



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

FACULTAD DE CIENCIAS

ALTERNATIVAS ESTADÍSTICAS PARA
EL CÁLCULO DEL VALOR EN RIESGO VaR

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

P R E S E N T A

ANDRÉS LUNA ARIAS

Tutor: Dr. Manuel Mendoza Ramírez

2006



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1. Datos del alumno
Luna
Arias
Andrés
53 92 04 28
Universidad Nacional Autónoma de México
Facultad de Ciencias
Actuaría
071132480
2. Datos del tutor
Dr.
Manuel
Mendoza
Ramírez
3. Datos del sinodal 1
Mat.
Margarita Elvira
Chavéz
Cano
4. Datos del sinodal 2
M. en C.
Jesús Agustín
Cano
Garcés
5. Datos del sinodal 3
Act.
Roberto
Canovas
Theriot
6. Datos del sinodal 4
Dr.
Rubén
Hernández
Cid
7. Datos del trabajo escrito
Alternativas Estadísticas para el Cálculo del Valor en Riesgo VaR
106 p
2006

AGRADECIMIENTOS

La matemática es la ciencia más exacta,
y sus conclusiones son susceptibles de
demostración absoluta. Pero ese se
debe exclusivamente a que la
matemática no intenta obtener
conclusiones absolutas. Todas las
verdades matemáticas son relativas,
condicionales

Charles Proteus Steinmetz (1923)

Este trabajo lo dedico especialmente a mi
Madre Imelda Arias, quien pese a todo tuvo la
visión de que sus hijos estudiaran.

A mis hermanos Jesús y Alberto

A mis sobrinos: Andrea, Gaby, Aurora y Betito

A los amigos y compañeros con los que he
compartido parte de la vida.

La verdad no puede explicarse de modo que se comprenda y no se crea

William Blake

| | |
|---|-----|
| RESUMEN..... | I |
| 1. INTRODUCCIÓN..... | 1 |
| 1.1 ADMINISTRACIÓN INTEGRAL DE RIESGOS..... | 7 |
| 1.2 Riesgo Financiero..... | 9 |
| 1.2.1 Clasificación Del Riesgo..... | 9 |
| 1.3 Tratamientos Del Riesgo De Mercado..... | 10 |
| 2. EL VALOR EN RIESGO..... | 13 |
| 2.1 Manejo de los datos..... | 13 |
| 2.2. Definición De VaR..... | 14 |
| 3. ENFOQUES BÁSICOS DEL CÁLCULO DEL VAR..... | 20 |
| 3.1 Los Modelos Riskmetrics y GARCH..... | 21 |
| 3.2 El Estimador Por Promedios Móviles Ponderados Exponencialmente (EWMA)..... | 22 |
| 3.3 Los Modelos De Heteroscedasticidad Autorregresiva Condicionada ARCH y GARCH..... | 28 |
| 3.3.1.a La función de verosimilitud..... | 39 |
| 4. LOS MÉTODOS NO-PARAMÉTRICOS..... | 49 |
| 4.1 Simulación Histórica..... | 49 |
| 4.2 Simulación por Monte Carlo..... | 53 |
| 4.2.1 Muestreo de Importancia Basado en la Aproximación Delta-gamma | 59 |
| 4.2.2 Muestreo Estratificado..... | 64 |
| 5. MODELOS ALTERNATIVOS..... | 66 |
| 5.1 El Déficit Esperado | 66 |
| 5.2 Medidas de Riesgo y Teoría del Valor Extremo (EVT) | 72 |
| 5.1.1 Distribución de los máximos por bloques..... | 74 |
| 5.1.2 Distribución de máximos sobre un umbral..... | 81 |
| 6. OPTIMIZACIÓN DEL PORTAFOLIO..... | 90 |
| 7. CONCLUSIONES..... | 101 |

Alternativas estadísticas al cálculo del Valor en Riesgo

RESUMEN

Durante la última década se han desarrollado nuevos métodos de medida y de gestión del riesgo de mercado. Una de estas medidas, conocida como *Valor en Riesgo* (VaR), ha cobrado especial importancia y actualmente es el patrón que siguen las instituciones financieras para el control de su riesgo de mercado. El enfoque VaR es atractivo porque es fácil de interpretar (el VaR está medido en unidades monetarias) y puede ser utilizado tanto para efectos de auditoría y control, como para producir una estimación de la cantidad necesaria de fondos propios, para cubrir las pérdidas asociadas al riesgo de mercado, de las actividades de negociación desarrolladas por las entidades financieras. Además, tiene la ventaja adicional de poder incorporar los efectos de la diversificación de las carteras.

A la consolidación de este enfoque como herramienta de gestión del riesgo de mercado, ha contribuido el grupo J. P. Morgan al hacer pública en 1994 la descripción de su sistema de medida del riesgo de mercado, denominado *RiskMetrics*, basado en la metodología VaR, incluyendo el conjunto de datos necesarios para su aplicación. También ha contribuido a la consolidación de estos sistemas el hecho de que las autoridades internacionales en supervisión bancaria (*Comité de Basilea y Unión Europea*), hayan sentado los acuerdos y bases estadísticas, y permitir a las entidades financieras la posibilidad de determinar la cantidad de fondos propios, necesarios para cubrir el riesgo de mercado de sus carteras de negociación, mediante modelos internos basados en la metodología VaR.

Las entidades supervisoras sugieren especificaciones estadísticas concretas para el cálculo del VaR; sin embargo, las instituciones financieras no están obligadas a utilizar ningún procedimiento concreto para estimar el VaR. Son las propias entidades las que deciden el procedimiento estadístico en que basarán sus estimaciones. Esta libertad de elección tiene sentido, ya que como indica Beder (1995) no existe hasta el momento un determinado enfoque VaR que sea superior al resto.

El objetivo principal de este trabajo es presentar algunas de las metodologías utilizadas en el cálculo del Valor en Riesgo y otras medidas alternativas; sus principales supuestos, ventajas y desventajas. Pretendemos mostrar estas metodologías, profundizando y ejemplificando algunas de ellas y describiendo de manera general otras, de manera que constituyan un compendio de las alternativas estadísticas que son utilizadas en la práctica.

En la parte introductoria se sitúa el riesgo de mercado dentro del contexto general de la administración de riesgos. El resto del documento se compone de la siguiente manera: en la segunda sección se definen algunos conceptos básicos para el manejo de los datos financieros y el valor en riesgo. En la tercera sección, se presentan las metodologías básicas de estimación paramétrica, así como sus características. En la cuarta sección se

presentan los principales resultados de los métodos no paramétricos. En la quinta sección se introducen algunos métodos alternativos junto con sus aplicaciones en la definición de medidas de riesgo. Posteriormente, en la sexta sección se presenta un método de optimización. Finalmente, se presentan las conclusiones.

1. INTRODUCCIÓN

La esencia de la actividad de las instituciones financieras es la administración de riesgos. Es decir, el riesgo es un componente inevitable en la operación, en la toma de decisiones de inversión de sus activos y en la toma de decisiones en general, en la búsqueda de maximizar sus utilidades, asumiendo ciertos niveles de riesgo que tienen una relación directa con la productividad y la generación eficiente de utilidades de la institución.

Administrar estos riesgos de una manera eficiente e integral es fundamental, ya que la correcta identificación, medición, monitoreo y control de los riesgos permite a las instituciones optimizar el rendimiento de su capital, ajustado por el nivel de riesgo, además de optimizar las decisiones relativas a su operación, prevenir pérdidas y proteger el capital.

La globalización de los mercados, así como la desregulación de las economías, son tendencias mundiales que han hecho evidente la necesidad de que las instituciones cuenten con herramientas que les permitan identificar, medir y monitorear los riesgos a los que están expuestas. A raíz de la apertura de los mercados, del aumento del volumen de transacciones y de la volatilidad de las mismas, el análisis de estrategias de administración de riesgos se hace indispensable, considerando alternativas que les permitan protegerse contra los riesgos financieros o especular con ellos.

Las experiencias sufridas en diferentes partes del mundo, donde se han tenido pérdidas que han llevado a la quiebra a diversas instituciones bancarias y de otro tipo, muestran de acuerdo a Philippe Jorion un elemento común: la deficiente administración de los riesgos financieros, especialmente los relacionados con el mercado y el crédito.

Dichos quebrantos, aunados a la búsqueda de un sistema financiero seguro y confiable fueron las bases para la instrumentación de recomendaciones y medidas que determinan el capital necesario para proteger los recursos frente al riesgo de mercado, que incluyen conceptos y peculiaridades de los sistemas supervisores y contables de cada país.

Así en 1988, los banqueros centrales del grupo de los diez (compuesto por Bélgica, Canadá, Francia, Alemania, Italia, Japón, Holanda, Suecia, Inglaterra y los Estados Unidos de América, además de Luxemburgo y Suiza), emitieron los acuerdos de Basilea, cuyo propósito principal era proporcionar a los bancos comerciales de los países miembros, un mínimo de requerimientos de capital, para cubrir los riesgos de crédito, del 8% del total de los activos con riesgo de crédito del banco.

Además, de los requerimientos de capital, dichos acuerdos establecen límites a la “toma excesiva de riesgos”, es decir, se trata de poner restricciones sobre grandes riesgos, definidos como posiciones que exceden el 10% del capital de un banco y no permitiendo las posiciones que excedan el 25% del capital de una empresa.

Posteriormente en 1993, se emite el primer paquete de propuestas, de carácter consultivo, en torno al riesgo de mercado, donde el valor en riesgo se calcula para los portafolios expuestos al riesgo de las tasas de interés, al riesgo cambiario, al riesgo de

otros activos y al riesgo accionario, utilizando directrices específicas, donde, el valor en riesgo total del banco se obtiene como la suma de los diferentes valores en riesgo de acuerdo a los cuatro riesgos señalados, lo que se ha denominado el modelo estándar.

Finalmente en 1996, el Comité de Basilea bajo los auspicios del Banco Internacional de Pagos y con los resultados de la consulta, emite una enmienda al acuerdo sobre requerimientos de capital de 1988, la cual adiciona cargos de capital para cubrir el riesgo de mercado, a los cargos que cubren el riesgo de crédito. Dicha propuesta cubre los riesgos de pérdidas de posiciones, dentro y fuera de las cuentas de balance, que se tengan por movimientos en los precios de mercado.

El conjunto propuesto contiene, cuatro líneas de acción para medir el riesgo del mercado y el cálculo del cargo de capital para cubrir este riesgo, con dos metodologías, la estándar y la interna de cada banco, propuesta que de manera resumida se presenta a continuación.

Medidas de riesgo del mercado

El Riesgo del mercado puede medirse usando la metodología estándar o modelos internos de los bancos (sujeto a la aprobación del supervisor nacional) como sigue:

La metodología estándar

Este método usa un procedimiento para posiciones de deuda y acciones donde el riesgo específico y el riesgo general son calculados separadamente. El cargo de capital bajo este método será la suma aritmética de las medidas de cada riesgo de mercado.

La deuda

Los instrumentos cubiertos incluyen: los pagares (e instrumentos que se comportan de igual manera, incluso acciones preferentes no convertibles) y tasas de interés, derivadas de la cuenta comercial. Las posiciones equivalentes en asuntos idénticos (por ejemplo, mismo emisor, tasas cupón, liquidez, compra de futuros), así también, como swaps, forwards y futuros los cuales adicionan condiciones que deben ser compensadas. El cargo de capital para la deuda es la suma del cargo de riesgo específico y del cargo de riesgo general.

