Serie Técnica de Estudios - N° 008

Backtesting para modelos internos de medición de riesgos: Determinación estadística de la Tabla de Permanencia

Ximena Alvarez Castillo*.

Analista de Riesgos del Departamento de Estudios.

Dirección de Estudios y Análisis Financiero.

Enero de 2007

Superintendencia de Bancos e Instituciones Financieras
SBIF
Chile

*Se agradecen los útiles comentarios de Julio Acevedo, Gunther Held y Sergio Huerta. Por cierto que los errores que aún subsistan son de exclusiva responsabilidad de la autora.

La colección Documentos de Estudio es una publicación realizada por la Superintendencia de Bancos e Instituciones Financieras (SBIF) de Chile que tiene el objetivo de ofrecer visiones sobre los temas de regulación y supervisión bancaria que son de su competencia.

Al difundir estos trabajos, SBIF espera contribuir a la discusión académica y del mercado en relación a la situación del sistema financiero nacional, sus proyecciones y principales áreas de trabajo. De esta manera busca facilitar el intercambio de ideas y dar a conocer investigaciones, con carácter preliminar, para su discusión y comentarios.

La publicación de estos trabajos no está sujeta a la aprobación previa SBIF ni sus autoridades, por lo que sus contenidos, son de exclusiva responsabilidad de su o sus autores y no reflejan necesariamente la opinión de esta Superintendencia.

SBIF
Moneda 1123, Santiago de Chile
Casilla 15-D

Mail: publicaciones@sbif.cl
Web: www.sbif.cl

Resumen Ejecutivo

Las pruebas para evaluar el desempeño de las metodologías empleadas en el cálculo del Valor en Riesgo (VaR) son conocidas como pruebas de Backtesting. El Backtesting es un procedimiento estadístico utilizado para validar la calidad y la precisión de un modelo VaR, mediante la comparación de los resultados reales de las posiciones de *trading* y las medidas de riesgo generadas por los modelos.

En materia de regulación, el Comité de Basilea establece que los requerimientos mínimos de capital para riesgos de mercado correspondan al máximo entre (i) el VaR del día anterior; y (ii) el promedio de los VaR calculados para los últimos 60 días multiplicado por 3 más un coeficiente s_t (denominado "factor aditivo") relacionado con la calidad del modelo. Si el modelo no es del todo satisfactorio, s_t aumenta desde 0 hasta llegar a un valor máximo de 1. Este procedimiento genera incentivos para que los administradores de riesgo procuren mantener modelos VaR bien calibrados.

Conforme a la nueva normativa sobre riesgos de mercado que debe observar la banca chilena, esos mismos procedimientos deben aplicarse para calcular el riesgo de mercado afecto a límite para evaluar los modelos internos. De hecho, el resultado de los Backtesting es uno de los factores críticos que esta Superintendencia considerará para autorizar el uso de modelos VaR para fines de medir el riesgo afecto a límite normativo para las exposiciones en riesgo de mercado.

Este documento replica los criterios estadísticos propuestos por el Comité de Basilea para determinar los rangos, en términos del número de excepciones (existirá una excepción cuando el valor del VaR sea menor que la perdida real en el lapso de un día), que delimitan cada zona de la "Tabla de Permanencia" (Tabla que define los factores multiplicativos que deben ser aplicados al VaR, para el requerimiento de capital), para cualquier tamaño de muestra (días) y presenta un ejemplo para llevarla a la práctica. Con ello, se establece un referente metodológico para construir la Tabla de Permanencia para los factores de multiplicación a ser utilizados para efectos de computar el riesgo de mercado afecto a límite normativo.

En términos generales, el objetivo de este documento es presentar los procedimientos metodológicos que debieran seguirse para realizar Backtesting a modelos VaR que trabajen con muestras superiores a 250 días y, de esa manera, determinar los cambios que presentaría la Tabla de Permanencia prevista en el numeral 18 del Título I del Capítulo 12-9 de la Recopilación Actualizada de Normas.

Contenidos

Introducción

La enmienda del comité de Basilea para incorporar riesgo de mercado del primer acuerdo de capital[1] establece una metodología estándar para medir esos riesgos pero, además, permite que los bancos utilicen modelos de Valor en Riesgo (VaR), si éstos son aprobados por las autoridades de supervisión.

Si un banco está utilizando un modelo propio aceptado por el supervisor, el Comité de Basilea prevé que el requisito mínimo de capital (RMC) sea calculado de la siguiente manera:

$$RMC_{t+1} = Max\left[VaR_t \,; (3+s_t)\cdot \frac{1}{60}\sum_{t=1}^{60} VaR_{t-1} \right]$$

Es decir, el requisito será el máximo valor entre el VaR del día anterior y el promedio de los últimos 60 VaR diarios multiplicado por un factor que es igual a 3 más un coeficiente s_t (denominado “factor aditivo”) relacionado con la calidad del modelo. Si el modelo no es del todo satisfactorio, s_t irá aumentando desde 0 hasta llegar a un valor máximo de 1. Conforme al capítulo 12-9 de la Recopilación Actualizada de Normas (RAN) esa es, también, la mecánica que deben seguir los bancos para determinar el valor en riesgo afecto a límite normativo, una vez que se les autorice el uso de modelos internos para esos efectos.

Los argumentos del Comité de Basilea para la introducción de un factor multiplicativo de tales características es que ello constituye para las autoridades supervisoras una herramienta para premiar a las entidades bancarias en función de la evaluación del modelo interno de riesgo de mercado y, de ese modo, que los bancos procuren perfeccionar constantemente sus metodologías de medición de riesgo.

Para validar la calidad y la precisión de un modelo VaR, se utiliza una herramienta llamada Backtesting, que es un procedimiento estadístico que mide la calidad de la prueba.

