain speculative prices», *Journal of Business*, vol. 36, pp. 394-419.