- **Riesgo Específico**

El cargo de riesgo específico está diseñado para la protección contra movimientos adversos en el precio de una emisión de deuda particular relacionados a factores del emisor individual. Las deudas y derivados garantizados son clasificados en amplias categorías (gobierno, calificado, y otros) con un cargo de capital aplicado a las posiciones en cada categoría. El rango del cargo de capital va del 0% para la categoría gubernamental hasta el 8% para los otros.

- El riesgo General

El cargo de riesgo general esta diseñado para capturar el riesgo de pérdidas que provienen de los cambios en las tasas de interés del mercado. Un cargo por riesgo general se calculará separadamente para cada posición significativa del banco, teniendo dos métodos a elegir:

1. El método de madurez

Las posiciones largas y cortas en deuda y derivados garantizados son incluidas de acuerdo a su madurez en 13 bandas de tiempo (o vencimiento). La posición neta en cada banda de tiempo es el riesgo ponderado por un factor diseñado para reflejar la sensibilidad del precio de las posiciones a los cambios en las tasas de interés.

2. El método de duración

Logra una medida más exacta de riesgo del mercado general al calcular la sensibilidad del precio de cada posición separadamente.

El cargo de riesgo general es la suma del riesgo-ponderado de posiciones cortas y largas en el libro comercial, una proporción pequeña de las posiciones equivalentes en cada banda de tiempo (la agregación vertical 10% para el método de madurez; 5% para el método de duración), y una proporción más grande de las posiciones equivalentes que cruzan diferentes bandas de tiempo (la agregación horizontal).

Acciones

Los instrumentos cubiertos incluyen: las acciones ordinarias, coberturas convertibles que se comportan como las acciones, los compromisos de compra o venta de acciones y derivados de acciones. Pueden compensarse posiciones equivalentes con cada acción en cada mercado, mientras produzcan un solo corto. El cargo de capital para las acciones es la suma de los cargo de riesgo específico y de riesgo general.

- El riesgo específico

El riesgo específico es el riesgo del sostener una posición larga o corta en una acción particular, es decir, la posesión absoluta de acciones del banco (la suma de todas las posiciones cortas y largas de acciones). El cargo de riesgo específico es del 8% (o 4% si el portafolio es líquido y diverso). Un cargo de riesgo específico del 2% se aplicará a la posición corta o larga neta en un índice que comprende un portafolio diversificado de acciones.

- El riesgo general

El riesgo general es el riesgo del sostener una posición larga o corta en el mercado como un todo, es decir, la diferencia de la suma de los largos y la suma de los cortos (la posición neta global en un mercado accionario). El cargo de riesgo de mercado general es del 8% y es calculado en base al valor de mercado.

El riesgo del tipo de cambio (incluido el oro)

El método para calcular el requerimiento de capital para el riesgo del tipo de cambio, es realizado midiendo la posición neta en cada divisa y oro a la tasa spot, aplicando un cargo de capital del 8% a la posición abierta neta (es decir, el mayor de la posición neta larga o corta en divisas y el 8% de la posición en oro).

Riesgo de Otros Activos

El riesgo de otros activos, incluyendo los metales preciosos, sin el oro, puede medirse usando el proceso estándar o el proceso simplificado para bancos que manejan sólo una cantidad limitada de otros activos. Bajo el procedimiento estándar, las posiciones de cortos, largos y futuros en cada bien son escaladas de acuerdo a su madurez. El cargo de capital se calculará aplicando una tasa del 1.5% a las posiciones equivalentes y un cargo de capital aplicado a la posición neta en cada escalón. Bajo el método simplificado, un cargo de capital del 15% se aplicará a la posición neta de cada bien.

El tratamiento de opciones

Los bancos que solamente compren opciones, se les permite usar un procedimiento simplificado; sin embargo, los bancos que también suscriban opciones, deben usar un procedimiento intermedio o un modelo de administración de riesgo. Bajo el proceso estándar, las opciones deben ser aisladas y estar sujetas a cálculos separados de los cargos de capital sobre comercialización particular y ser sumados a otros cargos de capital. Los procedimientos intermedios son el “delta más” y el análisis de escenarios. Bajo el proceso “delta más”, las opciones delta-ponderadas serán incluidas en la metodología estándar para cada tipo de riesgo.

La metodología alternativa: los modelos internos

Se permite a los bancos usar medidas de riesgo derivadas de su propia administración de riesgo, sujetas a un conjunto general de normas y condiciones. La aprobación por la autoridad de supervisión sólo se concederá, si hay un número suficiente de personas (incluidas las áreas comercial, de control de riesgo, auditoría y áreas de apoyo) experimentadas en el uso de los modelos, el modelo habrá de producir un registro más exacto de las predicciones de pérdidas y el banco incluirá pruebas bajo escenarios extremos (stress tests).

- El Cálculo del cargo de capital bajo el modelo interno
 - Cada banco deberá registrar sobre una base diaria los requerimientos de capital, expresados como el valor en riesgo más alto del día anterior, medido según los parámetros especificados o como un promedio del valor en riesgo diario de los sesenta días de operación anteriores, multiplicado por el factor correspondiente.
 - El factor será propuesto por los supervisores en base a la valoración del sistema de administración de riesgo del banco, teniendo un valor mínimo de 3 e incrementos entre 0 a 1, en base a los resultados del backtesting, teniendo

un incremento extra igual a cero aquellos bancos que cumplan satisfactoriamente todas las normas cualitativas.

- Los bancos que usen sus modelos estarán sujetos a un cargo de capital separado para cubrir el riesgo específico de deuda comercializada y de las acciones que no están incorporadas en sus modelos. Sin embargo, los bancos que usen modelos, para el cargo de riesgo específico aplicado a deuda o acciones, deberán tener un cargo mayor o igual a la mitad del cargo de riesgo específico calculado bajo la metodología estándar.
 - Cualquier elemento de riesgo de mercado no contemplado por el modelo interno deberá permanecer sujeto al procedimiento de medición estándar.
 - Los cargos de capital evaluados bajo el proceso estándar y el modelo interno se agregarán según el método de la suma simple.
- Los requisitos para el uso de modelos internos:

Las normas cualitativas

- La existencia de una unidad independiente de control de riesgos con participación activa en la alta dirección.
- El modelo debe integrarse en la administración del día con día del riesgo y debe usarse junto como parte de la operación interna y los límites de exposición al riesgo.
- Deben realizarse rutinariamente rigurosos programas de pruebas bajo escenarios extremos y backtesting.
- Una revisión independiente para asegurar el cumplimiento de la administración y medición del riesgo debe llevarse a cabo en intervalos de tiempo regulares.
- Los procedimientos deben ser descritos para la aprobación interna y externa del proceso de medida de riesgo

La especificación de factores de riesgo de mercado

Los factores de riesgo contenidos en un sistema de medida de riesgo deben ser suficientes para capturar el riesgo inherente en el portafolio del banco, es decir, tasas de interés, tipo de cambio, precios de acciones y precios de otros activos.

○ *Las normas cuantitativas*

- El Valor en riesgo debe ser calculado diariamente usando un percentil al 99%, de confianza sobre la cola y un periodo de tenencia mínimo de 10 días de operación. Se permite a los bancos escalar, el VaR de un día, para opciones para un cierto periodo de tiempo.

- El periodo de observaciones históricas estará sujeto a un periodo mínimo de un año. Para bancos que usen esquemas ponderados u otros métodos para el periodo de observaciones históricas, el periodo efectivo de las observaciones debe ser por lo menos un año.
- Los bancos tendrán la posibilidad de reconocer las correlaciones empíricas entre las categorías de riesgo. El uso de estimaciones de la correlación por categoría de riesgo están sujetas a la aprobación de la metodología de estimación usada.
- Los Bancos deben actualizar sus conjuntos de datos por lo menos una vez cada trimestre.
- Los modelos deben capturar los riesgos asociados con las opciones dentro de las categorías de riesgo, con precisión (usando los factores delta/gamma si el proceso analítico es seleccionado)

El cálculo de capital

- El porcentaje mínimo de capital disponible para cubrir los riesgos de crédito y de mercado es del 8%.
- El denominador del porcentaje es calculado multiplicando la medida de riesgo del mercado por 12.5 (recíproco del 8%) y sumando los resultados al riesgo ponderado de crédito de los recursos. El numerador es el capital elegible, es decir, la suma del capital de primer nivel del banco, el capital de segundo nivel, bajo los límites permitidos por el acuerdo de 1988 y el capital de tercer nivel, consistente de la deuda de corto plazo. Se permitirá usar el capital de tercer nivel con el propósito de reunir los requerimientos de capital para los riesgos de mercado, pero sujeto a ciertas limitaciones cuantitativas.
- Si bien los reportes regulares deben, en principio, tener lugar sólo en intervalos de tiempo (en la mayoría de los países de forma trimestral), se espera de los bancos, un manejo del riesgo de mercado de sus portafolios, en que los requerimientos de capital se reúnan en una base continua, es decir, al cierre diario de los negocios.

El papel de supervisión (el uso de backtesting)

Backtesting representa la comparación de las pérdidas y ganancias diarias con los pronósticos del modelo de riesgo para calibrar la calidad y exactitud de los sistemas de medición del riesgo de los bancos. El backtesting debe ser aplicado para comprobar, que el porcentaje observado de resultados cubiertos por la medida de riesgo, es consistente con el nivel de confianza (99%). El proceso backtesting debe usar medidas de riesgo calibradas sobre un día de tenencia. El Comité insta a los bancos a desarrollar la capacidad para realizar backtests usando resultados hipotéticos (basados en los cambios en el valor del portafolio ocurridos cuando al final del día las posiciones permanecieron inalteradas) y los resultados comerciales reales.

El procedimiento adoptado por el Comité calcula el número de veces que los resultados comerciales no se cubren por las medidas de riesgo (excepciones) en una base trimestral con los datos de los últimos 12 meses. El proceso abarca un rango de posibles respuestas que son clasificadas en 3 zonas. Los límites se basan en una muestra de 250 observaciones.

- La zona verde

Los resultados del backtesting no sugieren ningún problema con la calidad o exactitud del modelo del banco (sólo se permiten cuatro excepciones).

- La zona Amarilla

Los resultados del backtesting plantean dudas, pero no existe una conclusión definitiva (sólo se permiten 9 excepciones). Los resultados en este rango son creíbles para los modelos exactos e inexactos. El número de excepciones definirá el número de supervisiones. El propósito del incremento del factor de la multiplicación es que el modelo cubra el nivel de confianza del 99%. Los resultados del backtesting en la zona amarilla generalmente implicaran un incremento en el factor de la multiplicación a menos que el banco pueda demostrar que tal aumento no se garantiza.

- La zona Roja

El resultado del backtesting indica un problema con el modelo de riesgo del banco (10 o más excepciones). Si el modelo del banco cae en esta zona, el supervisor aumentará el factor de la multiplicación automáticamente en un punto y empezará la investigación.

1.1 Administración Integral De Riesgos

Dentro del país y en concordancia con los acuerdos de Basilea, La Comisión Nacional Bancaria y de Valores con fundamento en lo dispuesto por los artículos 97 de la Ley de Instituciones de Crédito, así como el 4, fracciones II y VI, 6, 16, fracción I y 19 de la Ley de la Comisión Nacional Bancaria y de Valores, emitió las disposiciones de carácter prudencial en materia de administración integral de riesgos aplicables a las instituciones de crédito, con las cuales se establecen las bases de la administración de riesgos en nuestro país.

El objetivo principal de la administración de riesgos es asegurar que las actividades de operación e inversión de una institución, no se expongan a pérdidas que puedan amenazar la viabilidad futura de la misma. La administración integral de riesgos es, en última instancia, responsabilidad de la dirección y del consejo de administración de la institución, ya que es a éstos a quienes corresponde determinar la dirección estratégica a tomar, así como la tolerancia al riesgo que están dispuestos a asumir.