Dado que las técnicas de Backtesting permiten detectar defectos en los modelos de medición de riesgos de mercado, aquellos bancos que han desarrollado e introducido modelos VaR usan normalmente estas técnicas para contrastar la precisión de sus modelos. Adicionalmente, el Comité de Basilea, y los reguladores en general, exigen el uso de Backtesting en forma rutinaria a los bancos que usan metodologías VaR para determinar capitales mínimos regulatorios o, en el caso de la SBIF, para determinar el límite a las exposiciones de riesgos de mercado.

Este documento aplica criterios estadísticos (los mismos que adopta el Comité de Basilea) para determinar los rangos, en términos del número de excepciones (una se

[1] Amendment to the Capital Accord to incorporate market risks. Updated November 2005

produce excepción cuando el valor del VaR sea menor que la perdida real en el lapso de un día), que delimitarían cada zona de la "Tabla de Permanencia", esta última define los factores multiplicativos que deben ser aplicados al VaR, para cualquier tamaño de muestra superior o igual a 250 observaciones diarias.

El objetivo de este documento es presentar los procedimientos metodológicos que debieran seguirse para realizar Backtesting a modelos VaR que trabajen con muestras superiores a 250 días y, de esa manera, determinar los cambios que presentaría la Tabla de Permanencia prevista en el numeral 18 del Capítulo 12-9 de la Recopilación Actualizada de Normas (RAN).

El trabajo está organizado de la siguiente manera. En la primera sección, se analizan las consideraciones estadísticas que respaldan el cálculo del factor multiplicativo. En la sección siguiente se explica, a través de un ejemplo, el cambio en la Tabla de Permanencia al aumentar el tamaño de muestra. Una vez definidas las tres zonas correspondientes, en la cuarta sección se explica cómo se calcula el factor aditivo. Por último, se presentan las principales conclusiones.

1 Fundamentos teóricos para la Tabla de Permanencia

1.1 El factor multiplicativo

Stahl (1997) propuso una justificación teórica al factor multiplicativo del VaR, que hasta ese momento había sido interpretado como un compromiso político un tanto arbitrario. Usa dos argumentos, uno relacionado con la falta de especificación de las colas de la distribución ("fat tails"), y el otro, relacionado con la variación potencial de la distribución de retornos del portafolio según la ventana temporal de datos que se elija (en particular, debido a incrementos no identificados de la varianza de los retornos). Ambos problemas son abordados mediante la desigualdad de Chebychev, es decir, la probabilidad de estar a más de k desviaciones estándares de la media es a lo más $1/k^2$, la que es válida para toda distribución de probabilidades que tenga varianza finita:

$$P\left[\left|X-\mu_X\right|>k\sigma_X\right]\le\frac{1}{k^2}\quad;\quad k \text{ constante positiva.}$$

A partir de esta desigualdad se construye el argumento básico de Stahl y asumiendo normalidad determina que el factor multiplicativo[2] es dado por:

$$\frac{k}{z_{1-\alpha}}$$

Por ejemplo si se trabaja con un 95% de confianza, $\frac{1}{k^2}=0.05 \Rightarrow k=4.47$

Reemplazando el valor de k en la ecuación para calcular el factor multiplicativo que propone Stahl se obtiene un valor de 2.7, tal como se muestra a continuación:

$$\frac{k}{z_{1-\alpha}}=\frac{4.47}{1.64}=2.7$$

Ahora, si se trabaja con un 99% de confianza, $\frac{1}{k^2}=0.01 \Rightarrow k=10$

Reemplazando k en la ecuación para calcular el factor multiplicativo, se obtiene el siguiente resultado:

$$\frac{k}{z_{1-\alpha}}=\frac{10}{2.33}=4.3$$

A este nivel de generalidad resulta un factor multiplicativo muy conservador (entre 3 y 4 según los cálculos anteriores), para cubrir hasta la más exóticas de las distribuciones.

[2] Markus Leippold, Paolo Vanini (2002)

Es así como el multiplicador definido por Basilea, para modelos VaR que son estadísticamente precisos, es de 3, lo que es válido para cualquier tamaño de muestra con un número de observaciones mayor o igual a 250 días.

1.2 Construcción de la Tabla de Permanencia

Teniendo en consideración las limitaciones estadísticas de los métodos de Backtesting, el Comité de Basilea introdujo una estructura simple para interpretar sus resultados, la que contempla un rango de posibles situaciones que dependen de la intensidad de la señal generada por el Backtesting. Estas situaciones son clasificadas en una tabla (la Tabla de Permanencia) que se divide en tres zonas, distinguidas por colores que representan el atributo o jerarquía de cada situación. La zona verde corresponde a resultados del Backtesting que no hacen pensar en un problema con la calidad o exactitud del modelo de medición de riesgos. La zona amarilla abarca resultados que plantean preguntas en este sentido, para las cuales no existen respuestas concluyentes. La zona roja comprende resultados que casi ciertamente indican un problema en el modelo de medición de riesgos.

1.2.1 Consideraciones estadísticas para definir las zonas

Al aplicar Backtesting para examinar el desempeño de un modelo VaR dentro de un período determinado, debe tenerse en cuenta que el período analizado puede estar influido por eventos aleatorios que hacen que las excepciones sumen un número superior al que se esperaría aun cuando el modelo sea preciso, o bien, que sumen un número dentro de lo esperado aun cuando el modelo sea impreciso. Considerando esto, el desempeño de un modelo VaR debe analizarse como una variable aleatoria, que correspondería al número de excepciones en n días.

La metodología propuesta por Basilea se basa en la consideración de que la ocurrencia de una excepción (0: el VaR es mayor o igual a la pérdida real, 1: el VaR es menor que la pérdida real) en determinado día es independiente del resultado de cualquier otro día. De ese modo, si se trabaja con una muestra de 250 observaciones (período que corresponde aproximadamente a los días hábiles bancarios en un año) y con un nivel de confianza del 99%, la probabilidad de tener "x" excepciones en 250 días, con una probabilidad de éxito de p = 1% sigue una distribución Binomial[3] (250, 0.01). Esto da origen a una clasificación por zonas como la mostrada en la Tabla N°1, que es la Tabla de Permanencia que se adopta en la actual normativa SBIF.

Los límites de las tres zonas son definidos de manera de balancear dos tipos de errores estadísticos: el error Tipo I, dado por la probabilidad de rechazar, en base al resultado de Backtesting, la hipótesis de que un modelo es preciso cuando en realidad lo es; y el error Tipo II, dado por la probabilidad de no rechazar, en base al resultado del Backtesting, la hipótesis de que un modelo es preciso cuando en realidad no lo es.

[3] $P(X = x) = \binom{250}{x} \cdot 0.01^{x} \cdot (1-0.01)^{250-x}$

En términos formales la definición de los errores Tipo I y Tipo II es la siguiente:
H_0: El modelo es preciso

	H_0: Verdadera	H_0: Falsa
H_0 rechazada	**Error Tipo I (α)**	Decisión correcta
H_0 no rechazada	Decisión correcta	**Error Tipo II (β)**

$$\alpha = P(Rechazar\ H_0 \mid H_0\ \ \text{Verdadera})$$

$$\beta = P(No\ Rechazar\ H_0 \mid H_0\ \ Falsa)$$

Una hipótesis estadística es una asunción relativa a una o varias poblaciones, que puede ser cierta o no. Las hipótesis estadísticas se pueden contrastar con la información extraída de las muestras y tanto si no se rechazan como si se rechazan se puede cometer un error.

La hipótesis nula se representa por H_0. Rechazar H_0 implica no rechazar una hipótesis alternativa (H_1).

Tabla N°1. Tabla de Permanencia para una muestra de 250 días

Zona	Número de Excepciones Registrados (x)	Factor Aditivo (s_t)	Factor Multiplicativo	$P(X \leq x)$
Verde	0	0	3,00	0,0811
	1			0,2858
	2			0,5432
	3			0,7581
	4			0,8922
Amarilla	5	0,40	3,40	**0,9588**
	6	0,50	3,50	0,9863
	7	0,65	3,65	0,9960
	8	0,75	3,75	0,9989
	9	0,85	3,85	0,9997
Roja	10 o más	1,00	4,00	**0,9999**

La Tabla N°2 muestra las probabilidades exactas de obtener un número particular de excepciones en una muestra de 250 observaciones independientes bajo varios supuestos sobre el porcentaje real de resultados que captura el modelo (entre otros, que las excepciones siguen una distribución Binomial). El lado izquierdo de la Tabla N°2 informa probabilidades asociadas con un modelo preciso (es decir, que tienen un verdadero nivel de cobertura del 99%); de ese modo, la columna “exacta” revela que con una cobertura de 99% pueden esperarse exactamente cinco excepciones en 6.7% de las muestras.

Tabla N° 2. Errores Tipo I y Tipo II para una muestra de 250 días

Modelo Preciso			Modelo impreciso: Alternativas posibles de niveles de cobertura								
Excepciones	Cobertura = 99%		Excepciones	Cobertura = 98%		Cobertura = 97%		Cobertura = 96%		Cobertura = 95%	
	Exacta	Tipo I		Exacta	Tipo II	Exacta	Tipo II	Exacta	Tipo II	Exacta	Tipo II
0	8,1%	100,0%	0	0,6%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
1	20,5%	91,9%	1	3,3%	0,6%	0,4%	0,0%	0,0%	0,0%	0,0%	0,0%
2	25,7%	71,4%	2	8,3%	3,9%	1,5%	0,4%	0,2%	0,0%	0,0%	0,0%
3	21,5%	45,7%	3	14,0%	12,2%	3,8%	1,9%	0,7%	0,2%	0,1%	0,0%
4	13,4%	24,2%	4	17,7%	26,2%	7,2%	5,7%	1,8%	0,9%	0,3%	0,1%
5	6,7%	10,8%	5	17,7%	43,9%	10,9%	12,8%	3,6%	2,7%	0,9%	0,5%
6	2,7%	4,1%	6	14,8%	61,6%	13,8%	23,7%	6,2%	6,3%	1,8%	1,3%
7	1,0%	1,4%	7	10,5%	76,4%	14,9%	37,5%	9,0%	12,5%	3,4%	3,1%
8	0,3%	0,4%	8	6,5%	86,9%	14,0%	52,4%	11,3%	21,5%	5,4%	6,5%
9	0,1%	0,1%	9	3,6%	93,4%	11,6%	66,3%	12,7%	32,8%	7,6%	11,9%
10	0,0%	0,0%	10	1,8%	97,0%	8,6%	77,9%	12,8%	45,5%	9,6%	19,5%
11	0,0%	0,0%	11	0,8%	98,7%	5,8%	86,6%	11,6%	58,3%	11,1%	29,1%
12	0,0%	0,0%	12	0,3%	99,5%	3,6%	92,4%	9,6%	69,9%	11,6%	40,2%
13	0,0%	0,0%	13	0,1%	99,8%	2,0%	96,0%	7,3%	79,5%	11,2%	51,8%
14	0,0%	0,0%	14	0,0%	99,9%	1,1%	98,0%	5,2%	86,9%	10,0%	62,9%
15	0,0%	0,0%	15	0,0%	100,0%	0,5%	99,1%	3,4%	92,1%	8,2%	72,9%

La parte derecha de la Tabla N°2 recoge las probabilidades asociadas a modelos cuyos verdaderos niveles de cobertura son 98%, 97%, 96%, y 95%, respectivamente.