La administración integral de riesgos, los modelos de medición de riesgo y sus sistemas de control interno deben de corresponder a las actividades propias de cada institución. Así mismo es importante evaluar los supuestos cuantitativos y cualitativos implícitos en el sistema, así como el grado de efectividad de la institución para controlar los mismos, los

reportes de información gerencial y otras formas de comunicación, con el propósito de que éstos satisfagan los objetivos y sean adecuados para el nivel de negocios de la institución.

La alta dirección y el consejo de administración deben tener un flujo de información continuo que les permita estar al tanto de los niveles de riesgo que está asumiendo la institución. Adicionalmente debe llevar a cabo la reevaluación periódica del grado de exposición aceptable para la institución, con relación al manejo y medición de riesgos, así como el cumplimiento de los límites establecidos y mantener la existencia de controles internos funcionales y un proceso extensivo de reportes y análisis de riesgos.

La administración integral de riesgos debe incluir al menos los siguientes puntos:

- La identificación y valuación de los distintos tipos de riesgos.
- El establecimiento de políticas, procedimientos y límites de riesgo.
- Monitoreo y reporte del cumplimiento de los límites establecidos.
- Delineación del capital asignado y de la administración del portafolio.
- Guías para el desarrollo de nuevos productos y la inclusión de nuevas exposiciones al riesgo dentro de la estructura existente.
- Aplicación de nuevos métodos de medición a los productos existentes.

Todos los riesgos mayores deben ser medidos explícita y consistentemente e integrarse a un sistema de medición integral de riesgos de la institución. Los sistemas y procedimientos deben reconocer que la medición de riesgos en muchos casos es una aproximación sujeta a variaciones por factores económicos y de mercado. Una práctica sana de medición de riesgo es mantener identificados de manera continua los cambios en las condiciones del mercado que pueden afectar negativamente el valor de los activos de la institución.

La evolución hacia mercados globales, así como la evolución de los instrumentos financieros, hace cada vez más difícil la medición de riesgos, no obstante, existen metodologías y procesos de análisis, algunos de los cuáles son ya de uso común, que permiten identificar los riesgos de ciertas actividades de negocios y agruparlos de manera genérica, dando como resultado una sola medida por tipo de riesgo. Además permiten que las instituciones midan su riesgo con base en las características específicas de sus carteras tomando en cuenta las consideraciones pertinentes con relación al perfil de riesgo y la estrategia global de la institución.

Un sistema integral de administración de riesgos debe considerar, además de lo mencionado en párrafos anteriores, lo siguiente:

- Tomar en cuenta de manera integral los distintos tipos de riesgo a los que la institución está expuesta
- Ser consistente con las recomendaciones formuladas internacional y nacionalmente
- Permitir a las instituciones identificar, medir, administrar y monitorear los distintos tipos de riesgos de una manera integral.

1.2 Riesgo Financiero

Los riesgos derivados de una inversión formada por activos financieros se ven reflejados en el hecho de que eventos inesperados afectan la estructura patrimonial del inversionista, provocando un cambio negativo tanto en sus activos como en sus obligaciones. El tratamiento de estos fenómenos consiste en identificar las exposiciones, es decir, determinar qué clase de riesgos afectan a las inversiones; establecer algún tipo de medida con la que puedan ser cuantificados; y, por último, idear una estrategia mediante la cual estos riesgos puedan ser controlados o atenuados.

Existen diversas clases de riesgos que afectan a las inversiones realizadas, como se verá a continuación:

1.2.1 Clasificación Del Riesgo

De acuerdo con las Disposiciones de Carácter Prudencial en Materia de Administración Integral de Riesgos Aplicables a las Instituciones de Crédito, emitidas por la Comisión Nacional Bancaria y de Valores el riesgo puede clasificarse de la siguiente manera:

- I) Riesgos cuantificables, son aquellos para los cuales es posible conformar bases estadísticas que permitan medir sus pérdidas potenciales, y dentro de éstos, se encuentran los siguientes:
 - a) Riesgos discrecionales, que son aquellos resultantes de la toma de una posición de riesgo, tales como el:
 - 1) Riesgo de crédito o crediticio, que se define como la pérdida potencial por la falta de pago de un acreditado o contraparte en las operaciones que efectúan las instituciones, incluyendo las garantías reales o personales que les otorguen, así como cualquier otro mecanismo de mitigación utilizado por las instituciones de crédito.
 - 2) Riesgo de liquidez, que se define como la pérdida potencial por la imposibilidad o dificultad de renovar pasivos o de contratar otros en condiciones normales para la institución, por la venta anticipada o forzada de activos a descuentos inusuales para hacer frente a sus obligaciones, o bien, por el hecho de que una posición no pueda ser oportunamente enajenada, adquirida o cubierta mediante el establecimiento de una posición contraria equivalente.
 - 3) Riesgo de mercado, que se define como la pérdida potencial por cambios en los factores de riesgo que inciden sobre la valuación o sobre los resultados esperados de las operaciones activas, pasivas o causantes de pasivo contingente, tales como tasas de interés, tipos de cambio e índices de precios, entre otros.
 - b) Riesgos no discrecionales, que son aquellos resultantes de la operación del negocio, pero que no son producto de la toma de una posición de riesgo, tales como el riesgo operativo que se define como la pérdida potencial por fallas o

deficiencias en los controles internos, por errores en el procesamiento y almacenamiento de las operaciones o en la transmisión de información, así como por resoluciones administrativas y judiciales adversas, fraudes o robos y comprende, entre otros, al riesgo tecnológico y al riesgo legal, en el entendido de que:

- 1) El riesgo tecnológico, se define como la pérdida potencial por daños, interrupción, alteración o fallas derivadas del uso o dependencia en el hardware, software, sistemas, aplicaciones, redes y cualquier otro canal de distribución de información en la prestación de servicios bancarios con los clientes de la institución.
 - 2) El riesgo legal, se define como la pérdida potencial por el incumplimiento de las disposiciones legales y administrativas aplicables, la emisión de resoluciones administrativas y judiciales desfavorables y la aplicación de sanciones, en relación con las operaciones que las instituciones llevan a cabo.
- II) Riesgos no cuantificables, que son aquellos derivados de eventos imprevistos para los cuales no se puede conformar una base estadística que permita medir las pérdidas potenciales.

El trabajo que aquí se presenta está enfocado al riesgo de mercado, dentro del cual confluyen distintos tipos de riesgo tal y como lo señala el acuerdo de Basilea.

1.3 Tratamientos Del Riesgo De Mercado

En la práctica, el tratamiento del riesgo de mercado referente a las inversiones financieras es y ha sido llevado a cabo mediante diversas alternativas. El objetivo, en esta sección, es mencionar brevemente los medios más importantes con los que los inversionistas han enfrentado y enfrentan esta clase de riesgo.

Diversificación de Carteras.

Uno de los medios más convencionales para tratar el riesgo de mercado es la conformación de un portafolio bien diversificado. Esto significa la elección adecuada de las proporciones en las que participará cada recurso financiero en la inversión. La elección de estas proporciones se realiza, de tal manera, que se tenga la expectativa de que ante escenarios adversos en las fluctuaciones de determinados precios correspondientes a recursos en posesión, existan posiciones en otros recursos que puedan compensar el impacto de dicho escenario. Esto significa, por ejemplo, mantener recursos cuyos precios tengan correlaciones negativas, es decir que las variaciones de sus precios tengan sentidos opuestos.

Bajo estas condiciones, las características particulares relativas de los recursos, deben ser tales que, las expectativas de variación de sus precios puedan compensarse, de modo que el riesgo de mercado quede atenuado.

El principal inconveniente de este método es el uso de expectativas de comportamiento de los precios de los recursos en cuestión. Lo importante es tomar conciencia que éstas no siempre se cumplen. Por lo tanto, es vital tener conocimiento de que una errónea formación de expectativas podría implicar consecuencias adversas en la estructura patrimonial del inversionista.

Opciones sobre Índices Bursátiles.

En determinados mercados se han desarrollado opciones de venta cuyo bien subyacente ya no es un recurso en particular sino un portafolio de recursos estrechamente relacionado con algún índice de mercado. Esto es, se considera la evolución del valor de este índice como la evolución del precio de una acción. Dichas opciones tienen la particularidad de garantizar un valor mínimo (vinculado con el precio de ejercicio) al índice considerado, en una fecha determinada.

Este tipo de instrumento es muy beneficioso en los casos en que la composición del portafolio a cubrir sea similar a la composición del índice considerado como bien subyacente. Por otro lado, puede presentarse un inconveniente, cuando la composición, y por lo tanto, la evolución del portafolio utilizado como referencia difiere de la composición del portafolio a asegurar, lo que implica una divergencia en los comportamientos de ambos y, consecuentemente, una deficiencia en la cobertura.

Otro inconveniente se presenta en los casos en los que el precio de ejercicio, según las características estandarizadas del mercado, no coincide con el valor mínimo aceptable que el inversionista desea para el portafolio en un horizonte futuro.

Reproducción de Opciones sobre Portafolios.

Dadas las dificultades e inconvenientes planteados para el caso de las Opciones sobre Índices Bursátiles, se han logrado técnicas para reproducir este tipo de instrumento mediante la formación de un portafolio compuesto por acciones y bonos. Dicho portafolio, que reproduce sintéticamente la evolución de la opción en cuestión, debe ser re-balanceado periódicamente para poder ajustarse a ésta según las condiciones sucesivas del mercado.

Esta ingeniosa idea permite obtener opciones sobre portafolio con la composición deseada y precio de ejercicio que coincida con el valor mínimo aceptable para la inversión en un momento futuro especificado.

Los inconvenientes de este método están relacionados con el re-balanceo del portafolio que reproduce la opción. Esto se da ya que las operaciones de re-balanceo implican costos de transacción, lo que puede ser un factor que incremente el costo total de esta cobertura.

Por otro lado, existe un problema relacionado con la determinación de la frecuencia con la que el portafolio debe ser re-balanceado.

Valor en Riesgo e Inmovilización de Capital.

La propuesta de Comité de Basilea señala que a partir del cálculo del Valor en Riesgo, las entidades financieras deben inmovilizar capital para utilizarlo como reserva, a fin de amortiguar los impactos derivados de los escenarios adversos en los mercados.

La medición del Valor en Riesgo debe llevarse a cabo utilizando varios horizontes de tiempo que pueden ir desde el muy corto plazo (un día), hasta el largo plazo (un año), tomando en cuenta que en mercados en crisis cerrar una posición o instaurar una cobertura a precios razonables puede ser muy difícil. Lo que significa mantener una posición por más tiempo de lo deseado, lo cual incrementa el riesgo, ya que la posibilidad de cambios de gran magnitud se incrementa con el tiempo y el perfil de riesgo de algunos instrumentos, como las opciones, varía substancialmente entre más se acerca la fecha de vencimiento.

En México existe un consenso de que la valuación de los portafolios de inversión debe hacerse a precios de mercado, debido a que el valor de mercado es el mejor parámetro de referencia para la valuación del riesgo de los portafolios de inversión.