La Tabla N°2 muestra además varias probabilidades de error. Bajo el supuesto de que el modelo cubre verdaderamente el 99% de los resultados (el nivel deseado de cobertura), la Tabla entrega la probabilidad de que, seleccionando un número dado de excepciones como un umbral para rechazar la precisión del modelo VaR, haya un rechazo erróneo de un modelo preciso (Error Tipo I). Por ejemplo, si el umbral que se fija corresponde a sólo una excepción, entonces se rechazarán modelos precisos el 91.9% del tiempo, porque de ellos sólo se escapará del rechazo un 8.1% de casos (aquellos que no generan excepción alguna). Cuando el umbral de excepciones aumenta, la probabilidad de este tipo de error declina.

Si el verdadero nivel de cobertura no fuese de 99%, el lado derecho de la Tabla N°2 muestra la probabilidad de que, seleccionando un número dado de excepciones como un umbral para rechazar la precisión del modelo, hay una aceptación errónea de un modelo impreciso (Error Tipo II), puesto que el verdadero nivel de cobertura es inferior a 99%. Por ejemplo, si el nivel real de cobertura del modelo es de 97%, y el umbral para rechazar se fija en siete o más excepciones, la Tabla indica que este modelo podría ser aceptado erróneamente como preciso el 37.5% de las veces.

Al interpretar la información de la Tabla N°2, es importante tener en cuenta que, aunque los modelos alternativos aparecen cerca de los niveles de cobertura deseados en términos de probabilidad (97% está cerca de 99%), la diferencia en términos de la medición de riesgo generada puede ser sustancial. Por ejemplo, en el caso de que los resultados se distribuyan como una Normal, el percentil 97 corresponde a 1.88 desviaciones estándar, mientras el percentil 99 corresponde a 2.33 desviaciones estándar, lo que equivale a un aumento de casi 25% en la dispersión. Así, distinguir entre modelos que proporcionan 99% de cobertura de aquellos que dicen proporcionar

99%, pero que en realidad sólo alcanzan un 97% de cobertura, es una cuestión que reviste gran importancia desde el punto de vista de la medición de riesgos y de la evaluación que se haga de la calidad de esas mediciones.

Es preciso señalar que esta metodología sólo se centra en el número de excepciones.

Existen por lo menos otros dos aspectos que deben ser evaluados con la debida atención: i) la ocurrencia sistemática de excepciones consecutivas (lo que colocaría en tela de juicio el supuesto de independencia de las excepciones y, por tanto, el propio Backtesting); y ii) la magnitud de las mismas (que sistemáticamente ocurran muy pocas pero grandes excepciones, puede ser más preocupante que ocurran más excepciones, pero de baja magnitud).

1.2.2 Delimitación de las zonas

Los resultados en la Tabla N°2 también muestran algunas de las limitaciones estadísticas de Backtesting. En particular, que no existe ningún umbral de excepciones que entregue tanto una probabilidad baja de rechazar erróneamente un modelo preciso como una probabilidad baja de aceptar erróneamente todos los modelos imprecisos. Es por ello que el método propuesto por el Comité de Basilea no se apoya en un único umbral y clasifica los resultados en tres categorías, que involucran la definición de dos umbrales. En la zona verde, los resultados de la prueba son consistentes con un modelo preciso y la probabilidad de aceptar erróneamente un modelo impreciso es baja. En la zona roja, los resultados de la prueba son sumamente improbables de haber sido generados por un modelo exacto, y la probabilidad de rechazar erróneamente un modelo preciso es casi nula o nula.

Entre esos dos extremos, sin embargo, hay una zona donde los resultados del Backtesting pueden ser consistentes con modelos precisos o imprecisos (zona amarilla).

La zona verde

Para un modelo de 250 observaciones que proporciona efectivamente 99% de cobertura, sería bastante probable que se produzcan a lo más cuatro excepciones, por lo que hay poca razón para preocuparse por los resultados del Backtesting que caigan en ese rango. Esta conclusión se apoya en los resultados mostrados en la Tabla N°2, que indican que, en esa zona, existe una probabilidad muy baja de aceptar erróneamente un modelo impreciso.

La zona amarilla

Los resultados que caen en esta zona son posibles tanto para los modelos precisos como imprecisos, aunque la Tabla N°2 sugiere que para este rango de excepciones los resultados son más probables para modelos imprecisos que para modelos precisos. Es más, la Tabla N°2 muestra que la probabilidad de aceptar erróneamente la hipótesis de que el modelo es impreciso aumenta con el número de excepciones.

El Comité propone que, dentro de la zona amarilla, el número de excepciones debe ser usado como criterio para aumentar los requisitos de capital. La Tabla N°1 muestra la pauta propuesta por el Comité para los aumentos del factor multiplicativo aplicable a los modelos VaR para fines de requisitos de capital, los que son reflejo de los resultados del Backtesting en una muestra de 250 observaciones.

Subyacente a este método, se visualiza una estructura de incentivos que propiciaría el desarrollo de modelos internos que tengan mayor precisión. Como se observa, en la Tabla N°1, los requisitos de capital aumentan con el número de excepciones; lo que, por lo demás, refuerza la noción de que entre más excepciones se verifiquen hay más dudas acerca de la cobertura del modelo y que, en consecuencia, los cargos de capital deben ser coherentes con esa constatación.

En efecto, la Tabla N°1 refleja la idea general de que el aumento en el factor de multiplicación debe ser suficiente para transformar un modelo con una cobertura inadecuada a uno con cobertura del 99%. Por ejemplo, cinco excepciones en una muestra de 250 implican sólo un 98% de cobertura. Así, el aumento en el factor multiplicativo debe ser suficiente para transformar un modelo con una cobertura de 98% en uno con cobertura de 99%. Es necesario destacar que estos cálculos se apoyan en supuestos estadísticos que probablemente no se mantengan en todos los casos. Así, si se supone que los resultados siguen una distribución Normal Estándar, la razón entre el percentil 99 y el percentil 98 es aproximadamente 1.13 y, por consiguiente, el aumento necesario en el factor multiplicativo, es aproximadamente 0.40[2]. Si la distribución real no es Normal, pero en cambio tiene colas pesadas, debieran exigirse aumentos mayores para alcanzar el percentil 99. La preocupación sobre las colas pesadas también es un factor importante en la elección de los incrementos específicos mostrados en la Tabla N°1.

Es importante, sin embargo, que estos aumentos no sean completamente automáticos. La Tabla Nº2 indica que los resultados en la zona amarilla no siempre implican un modelo impreciso. No obstante, para mantener los incentivos en la dirección correcta, los resultados del Backtesting que caigan en la zona amarilla generalmente deben asumirse como causal para un aumento en el factor multiplicativo, a menos que el banco pueda demostrar que ese aumento no procede. En otras palabras, en esas situaciones el peso de la prueba de la existencia de un problema no debe recaer en el supervisor, sino que debe ser el propio banco el que demuestre que su modelo es confiable. En tal situación, hay diferentes tipos de información adicional que podría ser pertinente para evaluar el modelo del banco.

En consecuencia, y como lo prevé el numeral 9 del capítulo 12.9 de la RAN, los bancos deben documentar todas las excepciones generadas por sus Backtesting, incluyendo las explicaciones para dichas excepciones. Esta documentación es importante para una adecuada acción del supervisor frente a los resultados que caigan en la zona amarilla. Los bancos también pueden implementar Backtesting para intervalos de confianza con percentiles distintos a 99, o pueden realizar otras pruebas estadísticas, no consideradas aquí, (en el Anexo a este documento, se presentan algunos ejemplos de estas otras pruebas), esa información puede proporcionar elementos útiles para evaluar un modelo VaR.

[2] $3 \cdot 1.13 = 3.40$; por lo tanto el factor aditivo para 5 excepciones es 0.40

La zona roja

Finalmente, en contraste con la zona amarilla donde los resultados del Backtesting pueden ser sometidos al juicio del supervisor, los resultados en la zona roja deben llevar generalmente a una presunción automática de que existe un problema en la confiabilidad o en la precisión del modelo del banco, ya que es sumamente improbable que un modelo preciso genere tantas excepciones independientes.

Como muestra la Tabla 1, si se verifican diez o más excepciones en una muestra de 250 observaciones el modelo caería en la zona roja y, automáticamente, el factor multiplicativo aplicable debe ser aumentado en uno (de tres a cuatro). Sin embargo, es preciso tener en cuenta que, aunque diez excepciones es un número muy alto para 250 observaciones, puede haber ocasiones, muy raras, en que haya una razón válida para que un modelo adecuado produzca tantas excepciones. En particular, cuando los mercados financieros están sujetos a cambios extremos, puede esperarse que muchas volatilidades y correlaciones, cambien sustancialmente. A menos que el modelo del banco esté diseñado para actualizar instantáneamente las estimaciones de volatilidad y de correlación, esos cambios pueden generar varias excepciones en un período corto de tiempo.

Como advierte el Comité de Basilea, esas excepciones debieran ser mutuamente dependientes y, por consiguiente, la acción de supervisión apropiada no podría ser la misma que la que se tomaría si se constatasen diez excepciones, pero cada una atribuible a una causalidad distinta. En este caso, una posible acción de la entidad supervisora puede ser simplemente exigir al banco que su modelo tome en cuenta tan rápidamente como sea posible los cambios en las condiciones de mercado. Debe enfatizarse, eso sí, que esa anormalidad sólo debe permitirse bajo circunstancias extraordinarias puesto que, como ya se indicó, la regla general es aplicar un aumento automático en los requisitos de capital cuando los resultados del Backtesting planteen alguna duda acerca de la precisión del modelo VaR.

2 Determinación de las zonas para muestras de tamaño de n días

De acuerdo a lo señalado, si se supone que el número de excepciones en n días sigue una distribución Binomial, para cada una de las excepciones que componen las zonas está asociada la probabilidad acumulada de que a lo más se presenten "x" excepciones en una muestra de n días (para una muestra de 250 días esto está recogido en la última columna de la Tabla N°1).

Como propone el Comité de Basilea[4], para otros tamaños de muestra los límites deben ser deducidos calculando las probabilidades Binomiales acumuladas asociadas a la cobertura verdadera del 99%, tal como se muestra en la Tabla Nº1. De esta manera, la zona amarilla comienza en un punto tal que la probabilidad acumulada de obtener "x" excepciones sea igual o superior al 95%. Similarmente, la zona roja comenzará en un punto tal que la probabilidad acumulada de obtener ese número sea igual o superior al 99.99%.

2.1 Determinación de los umbrales

Para determinar el umbral para la zona amarrilla es necesario encontrar el número de excepciones en un tamaño de muestra "n" en que la probabilidad acumulada sea igual o superior a 0.95, lo cual se obtiene de la siguiente ecuación:

$$P(X \leq x_A) = \sum_{i=0}^{x_A} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \geq 0.95 \qquad i = 0,\ldots\ldots,x_A$$

Al conocer el tamaño de muestra que utilice el banco en el cálculo de Backtesting, sólo queda como incógnita el número de excepciones que cumple esa igualdad (x_A). Despejando el valor de x_A, se podrá determinar el nuevo umbral para la zona amarilla y, por lo tanto, la zona verde termina en la excepción anterior ($x_A - 1$).

Para delimitar la zona roja se sigue el mismo procedimiento, pero esta vez se necesita despejar la siguiente ecuación:

$$P(X \leq x_R) = \sum_{i=0}^{x_R} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \geq 0.9999 \qquad i = 0,\ldots\ldots,x_R$$

Obteniendo el valor de x_R se encuentra el umbral que representa el punto donde comienza la zona roja y, por lo tanto, la zona amarilla termina en la excepción anterior ($x_R - 1$).