El procedimiento para realizar el análisis de riesgo de mercado, de los portafolios de inversión de las instituciones podría tipificarse de la siguiente manera:

- Determinación de las características de cada uno de los instrumentos de inversión para su valuación, tales como el valor nominal, el plazo, la tasa cupón, el número de cupones por vencer, el subyacente, la fecha de vencimiento, etc.
- Obtención de la información necesaria para la valuación del portafolio a partir de distintas fuentes de información, las cuáles pueden ser públicas o privadas.
- Análisis y selección del o los modelos a utilizar para la valuación de cada instrumento del portafolio.
- Determinación de los factores de riesgo y su relación con los modelos de valuación de tal forma que se puedan identificar tanto sus tendencias y volatilidades, ya sea para su proyección a futuro o para incorporar dicha información en la simulación de escenarios.
- Definición y obtención de las curvas de tasas de rendimiento para la valuación de instrumentos.
- Alimentación del sistema a partir de las bases de datos que contengan toda la información disponible.
- Respetar los criterios nacionales (dictados por la Comisión Nacional Bancaria y de Valores) e internacionales (Recomendaciones del Comité de Basilea) a fin de obtener resultados confiables y oportunos.

A la consolidación de este enfoque como herramienta de gestión del riesgo de mercado, ha contribuido de forma notable el grupo J.P. Morgan al hacer pública en 1994 la descripción de su sistema de medición del riesgo de mercado, denominado *RiskMetrics*, basado en la metodología VaR, así como el conjunto de datos necesarios para su aplicación. También, han sido importantes las contribuciones del Premio Nóbel de economía Robert F Engle y de Tim Bollerslev, quienes han hecho aplicable la metodología del VaR, a través del estudio y desarrollo de modelos autoregresivos.

2. EL VALOR EN RIESGO

2.1 Manejo de los datos

La mayoría de estudios financieros se realizan sobre rendimientos y no sobre precios. Por lo tanto, antes de definir la medida de riesgo utilizada que se presenta en este trabajo, es importante definir algunos conceptos básicos relacionados con los rendimientos.

En general, el cambio absoluto de un activo financiero está definido por el cambio en su precio durante un período de tiempo, $G_t = P_t - P_{t-1}$, sin embargo, esta definición, no es muy adecuada para la medición de riesgo, toda vez que el comparar estos cambios absolutos sólo se refieren al volumen monetario.

En términos de comparación, una medida más atractiva para los resultados de un activo son los rendimientos, ya que en estos los cambios se relativizan al valor inicial del activo, en este sentido, los *rendimientos simples* de un activo, para un período están dados por:

$$R_t = (P_t - P_{t-1}) / P_{t-1}.$$

Entonces el rendimiento bruto de un activo financiero será $1 + R_t$ de donde tenemos que los rendimientos de un activo de manera continua, que se denominan los rendimientos compuestos-continuos o logarítmicos están definidos como:

$$r_t = \ln(1 + R_t) = \ln(1 + (P_t - P_{t-1}) / P_{t-1}) = \ln(P_t / P_{t-1}) = p_t - p_{t-1}$$

donde $\ln(P_t) = p_t$. La definición de rendimiento logarítmico posee ventajas sobre los rendimientos simples, aunque sus valores se asemejan cuando las variaciones en los precios son pequeñas, tal y como se observa en el cuadro 1.

La definición de R_t y r_t describen los rendimientos para un periodo, mismas que se usaran para el cálculo del rendimiento multiperiodo.

Recordando nuestros cursos de matemáticas financieras, se tiene que una unidad con un crecimiento R_t , $t = 1, 2, \dots, k$, se comporta como sigue:

$$\begin{aligned} f(0) &= 1; \quad f(1) = 1 + R_1; \quad f(2) = 1 + R_1 + R_2(1 + R_1) = (1 + R_1)(1 + R_2), \\ f(3) &= (1 + R_1)(1 + R_2) + (1 + R_1)(1 + R_2)R_3 = (1 + R_1)(1 + R_2)(1 + R_3) \\ f(k) &= (1 + R_1)(1 + R_2) \dots (1 + R_k) = (P_1 / P_0)(P_2 / P_1) \dots (P_k / P_{k-1}) = (P_k / P_0) \end{aligned}$$

Entonces el rendimiento para el periodo k , considerando el periodo inicial en t , será igual a:

$$1 + R_{t,k} = (P_{t+k} / P_t) \text{ de donde}$$

$$R_{t,k} = ((P_{t+k} - P_t) / P_t)$$

De igual manera el rendimiento logarítmico será igual a:

$$r_{t,k} = \ln(1 + R_{t,k}) = p_{t+k} - p_t$$

Cuadro 1: Cambio de precios Absoluto, relativo y logarítmico del tipo de cambio de pesos por dólar

| Fecha | Pesos por dólar P_t | Cambio Absoluto de precios x 100 | Cambio de precios relativo (%) | Cambio logaritmo de precios (%) |
|------------|--------------------------|-------------------------------------|-----------------------------------|------------------------------------|
| 01/12/2005 | 10.5095 | -4.9000 | -0.4641 | -0.4652 |
| 02/12/2005 | 10.4640 | -4.5500 | -0.4329 | -0.4339 |
| 05/12/2005 | 10.4750 | 1.1000 | 0.1051 | 0.1051 |
| 06/12/2005 | 10.4265 | -4.8500 | -0.4630 | -0.4641 |
| 07/12/2005 | 10.4660 | 3.9500 | 0.3788 | 0.3781 |
| 08/12/2005 | 10.5120 | 4.6000 | 0.4395 | 0.4386 |
| 09/12/2005 | 10.6485 | 13.6500 | 1.2985 | 1.2902 |
| 13/12/2005 | 10.6180 | -3.0500 | -0.2864 | -0.2868 |
| 14/12/2005 | 10.7125 | 9.4500 | 0.8900 | 0.8861 |
| 15/12/2005 | 10.7320 | 1.9500 | 0.1820 | 0.1819 |
| 16/12/2005 | 10.7410 | 0.9000 | 0.0839 | 0.0838 |
| 19/12/2005 | 10.7505 | 0.9500 | 0.0884 | 0.0884 |
| 20/12/2005 | 10.6840 | -6.6500 | -0.6186 | -0.6205 |
| 21/12/2005 | 10.6270 | -5.7000 | -0.5335 | -0.5349 |
| 22/12/2005 | 10.6600 | 3.3000 | 0.3105 | 0.3100 |
| 23/12/2005 | 10.6900 | 3.0000 | 0.2814 | 0.2810 |
| 26/12/2005 | 10.6980 | 0.8000 | 0.0748 | 0.0748 |
| 27/12/2005 | 10.7515 | 5.3500 | 0.5001 | 0.4988 |
| 28/12/2005 | 10.7780 | 2.6500 | 0.2465 | 0.2462 |
| 29/12/2005 | 10.6650 | -11.3000 | -1.0484 | -1.0540 |

2.2. Definición De VaR

El valor en riesgo es una medida del máximo cambio potencial en el valor de un portafolio de instrumentos financieros con una probabilidad dada sobre un horizonte predeterminado. VaR responde la pregunta: ¿cuánto puedo perder con (1-c)% de probabilidad sobre un horizonte de tiempo dado?

Ejemplo 1:

Supongamos que un inversionista de Estados Unidos, tiene un activo por 140 millones de pesos. ¿Cuál es el VaR sobre un horizonte de un día, si se fija un 5% de probabilidad de que la pérdida comprendida sea mayor que lo que VaR proyectó?.

El primer paso es calcular la exposición en riesgo (es decir, valorar en términos de mercado el activo). Como la inversión es de un inversionista de los EE UU, entonces la exposición es igual al valor del mercado del activo en la base del inversionista. Si el tipo de cambio es de 10.40 PESOS/DÓLAR, el valor del mercado del activo se calcula como:

$$V_0 = 140,000,000 / 10.40 = 13.461 \text{ millones de dólares.}$$

La exposición al riesgo esta sobre el tipo de cambio, el cual requiere una estimación de cuánto puede moverse el tipo de cambio. La desviación estándar de los rendimientos del tipo de cambio PESOS/DÓLAR, medido históricamente, puede brindar una indicación del tamaño del movimiento de la tasa de cambio. En este ejemplo, considere una desviación estándar diaria de 0.565%. Entonces, bajo el supuesto de que los rendimientos se distribuyen como una normal $(0, \sigma^2)$ y dado el valor de la desviación estándar, un rendimiento negativo se calcula como -1.65 veces la desviación estándar, porque el cuantil al 5% de una distribución $N(0, 1)$ es -1.65, lo que es igual a -0.932%. (ver gráfica 1)

Entonces un valor negativo (al 95% de confianza) del activo se puede calcular como el valor inicial por el rendimiento negativo, y suponiendo que los rendimientos son compuestos-continuos, tenemos que:

$$V_1 = V_0 e^r = 13.461 * \exp(-0.00932) = 13.461 * 0.9907233 = 13.336$$

De donde el VaR de la posición es igual a ese valor menos el valor inicial, esto es

$$\text{VaR} = V_1 - V_0 = 13.336 - 13.461 = -0.125 \text{ millones de dólares}$$

Una manera alternativa de calcular el VaR, es $V_1 - V_0 = V_0 (\exp(r)-1)$, donde r es la tasa instantánea de rendimiento, se sabe que $\exp(r) = 1 + i$ y si se supone que $r = i$ entonces $\exp(r) \approx (1+r)$, de donde $V_0 (\exp(r)-1) \approx V_0 ((1+r)-1) = V_0 r$.

Entonces el VaR de la posición sería igual a 13.461 millones por $-0.00932 = -125,456$

Ejemplo 2:

Ahora suponga que se tienen los 140 millones de pesos en un bono gubernamental Mexicano a 10 años. ¿Cuál es el VaR sobre un horizonte de un día, si se considera un 5% de probabilidad de que la pérdida comprendida sea mayor que lo que VaR proyectó?

La única diferencia con el ejemplo anterior es que ahora aunado al riesgo cambiario se tiene el riesgo de la tasa de interés del bono. El valor del activo sigue siendo de 13.461 millones de dólares, pero ahora se atiende a dos factores de riesgo de mercado.

Si se usa una estimación de la desviación estándar de la tasa de interés del bono de 0.605%, entonces se puede calcular el riesgo de tasas de interés como:

$$13.461 \text{ millones por } -1.65 \text{ por } 0.605\% = -134,374.5, \text{ mientras que el riesgo del tipo de cambio es de } 13.461 \text{ millones por } -1.65 \text{ por } 0.565\% = -125,456$$

De manera lógica, se podría pensar que el riesgo total de la inversión es la suma de ambos riesgos, sin embargo, en general, ello no es verdad, por lo siguiente:

Sea $Z = Fx + Bn$, donde Fx y Bn son las variables aleatorias, de los rendimientos del tipo de cambio y de la tasa de interés del bono, respectivamente, entonces la varianza de esta variable es igual a:

$$\text{Var}(Z) = \text{Var}(Fx + Bn) = E(Fx+Bn - E(Fx+Bn))^2 = E(Fx-E(Fx) + Bn-E(Bn))^2 = E(Fx-E(Fx))^2 + E(Bn-E(Bn))^2 + 2(Fx-E(Fx))(Bn-E(Bn)) = \text{Var}(Fx) + \text{Var}(Bn) + 2 \text{Cov}(Fx, Bn)$$

Además si se supone que los rendimientos por tipo de cambio y tasa de interés, se distribuyen como normales $N(0, \sigma_1^2)$ y $N(0, \sigma_2^2)$ respectivamente, entonces la variable aleatoria Z se distribuye como una normal $(0, \sigma_1^2 + \sigma_2^2 + 2Cov(Fx, Bn))$. Suponga en este ejemplo, que se tiene una correlación entre el rendimiento del tipo de cambio de PESOS/DÓLAR y el rendimiento del bono a 10 años de -0.27 .