Siguiendo este procedimiento se pueden calcular los umbrales de las tres zonas de Backtesting, para cualquier tamaño de muestra que la entidad utilice en el cálculo del VaR (que, en todo caso, debe ser superior a 250).

---

4 Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements

2.2 Un Ejemplo

A modo de ejemplo, a continuación se obtiene la Tabla de Permanencia, utilizando la metodología expuesta anteriormente, para un tamaño de muestra de 500 días.

Delimitación de las zonas

Determinación del umbral para la zona amarilla:

$$P(X \leq x_A) = \sum_{i=0}^{x_A} \binom{500}{i} \cdot 0.01^i \cdot (1-0.01)^{500-i} \geq 0.95 \quad i = 0,.....,x_A$$

Al despejar de esta ecuación x_A el valor que resulta es 9, es decir, la zona amarilla comenzará en 9 excepciones y, por lo tanto, la zona verde estará comprendida entre 0 y 8 excepciones.

Para calcular el umbral de la zona roja se realiza el mismo procedimiento, pero esta vez se necesita despejar la siguiente ecuación:

$$P(X \leq x_R) = \sum_{i=0}^{x_R} \binom{500}{i} \cdot 0.01^i \cdot (1-0.01)^{500-i} \geq 0.9999 \quad i = 0,.....,x_R$$

Del resultado se desprende que la zona roja comienza en 15 excepciones, por lo que la zona amarilla está comprendida entre 9 y 14 excepciones. Los límites que demarcan las zonas obtenidas se muestran en la Tabla N° 3.

Cálculo del factor aditivo

Para obtener el factor aditivo en una muestra de 500 observaciones, como se mencionó anteriormente, es necesario calcular los valores particulares que hagan que el aumento en el factor multiplicativo sea suficiente para transformar un modelo con cobertura menor que 99% en uno con cobertura de 99%.

De acuerdo a lo ya expuesto, cinco excepciones en una muestra de 250 observaciones implican la cobertura de solamente 98% entonces, si la distribución de resultados se supone Normal Estándar, el cuociente entre los percentiles de 99 y 98 es aproximadamente 1.13, por lo que el aumento necesario en el factor multiplicativo es aproximadamente 0.40. Para efectos de cálculo del factor aditivo para una muestra de 500, el raciocinio es el mismo; es decir, 9 excepciones implican una cobertura aproximada de 98,2%, lo que envolvería calcular el cuociente entre el percentil del 99 y el 98,2, que es de 1.11, por lo que el aumento necesario en el factor multiplicativo es de aproximadamente 0.35, y así sucesivamente.

Se debe notar que el factor aditivo calculado para 15 excepciones (inicio de la zona roja) es de 0.71 y que un valor aproximado a 1 sólo se alcanza con 20 observaciones. Sin embargo, adoptar esos factores aditivos para la zona roja conduce a inconsistencias debido a que en una muestra de menor tamaño (de 250 observaciones, por ejemplo), factores aditivos de esa magnitud están dentro de la zona amarilla.

Es por ello que se debe aplicar un único factor aditivo a la zona roja. En particular, al pasar de la zona amarilla a la roja, para cualquier tamaño de muestra, el factor multiplicativo debe aumentar automáticamente en uno (de tres a cuatro).

Aunque lo anterior involucra que el salto sea más violento mientras mayor sea el número de observaciones, existe una justificación metodológica para ello: los umbrales son definidos en términos de un valor crítico para la probabilidad Binomial acumulada (que es discreta) de que ocurran a lo mas "x" excepciones en una muestra de tamaño n; en tanto que el factor aditivo se obtiene suponiendo que los retornos siguen una distribución Normal (que es continua) y está determinado en función del cuociente entre el percentil de cobertura verdadero del VaR y el percentil teórico para una muestra del mismo tamaño.

En todo caso, este procedimiento constituye un mayor incentivo para asegurar la precisión y confiabilidad de los modelos para los bancos que utilicen modelos internos, para la medición de riesgo de mercado, a efecto de calcular el límite normativo.

La Tabla de Permanencia que se obtendría para 500 observaciones se muestra en la Tabla N°3. Cabe advertir que, como hace Basilea, los valores de los factores aditivos se han redondeado hacia 0.05 ó 0.10; los cómputos exactos se muestran en la Tabla A1 del Anexo 1.

Tabla N°3. Tabla de Permanencia para una muestra de 500 días

Zona	Número de Excepciones Registrados (x)	Factor Aditivo (s_t)	Factor Multiplicativo	$P(X \leq x)$
Verde	0	0	3	0,0066
	1			0,0398
	2			0,1234
	3			0,2636
	4			0,4396
	5			0,6160
	6			0,7629
	7			0,8677
	8			0,9329
Amarilla	9	0,35	3,35	0,9689
	10	0,40	3,40	0,9868
	11	0,50	3,50	0,9948
	12	0,55	3,55	0,9981
	13	0,60	3,60	0,9994
	14	0,65	3,65	0,9998
Roja	15	1,00	4,00	0,9999

3 Conclusión

El Backtesting, además de constituir un intento de evaluar estadísticamente la precisión de los modelos VaR, puede ser usado como una herramienta para incentivar a los bancos a que evalúen constantemente sus modelos internos de medición de riesgos de mercado y procuren que éstos sean lo suficientemente precisos.

Esos incentivos consisten en premiar con un menor factor multiplicativo y, por lo tanto, con mayor holgura para cumplir el límite normativo (o con menores requerimientos de capital cuando se introduzcan cargos por riesgos de mercado) a aquellos bancos que tengan modelos de medición estadísticamente precisos (que acusen el menor número posible de excepciones, de modo que sean clasificados en la zona verde) y penalizar con factores multiplicativos mayores a aquellos bancos que tengan modelos que no sean estadísticamente precisos, o sea, que acusen un número mayor de excepciones, de modo que caigan en algún tramo de la zona amarilla o roja.