Entonces el VaR se calcula como:

$$\begin{aligned} \text{VaR} &= 13.461(-1.65)\sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho_{1,2}\sigma_1\sigma_2} \\ &= -22.21065\sqrt{(0.00565)^2 + (0.00605)^2 + 2(-0.27)(0.00565)(0.00605)} \\ &= -157,157 \text{ dólares} \end{aligned}$$

De manera general si queremos calcular el Valor en riesgo de un portafolio sobre un horizonte de 1 día con una probabilidad del 5% de que la pérdida real en el valor del portafolio sea mayor que la estimación de VaR, se siguen los siguientes pasos:

1. Valuar el portafolio actual a precios de mercado. V_0
2. Hacer una estimación del rendimiento de 1 día del portafolio (\tilde{r}), con una probabilidad del 5% de que el rendimiento actual sea menor que el rendimiento estimado. Esto es: $\Pr(r < \tilde{r}) = 5\%$.
3. Calcular el valor futuro del portafolio, como $V_1 = V_0 \exp(\tilde{r})$. De donde, la estimación del Valor en Riesgo simplemente es $V_1 - V_0 = V_0(\exp(\tilde{r}) - 1)$ y como $\exp(\tilde{r}) \approx 1 + \tilde{r}$, entonces VaR es aproximadamente igual a $V_0\tilde{r}$.

De manera general para cualquier valor de t , podemos definir probabilísticamente el VaR, en un horizonte temporal $[t, t+1]$ con una probabilidad establecida c , lo que también se conoce como el nivel de confianza, de la siguiente forma.

Sea V_t el valor inicial del portafolio (usualmente V_t representa su valor presente) y sea V_{t+1} el valor del portafolio transcurrido el horizonte temporal en consideración (un día, una semana, un mes etc.). La definición implícita del VaR en términos probabilísticos es:

$$\Pr(V_{t+1} - V_t < -\text{VaR}) = 1 - c \quad (1)$$

donde el signo negativo se usa porque generalmente el VaR se considera positivo. Si se supone, como en los ejemplos anteriores, que los rendimientos se distribuyen como una normal $(0, \sigma^2)$, entonces, se puede encontrar una expresión para el VaR como sigue, suponemos que $V_{t+1} = V_t \exp(r_{t+1})$ y se sabe que si r es la tasa instantánea de crecimiento, $\exp(r) = 1+i$, por lo que si hacemos $r = i$, se tiene que $\exp(r) \cong 1+r$

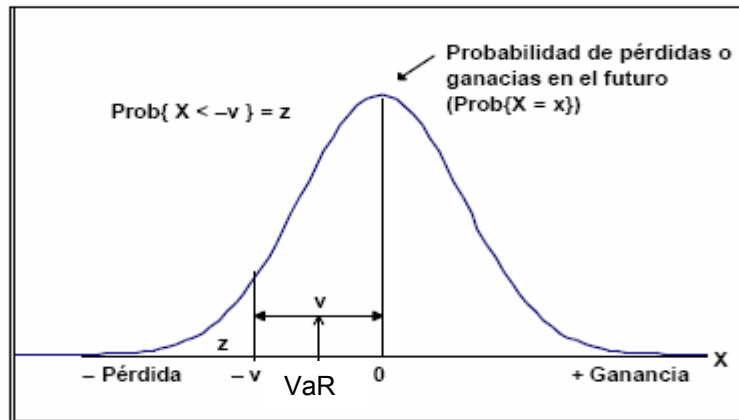
$$\begin{aligned} \Pr(V_t(\exp(r_{t+1})-1) < -\text{VaR}) &\cong \Pr(V_t(1+r_{t+1})-1) < -\text{VaR}) = \Pr(1+r_{t+1}-1 < -\text{VaR}/V_t) \\ &= \Pr(r_{t+1} < -\text{VaR}/V_t) = 1-c \end{aligned} \quad (2)$$

Si llamamos q_{1-c} al percentil $(1-c)$ de la variable rendimiento del portafolio después del horizonte temporal, entonces (2) nos conduce a la siguiente expresión para VaR

$$\text{VaR} \cong -V_t(q_{1-c}) = -V_t\sigma c_{1-c} \quad \text{donde } c_{1-c} \text{ es el cuantil al } 1-c \% \text{ de una normal } (0,1) \quad (3)$$

por lo que en términos estadísticos el VaR se corresponde con el percentil σ_{1-c} de la distribución de pérdidas y ganancias del portafolio y en este sentido la vocación de VaR es *predictiva*, ya que es una estimación sobre el valor futuro del portafolio transcurrido el horizonte temporal.

Gráfica 1: Representación Gráfica del Concepto del Valor en Riesgo



Las ecuaciones anteriores asignan propiedades probabilísticas al VaR, sin embargo, también resultan interesantes sus posibles propiedades frecuentistas, ya que éstas pueden servir como herramienta de gestión a mediano y largo plazo. Analizando (1) podemos intuir que si calculamos el VaR sobre un alto número de periodos (moviendo t), la proporción de periodos donde el cambio del valor del portafolio no sea superior al VaR será aproximadamente $(1-c)$. Pero debemos considerar que esta propiedad frecuentista del VaR *sólo* es cierta bajo la hipótesis de independencia entre los periodos, e igualdad de condiciones, lo que generalmente no se cumple.

Debido a que el VaR es un cálculo estadístico de las pérdidas potenciales de un portafolio, independientemente de cuál sea el método utilizado para su estimación, su valor depende de la elección del periodo de mantenimiento del portafolio (*holding period*) y del nivel de confianza c seleccionado.

El periodo de mantenimiento es el horizonte temporal para el que se quiere la estimación del VaR, con lo que implícitamente se supone que las posiciones del portafolio se mantienen constantes durante dicho periodo. Los periodos habitualmente utilizados oscilan entre un día y un mes, aunque se puede utilizar otros periodos más largos como trimestres e incluso años. El *Comité de Basilea* establece un periodo de dos semanas (diez días hábiles), aunque en la práctica se utilizan periodos de un día, siendo posible trabajar con periodos inferiores a éste.

En la elección del periodo de mantenimiento deben considerarse diversos aspectos, como la liquidez de los mercados en los que opera la institución o la posibilidad de que la entidad reajuste las posiciones del portafolio. En el primer caso, el periodo de mantenimiento ideal sería el tiempo necesario para asegurar la liquidación de las posiciones en ese mercado. En cuanto al segundo aspecto, si el periodo de mantenimiento es demasiado largo, es probable que se realicen cambios en la

composición del portafolio, de manera que estos cambios hacen que la medida del VaR sea irreal, dado que se ha violado el supuesto de condiciones iguales. Por tanto, cuanto más corto sea el periodo de mantenimiento, más real será el supuesto de que el portafolio se mantiene inalterado durante dicho periodo.

La elección del nivel de confianza (c) dependerá del uso se quiera dar al cálculo del VaR:

Si la medida del VaR se utiliza para determinar los fondos requeridos, los parámetros del VaR, incluido el nivel de confianza, vienen determinados por quien regula la actividad financiera. Por ejemplo, el Comité de Basilea establece que el cálculo del VaR debe realizarse diariamente sobre un horizonte de planificación de dos semanas (diez días de negociación), utilizando un nivel de confianza del 99% ($c=0.99$) y un periodo mínimo de observación de un año.

En el caso de que el VaR se utilice como sistema interno para la gestión de riesgos, el nivel de confianza dependerá básicamente de la aversión al riesgo propuesto por la administración.

Resumiendo podemos decir que el VaR, presenta las siguientes ventajas y desventajas:

Ventajas del VaR

- El propósito primario de VaR es estimar el riesgo de mercado.
- VaR puede usarse como un reporte para notificar a la Dirección y al Consejo de Administración de los riesgos involucrados en operaciones, en un punto dado del tiempo.
- VaR puede ser usado para asignar recursos que cubran el riesgo entre comercialización o estrategia.
- Los valores obtenidos a través del VaR, permiten un mejor desempeño. Por ejemplo, si una institución gana 2.5 millones arriesgando sólo 100,000 por día, mientras que otra institución gana 3 millones por día arriesgando 1 millón, la primera institución tiene mejor desempeño. El desempeño se evalúa como los ingresos por peso arriesgado en lugar de sólo las ganancias.

Desventajas del VaR

- VaR no indica el máximo que una compañía puede perder. Sólo el máximo probable que una compañía puede perder bajo las condiciones actuales de mercado. VaR no mide volumen de riesgo o estabilidad financiera.
- VaR no informa acerca de la rentabilidad o rendimientos sobre la inversión. Sin embargo, puede ser un componente para determinar el valor ajustado por riesgo, del portafolio.
- VaR es sólo una estimación estadística de las pérdidas y no un resultado exacto de cuanto se pierde o se debe cubrir con capital propio.
- VaR no puede predecir la peor pérdida, ni acumula pérdidas.
- VaR no es una medida que abarque todo el desempeño de la institución, sólo tiene significado si la Dirección ha establecido una tolerancia de riesgo de acuerdo con las pautas de administración

Hasta el momento se han propuesto en la literatura numerosos métodos para resolver aproximadamente (3), aunque la inmensa mayoría de éstos, son variantes de tratamientos estadísticos clásicos y tratamientos estadísticos no paramétricos.

3. ENFOQUES BÁSICOS DEL CÁLCULO DEL VaR

Mientras que VaR es un concepto fácil e intuitivo, medirlo es un problema estadístico muy desafiante. Aunque los modelos existentes para VaR emplean diferentes metodologías, todos siguen una estructura general común, que puede resumirse en tres puntos: 1) Valorar el portafolio a precios de mercado, 2) Estimar la distribución de los rendimientos del portafolio, 3) Calcular el VaR del portafolio.

Las principales diferencias entre los modelos de VaR se relacionan al punto 2, que es la manera de cómo estimar los posibles cambios en el valor del portafolio. Dichos modelos se pueden clasificar en dos categorías¹:

- Paramétricos (RiskMetrics, familias ARCH y GARCH, Teoría del Valor Extremo, Déficit Esperado, métodos robustos y cuasi-máxima verosimilitud GARCH)
- No-paramétricos (la Simulación Histórica, El Modelo Híbrido y la Simulación de Monte Carlo)

Los resultados que cada uno de estos métodos ofrece, pueden ser muy diferentes, Beder [9] aplicó ocho metodologías de VaR² a tres portafolios hipotéticos, cuyos resultados mostraron diferencias muy grandes, con estimaciones de VaR que varían por más de 14 veces para el mismo portafolio. Entonces, para decidir qué metodología escoger, es necesario entender los supuestos subyacentes, así como los modelos matemáticos y técnicas cuantitativas que usan. Sólo después de este paso preliminar, se estará en posibilidad de escoger el modelo que se piense, puede ser más apropiado a los objetivos.

La base de las metodologías VaR viene de las características de los datos financieros. Los hechos empíricos sobre los mercados financieros están descritos y especificados, en los trabajos de Mandelbrot [51] y Fama [32] quienes con amplias explicaciones, detectan, entre otras, las siguientes características:

1. Las distribuciones de los rendimientos financieros son leptokurticas, presentan colas más pesadas y crestas más altas que la distribución normal.
2. Los rendimientos de las acciones son típicamente sesgados negativamente.
3. Los rendimientos al cuadrado presentan autocorrelación significativa, es decir, los factores de volatilidad del mercado tienden a agruparse. Ésta es una característica muy importante de los rendimientos financieros, que permite considerar las volatilidades del mercado como cuasi-estables, cambiantes a la larga, pero estables en el corto plazo. La mayoría de los modelos VaR hace uso de esta cuasi-estabilidad para evaluar el riesgo del mercado.