Aunque la normativa sobre riesgos de mercado sólo impone límites a los riesgos que los bancos pueden asumir, la aplicación de Backtesting también está concebida como una herramienta para estimular a los bancos a mantener modelos precisos; para lo cual se adopta la Tabla de Permanencia, con los correspondientes factores aditivos que se aplicarían en caso de que el banco no demuestre fehacientemente que la ocurrencia de las “excepciones” registradas no derivan de imprecisiones del modelo.

Este documento describe la metodología para construir la Tabla de Permanencia para cualquier tamaño de muestra y presenta un ejemplo para llevarla a la práctica. Con ello, se establece el referente metodológico para construir la Tabla de Permanencia para las entidades que requieran de cómputos específicos para los factores de multiplicación a ser utilizados para efectos de calcular el riesgo de mercado afecto a límite normativo.

Bibliografía

Basle Committee on Banking Supervision (1996), "Amendment to the Capital Accord to Incorporate Market Risks", Basilea.

Basel Committee on Banking Supervision (1996). "Supervisory Framework for the use of Backtesting in Conjunction with the Internal Models Approach to Market Risk Capital Requirements.", Basilea.

Markus Leippold, Paolo Vanini (2002). "Half as many cheers – the multiplier reviewed". Available at Social Science Reseach Network: http://ssrn.com/abstract=401881 or DOI: 10.2139/ssrn.401881

Stahl G., 1997, Three Cheers, Risk Magazine, Vol. 10, No. 5, 67-69.

Banco Central de Chile, Compendio de Normas Financieras, Capítulo III.B.2.

Superintendencia de Bancos e Instituciones Financieras, Recopilación Actualizada de Normas, Capítulo 12-9.

Anexos

Anexo 1.

Tabla A1. Tabla de Permanencia para una muestra de 500 días

Zona	Tabla de Permanencia: Número de Excepciones Registrados (x)	Tabla de Permanencia: Factor Aditivo (s_t)	Tabla de Permanencia: Factor Multiplicativo	$P(X \leq x)$
Verde	0	0	3	0,0066
	1			0,0398
	2			0,1234
	3			0,2636
	4			0,4396
	5			0,6160
	6			0,7629
	7			0,8677
	8			0,9329
Amarilla	9	0,33	3,33	0,9689
	10	0,40	3,40	0,9868
	11	0,47	3,47	0,9948
	12	0,53	3,53	0,9981
	13	0,59	3,59	0,9994
	14	0,65	3,65	0,9998
Roja	15	1	4	0,9999

Tabla A2

Modelo Exacto		
Excepciones	Cobertura = 99%	
n = 500	Exacta	Tipo I
0	0,5%	100,0%
1	2,8%	99,3%
2	7,4%	96,0%
3	12,9%	87,7%
4	16,8%	73,6%
5	17,6%	56,0%
6	15,2%	38,4%
7	11,3%	23,7%
8	7,3%	13,2%
9	4,2%	6,7%
10	2,2%	3,1%
11	1,0%	1,3%
12	0,4%	0,5%
13	0,2%	0,2%
14	0,1%	0,1%
15	0,0%	0,0%

Tabla A3

Modelo inadecuado: Alternativas posibles de niveles de cobertura								
Excepciones	Cobertura = 98%		Cobertura = 97%		Cobertura = 96%		Cobertura = 95%	
n = 500	Exacta	Tipo II	Exacta	Tipo II	Exacta	Tipo II	Exacta	Tipo II
0	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
1	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
2	0,2%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
3	0,7%	0,3%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
4	1,8%	1,0%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%
5	3,7%	2,8%	0,2%	0,1%	0,0%	0,0%	0,0%	0,0%
6	6,2%	6,5%	0,4%	0,3%	0,0%	0,0%	0,0%	0,0%
7	9,0%	12,8%	1,0%	0,7%	0,0%	0,0%	0,0%	0,0%
8	11,3%	21,7%	1,9%	1,7%	0,1%	0,1%	0,0%	0,0%
9	12,6%	33,1%	3,2%	3,5%	0,3%	0,2%	0,0%	0,0%
10	12,6%	45,7%	4,8%	6,7%	0,5%	0,4%	0,0%	0,0%
11	11,5%	58,3%	6,6%	11,5%	1,0%	1,0%	0,1%	0,0%
12	9,6%	69,8%	8,3%	18,1%	1,7%	2,0%	0,1%	0,1%
13	7,3%	79,3%	9,6%	26,4%	2,6%	3,6%	0,3%	0,3%
14	5,2%	86,7%	10,4%	36,0%	3,8%	6,2%	0,5%	0,6%
15	3,4%	91,9%	10,4%	46,4%	5,1%	10,0%	0,9%	1,1%

Las Tablas A2 y A3 reportan simultáneamente las probabilidades exactas de obtener un cierto número de excepciones de una muestra de 500 días independientes, bajo distintos niveles de cobertura, como asimismo las probabilidades de cometer un error del Tipo I o Tipo II, derivadas de dichas probabilidades exactas.

Anexo 2.

Pruebas estadísticas alternativas para evaluar el desempeño del VaR

Junto con el gráfico de Backtesting, existen pruebas estadísticas que permiten evaluar el desempeño del VaR. Una de estas pruebas está asociada a la variable aleatoria que cuenta el número de veces en las cuales las pérdidas exceden el VaR.

Si se asume que estos eventos son independientes, esta variable sigue una distribución Binomial con parámetros n y p, que corresponden al número de observaciones consideradas dentro de Backtesting y la probabilidad de falla (p = α), respectivamente. Por ejemplo, para el tamaño mínimo de días en el calculo de Backtesting que permite el Comité de Basilea, ésta corresponde a una distribución Binomial(250; 0.01). Dicha distribución es de la forma:

$$(1) \qquad P(X = x|250{,}0.01) = \binom{250}{x}(0.01)^{x}(0.99)^{250-x}$$

A partir de este resultado es posible realizar varias pruebas estadísticas para evaluar el desempeño del VaR, tal como se muestra a continuación.

a) Test de proporción de fallas (Kupiec, 1995)

La prueba de proporción de fallas de Kupiec, *POF* por sus siglas en inglés, evalúa la hipótesis nula que la probabilidad de falla sea igual a α. Es decir, si se calcula el VaR con un nivel de confianza del 99%, entonces H_0: 0.01. En términos prácticos, esta prueba estadística mide si el nivel de significancia propuesto por el VaR es consistente con la proporción de fallas que presenta el modelo.

Lo que hace el test es modelar diferencias entre los resultados reales y las medidas de riesgo generadas por el modelo VaR mediante una distribución Binomial. Si el VaR es inferior a los resultados reales, se define el evento como "fracaso" con probabilidad (p). En el caso opuesto, cuando las pérdidas cambiarias son inferiores al VaR, entonces se define ese evento como un "éxito" con probabilidad (1-p). La probabilidad de que el número de fracasos sea igual a "x" en una muestra de tamaño "n" se determina a partir de la distribución Binomial:

$$P(X = x|n,p) = \binom{n}{x} p^{x}(1-p)^{n-x}$$

La probabilidad de fracaso de los modelos VaR se estima a través del método de máxima verosimilitud (ML), cuya estimación consiste en tomar logaritmos a la distribución Binomial y maximizar dicha función con respecto a la probabilidad (p).

De este modo se debe maximizar la siguiente expresión:

$$\mathrm{Log}(\mathrm{P}) = \mathrm{Log}\begin{pmatrix} \mathrm{n} \\ \mathrm{x} \end{pmatrix} + \mathrm{x} \cdot \mathrm{Log}(\mathrm{p}) + (\mathrm{n} - \mathrm{x}) \cdot \mathrm{Log}(1 - \mathrm{p})$$

La condición de primer orden de la maximización es la siguiente:

$$\frac{\partial Log(P)}{\partial p} = x \cdot \frac{1}{p} - (n - x) \cdot \frac{1}{1 - p} = 0$$

Al simplificar la condición de primer orden, se obtiene que la probabilidad de fracaso del modelo VaR equivalente a la proporción de fallas del modelo (Porcentaje de veces que el VaR no predice las máximas pérdidas potenciales):

$$\hat{\mathrm{p}} = \frac{\mathrm{x}}{\mathrm{n}}$$

Una vez obtenida la estimación por máxima verosimilitud se debe establecer una comparación estadística entre la probabilidad teórica (p) y la probabilidad estimada ($\hat{\mathrm{p}}$), lo cual permitirá contrastar la hipótesis nula del test. Para esto se emplea una prueba de razón verosimilitud de la siguiente forma:

$$(2)\ \ \mathrm{LR}_{\mathrm{POF}} = -2\,\mathrm{Log}\left(\frac{\mathrm{p}^{\mathrm{x}}(1-\mathrm{p})^{\mathrm{n}-\mathrm{x}}}{\hat{\mathrm{p}}^{\mathrm{x}}(1-\hat{\mathrm{p}})^{\mathrm{n}-\mathrm{x}}} \right)$$

Donde “x” es el número de excepciones o fallas, “n” el número de observaciones incluidas en el *Backtesting* y $\hat{\mathrm{p}} = \frac{\mathrm{x}}{\mathrm{n}}$. El numerador de (2) corresponde al valor de la función de verosimilitud bajo la hipótesis nula, mientras que el denominador corresponde a la función de verosimilitud evaluada en el estimador ML no restringido de p, $\hat{\mathrm{p}}$. La estadística de esta prueba se distribuye asintóticamente Chi-cuadrado con un grado de libertad.

Utilizando el modelo Binomial, comentado en el test anterior, el estimador de máxima verosimilitud de p, $\hat{\mathrm{p}}$, está definido por:

$$(3)\quad \hat{\mathrm{p}} = \frac{\mathrm{x}}{\mathrm{n}}$$

Y el estimador de la varianza es:

$$(4)\quad \hat{\mathrm{V}}(\mathrm{p}) = \frac{\hat{\mathrm{p}}(1-\hat{\mathrm{p}})}{\mathrm{n}}$$

Teniendo en cuenta que (3) es un estimador de máxima verosimilitud, éste sigue asintóticamente una distribución normal. Por lo tanto, es posible establecer intervalos de confianza para p, de la forma: $\left[\hat{p} \pm z_{\alpha/2}\sqrt{\hat{V}(p)}\right]$, con el objeto de evaluar si p = α se encuentra dentro del intervalo. El resultado anterior se puede obtener análogamente a través de la hipótesis nula H_0: p = α, mediante el uso de la siguiente estadística:

(5) $$Z_0 = \frac{\hat{p} - (1-\alpha)}{\sqrt{\hat{V}(p)}}.$$

Asintóticamente, la distribución de la estadística presentada en (5) es Normal estándar.

b) Estimación directa a partir de la distribución Binomial.

Esta prueba también se basa en la ecuación (1) pero, a diferencia de las pruebas anteriores, tiene la ventaja de no depender de resultados asintóticos.

La idea detrás de esta prueba es construir un intervalo de confianza $(1-\tilde{\alpha})\cdot 100\%$ del número de fallas bajo H_0: p = α utilizando la distribución Binomial. Si el número de fallas observado, x, se encuentra dentro de este intervalo la hipótesis nula no es rechazada con un nivel de significancia $\tilde{\alpha}\cdot 100\%$.