¹ El número y tipo de planteamientos para la estimación de VaR están creciendo y es imposible tener en cuenta todos ellos. En particular, la prueba de extremos no se discutirá en este trabajo. El lector interesado puede consultar el sitio web www.gloriamundi.org para profundizar sobre las contribuciones al VaR.

² El Déficit Esperado y la Teoría del Valor Extremo no son considerados por Beder, porque son métodos posteriores a su estudio.

3.1 Los Modelos Riskmetrics y GARCH

Modelos como RiskMetrics, ARCH y GARCH, que veremos con más detalle en la sección 3.2 y 3.3, proponen una parametrización específica para el comportamiento de los rendimientos y en el caso de los dos últimos de la varianza o volatilidad. La familia de modelos ARCH la desarrollo Engle [27] y posteriormente Bollerslev [11] los generalizó y se han aplicado con éxito a datos financieros (vea Bollerslev, Engle y Nelson [12] para profundizar más sobre este aspecto). A manera de ejemplo un modelo GARCH(1,1), para datos financieros, se puede describir como sigue:

$$\begin{aligned} r_t &= \sigma_t \varepsilon_t & \varepsilon_t &\sim i.i.d N(0,1) \\ \sigma_t^2 &= w + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \quad (4)$$

Este modelo tiene dos elementos cruciales: la particular especificación de la ecuación de la varianza y el supuesto de que los residuales estandarizados (ε_t) son independientes e idénticamente distribuidos (i.i.d) como una $N(0,1)$. La ecuación de la varianza es consecuencia de las características de los datos financieros, señaladas anteriormente, en tanto que el supuesto de i.i.d. de las (ε_t) y su distribución, es un dispositivo necesario para estimar los parámetros desconocidos. La distribución generalmente usada es la normal estándar³. Sólo después de que este supuesto sobre la distribución se ha impuesto, es posible escribir una función de verosimilitud y conseguir una estimación de los parámetros desconocidos.

Bajo el modelo de RiskMetrics la varianza es estimada usando Promedios Móviles Ponderados Exponencialmente (EWMA), es decir un filtro de promedios móviles, pero con ponderaciones decrecientes, y que corresponde a un modelo GARCH Integrado (IGARCH) de la forma $\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1-\lambda) r_{t-1}^2$, RiskMetrics también asume que los residuales son normalmente distribuidos.

Hasta ahora, sólo nos hemos referido a un portafolio compuesto por un solo recurso, si el portafolio está formado por N activos, el valor de σ_t se corresponde con una combinación de las varianzas y covarianzas de los rendimientos de los activos de dicho portafolio (ver ejemplo 2). Sea $R_t = (r_1^t, r_2^t \dots r_N^t)$ el vector formado por los rendimientos individuales de cada recurso, y $\mathbf{w} = (w_1, w_2, \dots, w_N)$ el vector del porcentaje de participación de cada instrumento dentro del portafolio, de manera que la suma de los w_i es igual a 1, tenemos que r_t es la media ponderada con los pesos w_i de los rendimientos, esto es

$$r_t = \sum_{i=1}^N w_i r_i^t$$

de donde, si calculamos la varianza de la expresión anterior tenemos

³ Este supuesto es también motivado por el importante resultado presentado por Bollerslev y Woolridge [13] sobre el GARCH cuasi-maximo verosímil

$$\sigma_t^2 = \text{Var}(r_t) = \text{Var}\left(\sum_{i=1}^N w_i r_i^t\right) = \sum_{i=1}^N w_i^2 \sigma_{i,t}^2 + 2 \sum_{i=1}^N w_i w_j \sigma_{ij,t}$$

donde $\sigma_{i,t}^2$ es la varianza de r_i y σ_{ijt} la covarianza entre r_i y r_j matricialmente, lo podemos escribir como

$$\sigma_t^2 = \mathbf{w}' \Omega_t \mathbf{w}, \quad \text{donde } \Omega_t \text{ es la matriz de varianzas-covarianzas de los rendimientos.}$$

Cuando el portafolio incluye una gran cantidad de activos o instrumentos más complejos, es necesario recurrir a técnicas que permitan reducir el número de factores de riesgo. La solución habitual es reducir la dimensionalidad de Ω_t como lo propone la metodología de Riskmetrics, que utiliza un sistema denominado "mapping" o cartografía de flujos de cajas para portafolios que incorporan instrumentos de renta fija, como los bonos. De esta forma, los activos de renta fija son descompuestos en flujos de caja en función de la corriente futura de pagos que generan dichos activos. Una vez descompuestos los activos en flujos de caja, estos deben ser asignados ("cartografiados") a una serie de vértices, que se corresponden con determinados vencimientos de la estructura temporal mediante activos de referencia para los que se dispone de datos. De esta forma los distintos instrumentos pueden ser descritos mediante combinaciones de tales activos de referencia.

Asimismo Navarro y Nave [58] y [59] desarrollaron un modelo bifactorial de duraciones, basado en el trabajo de Elton, Gruber y Michaely (EGM) [25], donde se asume que el comportamiento de la Estructura Temporal de los Tipos de Interés (ETTI) puede ser explicado mediante un número reducido de variables clave. En EGM se asume que las variables clave son algunos tipos de interés al contado que permiten aproximar las variaciones no esperadas en cualquier tipo de interés.

Las fórmulas planteadas hasta ahora expresan la relación entre el VaR y los parámetros del modelo (básicamente Ω_t) que evidentemente deben ser estimados pues son desconocidos.

Estimar la matriz de varianzas-covarianzas es un proceso complicado, debido al gran número de factores que intervienen en la misma. El principal problema es la inestabilidad de la volatilidad en periodos cortos de tiempo, ya que se observa que los precios pueden tener movimientos muy bruscos, cuando el mercado reacciona abruptamente.

Entre los métodos de estimación de la matriz de varianzas y covarianzas más utilizados está el de Promedios Móviles Ponderados Exponencialmente (EWMA por sus siglas en inglés), utilizado por JPMorgan en su sistema RiskMetrics [57] y los modelos de heteroscedasticidad condicional autorregresiva condicionada ARCH y GARCH.

3.2 El Estimador Por Promedios Móviles Ponderados Exponencialmente (EWMA)

Una clase ampliamente de modelos que describen la evolución de los rendimientos esta basado en la noción de que el logaritmo de los precios financieros se modela como una caminata aleatoria (random walk).

El modelo para describir el logaritmo de los precios del recurso es de la siguiente forma:

$$p_t = \mu + p_{t-1} + \sigma \varepsilon_t \quad \varepsilon_t \sim \text{iid } N(0,1), \quad t = 1, 2, \dots,$$

Dicha ecuación postula la evolución del logaritmo de los precios y su distribución, en un punto en el tiempo, donde el logaritmo del precio actual depende de un parámetro fijo μ , del logaritmo del precio de periodo anterior, y una variable aleatoria normalmente distribuida multiplicada por el parámetro σ .

Al modelar el logaritmo de los precios, se tiene un modelo para los rendimientos, es decir, $r_t = \mu + \sigma \varepsilon_t$. De donde, podemos derivar una expresión para los precios, de la siguiente forma

Este modelo del logaritmo de los precios, supone una varianza constante, es decir σ , no cambian con el tiempo, lo cual de acuerdo con estudios empíricos es poco realista, por lo que en dicho modelo de caminata aleatoria se incluye la varianza variante en el tiempo como sigue:

$$p_t = \mu + p_{t-1} + \sigma_t \varepsilon_t \quad \varepsilon_t \sim \text{iid } N(0,1) \text{ ó}$$

$$r_t = \mu + \sigma_t \varepsilon_t \quad \varepsilon_t \sim \text{iid } N(0,1)$$

Los supuestos de la metodología de Riskmetrics, sobre el modelo que describe el comportamiento de los rendimientos se pueden resumir de la siguiente manera:

- Los rendimientos presentan varianza heterocedastica, autocorrelacionada y cambiante en el tiempo
- Las covarianzas de los rendimientos son autocorrelacionadas
- El que los rendimientos sean normalmente distribuidos significa que:
 - Sólo la media y la varianza son necesarios para describir la forma de la distribución
 - La suma de diferentes rendimientos normales tiene una distribución normal, lo que facilita la descripción de los rendimientos del portafolio, los cuales son sumas de rendimientos subyacentes
- La media de la distribución de los rendimientos es igual a cero lo que implica que $\mu = 0$

Con estos supuestos, si consideramos N recursos, los rendimientos se generan de acuerdo a un modelo de caminata aleatoria de la forma siguiente:

$$r_{i,t} = \sigma_{i,t} \varepsilon_{i,t} \quad \varepsilon_{i,t} \sim N(0,1) \quad i=1, \dots, N$$

$$\varepsilon_t \sim NMV(0, \Omega_t) \quad \varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{Nt}]$$

Donde Ω_t es la matriz NxN de varianzas y covarianzas, esto es, es la matriz donde la varianza de cada rendimiento y las covarianzas de los rendimientos son dependientes del tiempo.

Entonces bajo el supuesto de que los rendimientos tienen una distribución normal, lo que tenemos que encontrar es un estimador de la varianza σ_i^2 y de las covarianzas de los rendimientos a fin de tener la estimación de la matriz Ω_t .

Dado que el modelo de los rendimientos requiere re-estimar la varianza en cada periodo, podemos suponer que en el periodo T, se puede estimar la varianza en este periodo, a partir de la estimación de la varianza y del cuadrado del rendimiento del periodo T-1. Una manera de obtener el estimador es modificar el anterior estimador más alguna fracción ($0 < \lambda < 1$) de la diferencia del rendimiento al cuadrado menos el estimador de la varianza en el anterior periodo, esto es, para cualquier $i = 1, \dots, N$ se tiene

$$\begin{aligned}\hat{\sigma}_{i,T}^2 &= \hat{\sigma}_{i,T-1}^2 + (1-\lambda) [r_{i,T-1}^2 - \hat{\sigma}_{i,T-1}^2] \\ &= (1-\lambda)r_{i,T-1}^2 + \lambda\hat{\sigma}_{i,T-1}^2 \\ &= (1-\lambda)r_{i,T-1}^2 + (\lambda) [(1-\lambda)r_{i,T-2}^2 + \lambda\hat{\sigma}_{i,T-2}^2] \\ &= (1-\lambda)r_{i,T-1}^2 + \lambda(1-\lambda)r_{i,T-2}^2 + (\lambda)^2\hat{\sigma}_{i,T-2}^2 \\ &= (1-\lambda)\sum_{k=1}^T (\lambda)^{k-1} r_{i,T-k}^2 + (\lambda)^T \hat{\sigma}_{i,0}^2\end{aligned}$$

Para un valor grande de T, el valor de $(\lambda)^T \hat{\sigma}_{i,0}^2$ lo podemos considerar cero y entonces tenemos que

$$\hat{\sigma}_{i,T}^2 = (1-\lambda)\sum_{k=1}^T \lambda^{k-1} r_{i,T-k}^2$$

Este método de estimación permite obtener información acerca de la varianza de una serie histórica de observaciones, asignando distintas ponderaciones a las mismas, de tal manera que las observaciones recientes tienen mayor peso. Este estimador tiene dos ventajas importantes respecto al estimador de promedios móviles simple

$$F_{i,T}^2 = \frac{1}{T} \sum_{k=1}^T r_{i,T-k}^2$$

Primero incorpora el hecho de que la varianza reacciona más rápido ante cambios ocurridos en el pasado cercano del mercado, ya que asigna mayor peso al pasado reciente que al pasado más lejano. Segundo, cuando una observación o serie de observaciones muestran un cambio brusco, la varianza disminuye exponencialmente conforme la ponderación de esa o esas observaciones disminuye. En contraste, el uso de promedios móviles simple, ocasiona que cuando una observación o serie de observaciones muestran un cambio brusco el promedio móvil no registra este efecto de inmediato dado que asigna la misma ponderación a todas las observaciones, tal y como se observa en el cuadro siguiente:

| Fecha | Precio | Rendimiento | Rendimiento Al cuadrado (A) | Ponderación simple (B) | Ponderación Exponencial $\lambda = 0.94$ (C) | Promedio simple (AxB) | Promedio ponderado (AxC) |
|---|---------|-------------|-----------------------------|------------------------|--|-----------------------|--------------------------|
| 01/12/05 | 10.5095 | -0.4652 | 0.2164 | 0.05 | 0.0185 | 0.0108 | 0.0040 |
| 02/12/05 | 10.4640 | -0.4339 | 0.1883 | 0.05 | 0.0197 | 0.0094 | 0.0037 |
| 05/12/05 | 10.4750 | 0.1051 | 0.0110 | 0.05 | 0.0210 | 0.0006 | 0.0002 |
| 06/12/05 | 10.4265 | -0.4641 | 0.2154 | 0.05 | 0.0223 | 0.0108 | 0.0048 |
| 07/12/05 | 10.4660 | 0.3781 | 0.1430 | 0.05 | 0.0237 | 0.0071 | 0.0034 |
| 08/12/05 | 10.5120 | 0.4386 | 0.1923 | 0.05 | 0.0252 | 0.0096 | 0.0049 |
| 09/12/05 | 10.6485 | 1.2902 | 1.6645 | 0.05 | 0.0268 | 0.0832 | 0.0447 |
| 13/12/05 | 10.6180 | -0.2868 | 0.0823 | 0.05 | 0.0286 | 0.0041 | 0.0023 |
| 14/12/05 | 10.7125 | 0.8861 | 0.7851 | 0.05 | 0.0304 | 0.0393 | 0.0238 |
| 15/12/05 | 10.7320 | 0.1819 | 0.0331 | 0.05 | 0.0323 | 0.0017 | 0.0011 |
| 16/12/05 | 10.7410 | 0.0838 | 0.0070 | 0.05 | 0.0344 | 0.0004 | 0.0002 |
| 19/12/05 | 10.7505 | 0.0884 | 0.0078 | 0.05 | 0.0366 | 0.0004 | 0.0003 |
| 20/12/05 | 10.6840 | -0.6205 | 0.3850 | 0.05 | 0.0389 | 0.0193 | 0.0150 |
| 21/12/05 | 10.6270 | -0.5349 | 0.2862 | 0.05 | 0.0414 | 0.0143 | 0.0118 |
| 22/12/05 | 10.6600 | 0.3100 | 0.0961 | 0.05 | 0.0440 | 0.0048 | 0.0042 |
| 23/12/05 | 10.6900 | 0.2810 | 0.0790 | 0.05 | 0.0468 | 0.0039 | 0.0037 |
| 26/12/05 | 10.6980 | 0.0748 | 0.0056 | 0.05 | 0.0498 | 0.0003 | 0.0003 |
| 27/12/05 | 10.7515 | 0.4988 | 0.2488 | 0.05 | 0.0530 | 0.0124 | 0.0132 |
| 28/12/05 | 10.7780 | 0.2462 | 0.0606 | 0.05 | 0.0564 | 0.0030 | 0.0034 |
| 29/12/05 | 10.6650 | -1.0540 | 1.1108 | 0.05 | 0.0600 | 0.0555 | 0.0667 |
| Estimación de la desviación estándar | | | | | | | |
| Promedios móviles simples | | | | | | 0.5394 | |
| Promedios móviles exponencialmente ponderados | | | | | | | 0.4602 |

Otra característica de este estimador es que puede expresarse de manera recursiva, lo cual facilita el cálculo de las estimaciones de la varianza, así el pronóstico para la varianza en el periodo t+1 a partir de información hasta el periodo t, se puede calcular de la siguiente manera, para cualquier $i = 1, \dots, N$ se tiene:

$$\begin{aligned}\hat{\sigma}_{i,t+1}^2 &= (1-\lambda) \sum_{j=1}^{\infty} \lambda^j r_{i,t-j}^2 = (1-\lambda)(r_{i,t}^2 + \lambda r_{i,t-1}^2 + \lambda^2 r_{i,t-2}^2 + \dots) \\ &= (1-\lambda)r_{i,t}^2 + \lambda(1-\lambda)(r_{i,t-1}^2 + \lambda r_{i,t-2}^2 + \lambda^2 r_{i,t-3}^2 + \dots) \\ &= (1-\lambda)r_{i,t}^2 + \lambda \hat{\sigma}_{i,t}^2\end{aligned}$$

De igual manera, el estimador para las covarianzas está dado por:

$$\hat{\sigma}_{i,k,t}^2 = (1-\lambda) \sum_{j=1}^T \lambda^{j-1} (r_{i,t})(r_{k,t})$$

El cual puede expresarse en forma recursiva como sigue:

$$\hat{\sigma}_{ik,t+1}^2 = \lambda \hat{\sigma}_{ik,t}^2 + (1-\lambda)r_{i,t}r_{k,t}$$

Estos resultados corresponden a las estimaciones tanto de la varianza como de la covarianza con horizonte de t a t+1 que puede considerarse como de un día.

En muchos casos el interés dentro del análisis de riesgo es el de pronosticar horizontes mayores a un día. A continuación se describe cómo obtener los pronósticos tanto para las varianzas, como para las covarianzas en el método EWMA para horizontes de tiempo más largos, pero que es aplicable con otros métodos de estimación. En general si denotamos como h (h días) el horizonte de tiempo para el cuál se desea hacer el pronóstico, los pronósticos se pueden calcular como sigue:

Recordemos que para RiskMetrics el logaritmo de los precios se genera de acuerdo al modelo:

$$p_t = p_{t-1} + \sigma_t \varepsilon_t \quad \varepsilon_t \sim \text{iid } N(0, 1)$$

de donde se sigue que $p_{t+h} - p_t = \sigma_{t+1} \varepsilon_{t+1} + \dots + \sigma_{t+h} \varepsilon_{t+h}$ que en términos de rendimiento lo podemos expresar como:

$$r_{t+h} = \sum_{i=1}^h \sigma_{t+i} \varepsilon_{t+i}$$

Tomando la varianza de la ecuación anterior en el tiempo t se tiene que:

$$\sigma_{t+h}^2 = \text{Var}(r_{t+h}) = \text{Var}\left(\sum_{i=1}^h \sigma_{t+i} \varepsilon_{t+i}\right) = \sum_{i=1}^h \sigma_{t+i}^2 \text{Var}(\varepsilon_{t+i}) = \sum_{i=1}^h \sigma_{t+i}^2 = h \sigma_{t+1}^2, \text{ esta última}$$

igualdad de obtiene como sigue:

Para h = 1 se tiene que $r_{t+1} = \sigma_{t+1} \varepsilon_{t+1}$ de donde $\text{Var}(r_{t+1}) = \text{Var}(\sigma_{t+1} \varepsilon_{t+1}) = \sigma_{t+1}^2$, para h = 2 se tiene que

$r_{t+2} = \sigma_{t+1} \varepsilon_{t+1} + \sigma_{t+2} \varepsilon_{t+2}$, pero del modelo de estimación de la varianza, se tiene que

$\hat{\sigma}_{t+2}^2 = \lambda \hat{\sigma}_{t+1}^2 + (1-\lambda)r_{t+1}^2$ de donde, tomando la esperanza se tiene que

$$\sigma_{t+2}^2 = E(\hat{\sigma}_{t+2}^2) = E(\lambda \hat{\sigma}_{t+1}^2 + (1-\lambda)r_{t+1}^2) = \lambda \sigma_{t+1}^2 + (1-\lambda)E(r_{t+1}^2) = \lambda \sigma_{t+1}^2 + (1-\lambda)\sigma_{t+1}^2 = \sigma_{t+1}^2$$

Entonces $r_{t+2} = \sigma_{t+1} \varepsilon_{t+1} + \sigma_{t+2} \varepsilon_{t+2} = \sigma_{t+1} \varepsilon_{t+1} + \sigma_{t+1} \varepsilon_{t+2} = \sigma_{t+1} (\varepsilon_{t+1} + \varepsilon_{t+2})$

$$\text{Var}(r_{t+2}) = \text{Var}(\sigma_{t+1} (\varepsilon_{t+1} + \varepsilon_{t+2})) = \sigma_{t+1}^2 \text{Var}(\varepsilon_{t+1} + \varepsilon_{t+2}) = 2\sigma_{t+1}^2$$

De donde se sigue que

$$\hat{\sigma}_{i,t+h}^2 = h * \hat{\sigma}_{i,t+1}^2 \quad i = 1, \dots, N$$

$$\hat{\sigma}_{ij,t+h}^2 = h * \hat{\sigma}_{ij,t+1}^2 \quad i \neq j \quad i, j = 1, \dots, N$$

Las dos expresiones anteriores indican que los pronósticos, en el horizonte de pronóstico, se calculan como múltiplos del pronóstico a 1 día. Para diferenciar los pronósticos a 1 día con respecto a los pronósticos a un mes, lo que suele hacerse es utilizar distintos valores de λ al momento de pronosticar.

Un aspecto importante de la estimación de un promedio móvil ponderado exponencialmente es seleccionar adecuadamente el valor del factor de decaimiento (λ), lo cual depende de que tan rápido cambie el nivel promedio de la serie en el tiempo. Si el nivel de la serie cambia muy rápido, deberán seleccionarse valores de λ cercanos a uno, cuando el nivel de la serie cambia más lento el valor de λ deberá ser menor. También es importante determinar para un valor dado de λ , cuanta información se requiere para obtener estimaciones adecuadas con este método.

En la ecuación del estimador de la varianza decimos que para T grande $(\lambda)^T \hat{\sigma}_{i,0}^2$ lo podemos considerar cero, entonces sea TL el nivel de tolerancia, esto es, el valor de T a partir del cual los valores de la ponderación son poco significativos y que denominaremos como

$$TL = (1-\lambda) \sum_{j=T}^{\infty} \lambda^j \text{ esto es, } TL = (1-\lambda) \lambda^T (1 + \lambda + \lambda^2 + \lambda^3 + \dots)$$

Como $0 < \lambda < 1$, tenemos que la serie $(1 + \lambda + \lambda^2 + \lambda^3 + \dots)$ converge a $1 / (1 - \lambda)$, entonces

$$TL = (1-\lambda) \lambda^T / (1-\lambda), \text{ de donde } TL = \lambda^T$$

y por tanto $T = (\text{Log}(TL) / \text{Log}(\lambda))$

Este resultado permite calcular para, valores de λ y de (TL), la cantidad de datos en días que requiere el método de EWMA para la estimación de la volatilidad, (ver tabla 2).

TABLA 2

| Valor de λ | TL = .00001 | TL = .0001 | TL = .001 | TL = .01 |
|--------------------|-------------|------------|-----------|----------|
| 0.85 | 71 | 57 | 43 | 28 |
| 0.86 | 76 | 61 | 46 | 31 |
| 0.87 | 83 | 66 | 50 | 33 |
| 0.88 | 90 | 72 | 54 | 36 |
| 0.89 | 99 | 79 | 59 | 40 |
| 0.90 | 109 | 87 | 66 | 44 |
| 0.91 | 122 | 98 | 73 | 49 |
| 0.92 | 138 | 110 | 83 | 55 |
| 0.93 | 159 | 127 | 95 | 63 |
| 0.94 | 186 | 149 | 112 | 74 |
| 0.95 | 224 | 180 | 135 | 90 |
| 0.96 | 282 | 226 | 169 | 113 |
| 0.97 | 378 | 302 | 227 | 151 |
| 0.98 | 570 | 456 | 342 | 228 |
| 0.99 | 1146 | 916 | 687 | 458 |

Con la estimación de la varianza se tiene que el VaR resultante de este enfoque es:

$$\text{VaR} = V_t z_{1-c} \sigma_t \quad (5)$$

donde z_{1-c} es el cuantil de la distribución $N(0,1)$ al $1-c$ % de probabilidad

Merece la pena hacer dos observaciones a la anterior fórmula. La primera es que implícitamente σ_t es una estimación de σ que *cambia* con el tiempo, pues para calcular σ_t se utilizan distintas secuencias de observaciones históricas cada vez. Por otro lado, y más importante, al sustituir en (3) σ por σ_t pasamos literalmente a desconocer las propiedades probabilísticas del VaR resultante, puesto que evidentemente estas características concretas no se conservan.

3.3 Los Modelos De Heteroscedasticidad Autorregresiva Condicionada ARCH y GARCH

La modelación y pronóstico de la volatilidad o matriz de varianzas y covarianzas con modelos ARCH, que modelan la varianza condicionada a la información anterior, fue introducido por Robert F. Engle [27] que partió de la idea de modelar y pronosticar la volatilidad como una función variante en el tiempo de la información pasada.

La clase de modelos GARCH de los cuales el GARCH(1,1) es el más utilizado, fue introducido por Tim Bollerslev [11], modelo que también fue estudiado de manera independiente por Taylor [74]. Dichos modelos, incluyen el pronóstico de la volatilidad y se han estudiado exhaustivamente por diferentes autores, por lo que en este trabajo sólo se resaltan algunos de los rasgos más importantes de dichos modelos que permiten la aplicación empírico práctica.

Es difícil obtener pronósticos óptimos de la volatilidad sin la noción de un modelo, o del proceso generador de los datos, tal y como lo hace la metodología de RiskMetrics y si a ello agregamos que, en el contexto de pronosticar la varianza de los rendimientos de varios periodos, se supone que los correspondientes estimadores de la varianza para un periodo son consecutivamente no correlacionados de manera que el estimador de la varianza sea igual a la suma de los sucesivos estimadores de la varianza, esto es.

$$\text{Var}(r_{t+k} + r_{t+k-1} + \dots + r_{t+1} | \mathfrak{F}_t) = \sigma_{t+k|t}^2 \approx k \hat{\sigma}_t^2$$

Significa que la volatilidad de los rendimientos multi-periodo es un múltiplo del pronóstico de un periodo por el horizonte k . En contraste, los modelos autoregresivos introducen la característica de varianza condicional (de ahí su nombre) a los cuales nos referiremos en los siguientes párrafos.

Para definir la clase de modelos GARCH, considere la descomposición de r_t para un periodo con media condicional $\mu_t = E(r_t | \mathfrak{F}_{t-1})$, y varianza condicional $\sigma_t^2 = \text{Var}(r_t | \mathfrak{F}_{t-1})$, donde \mathfrak{F}_{t-1} es el conjunto de la información pasada hasta el periodo t bajo el siguiente modelo:

$$r_t = \mu_t + \sigma_t \varepsilon_t \quad \varepsilon_t \text{ i.i.d con } E(\varepsilon_t) = 0 \text{ y } \text{Var}(\varepsilon_t) = 1 \quad (6)$$

El modelo GARCH(1,1) para la varianza condicional se define por la relación recursiva,

$$\sigma_t^2 = w + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (7)$$

Donde los parámetros son no-negativos ($\alpha \geq 0$, $w > 0$ y $\beta \geq 0$), a fin de asegurar que la varianza condicional sea positiva para todas las r_t del proceso.

Por sustituciones recursivas, el modelo GARCH(1,1) puede expresarse alternativamente como un modelo ARCH(∞), de la siguiente manera:

$$\sigma_t^2 = w(1 - \beta)^{-1} + \alpha \sum_{i=1}^{\infty} \beta^{i-1} r_{t-i}^2$$

Esto reduce el modelo a un filtro de promedios móviles ponderados exponencialmente con $w=0$, $\alpha=(1-\lambda)$ y $\beta = \lambda$, este modelo GARCH en el cual $\alpha+\beta=1$, es referido por algunos autores como un proceso GARCH Integrado, o IGARCH(1,1). Es importante señalar que, en la clase de modelos GARCH(1,1), y de manera general en los procesos ARCH, la noción del proceso generador de datos esta implícita en los supuestos de la distribución de las ε_i . Esto significa que la construcción de pronósticos óptimos de la varianza está dentro del contexto del modelo, al definir el comportamiento de la varianza de manera explícita.

En particular, se sigue directamente de la formulación del modelo que el pronóstico óptimo de la varianza, en un sentido de error cuadrático medio, para un periodo, es igual a σ_{t+1}^2 . Las expresión correspondiente para pronósticos de periodos de largo plazo σ_{t+h}^2 , con $h > 1$, se construye por procedimientos recursivos.

Al igual que en el modelo de Riskmetrics, se supone que la media condicional es igual a cero, es decir, $\mu_t=0$, además de que $\alpha + \beta < 1$ y dado que la varianza no-condicional del proceso existe, esta se puede expresar de acuerdo al teorema de Bollerslev [11], como:

$$\sigma^2 = w(1 - \alpha - \beta)^{-1} \quad (8)$$

El pronostico de la varianza en el periodo $t+h$ se puede calcular de la siguiente forma:

$\sigma_{t+1}^2 = w + \alpha r_t^2 + \beta \sigma_t^2$ de donde si tomamos la esperanza condicional, se tiene

$$E[\sigma_{t+1}^2 | \mathfrak{F}_t] = E[w + \alpha r_t^2 + \beta \sigma_t^2] = w + \alpha E[r_t^2 | \mathfrak{F}_t] + \beta E[\sigma_t^2 | \mathfrak{F}_t]$$

$= w + \alpha r_t^2 + \beta \sigma_t^2$, entonces dado que:

$$E[r_{t+1}^2 | \mathfrak{F}_t] = E[\sigma_{t+1}^2 | \mathfrak{F}_t] = \sigma_{t+1}^2 \text{ se tiene que}$$

$$E[\sigma_{t+2}^2 | \mathfrak{F}_t] = w + \alpha E[r_{t+1}^2 | \mathfrak{F}_t] + \beta E[\sigma_{t+1}^2 | \mathfrak{F}_t]$$

$$= w + (\alpha + \beta) E[\sigma_{t+1}^2 | \mathfrak{F}_t] = w + (\alpha + \beta) \sigma_{t+1}^2$$

Entonces, para h mayor que 2 se tiene que

$$\begin{aligned}
\sigma_{t+h}^2 &= w + (\alpha + \beta) \sigma_{t+h-1}^2 = w + (\alpha + \beta) (w + (\alpha + \beta) \sigma_{t+h-2}^2) \\
&= w \sum_{j=1}^{h-1} (\alpha + \beta)^j + (\alpha + \beta)^{h-1} \sigma_{t+1}^2 \\
&= w \left(\frac{1 - (\alpha + \beta)^{h-1}}{1 - (\alpha + \beta)} \right) + (\alpha + \beta)^{h-1} \sigma_{t+1}^2 \\
&= \sigma^2 (1 - (\alpha + \beta)^{h-1}) + (\alpha + \beta)^{h-1} \sigma_{t+1}^2 \\
&= \sigma^2 + (\alpha + \beta)^{h-1} (\sigma_{t+1}^2 - \sigma^2) \quad (9)
\end{aligned}$$

Expresión que muestra que los pronósticos revierten a largo plazo la varianza no condicional a una tasa exponencial sujeta al valor de $\alpha + \beta$.

Además si los rendimientos son no correlacionados, la varianza condicional a la información anterior para el pronóstico del periodo de t a $t+k$, es la suma de las varianzas condicionales de cada periodo, esto es:

$$\begin{aligned}
\sigma_{t:t+k}^2 &= \sum_{j=1}^k \sigma^2 + (\alpha + \beta)^j (\sigma_{t+1}^2 - \sigma^2) \\
&= k \sigma^2 + (\sigma_{t+1}^2 - \sigma^2) \sum_{j=1}^k (\alpha + \beta)^j \\
&= k \sigma^2 + (\sigma_{t+1}^2 - \sigma^2) \left(\frac{1 - (\alpha + \beta)^{k+1}}{1 - (\alpha + \beta)} \right)
\end{aligned}$$

De esta última expresión y de la expresión (9) se sigue que

$$E(k \hat{\sigma}_{t+1}^2) = k \sigma^2 = E(\hat{\sigma}_{t:t+k}^2)$$

Sin embargo, comparando la varianza del pronóstico escalado por k con la varianza del pronóstico óptimo se tiene que:

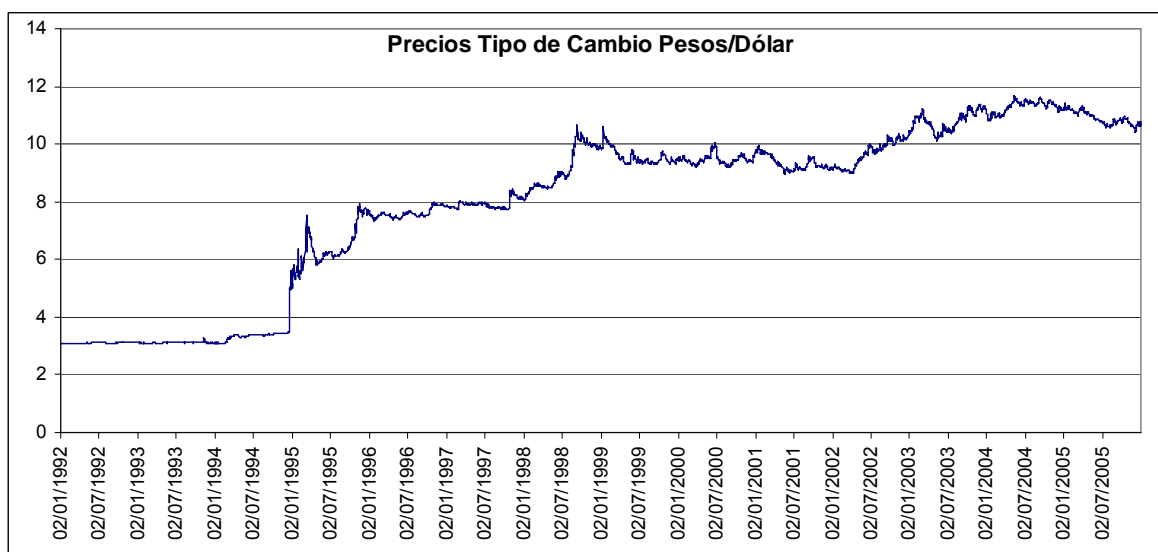
$$Var(k \hat{\sigma}_{t+1}^2) = k^2 Var(\hat{\sigma}_{t+1}^2) > \left(1 - (\alpha + \beta)^k\right)^2 (1 - \alpha - \beta)^{-2} Var(\hat{\sigma}_{t+1}^2) = Var(\hat{\sigma}_{t:t+k}^2)$$

De donde el pronóstico escalado sobre la estimación de un día tranquilo subestima el verdadero riesgo, mientras que el riesgo es sobrestimado cuando se escala sobre el pronóstico de un día malo, lo que significa que no es muy prudente escalar el riesgo como proceso de administración del riesgo.

Ejemplo de un modelo GARCH(1,1)

Con el paquete MatLab 7.0 se estimó un modelo GARCH(1,1) del tipo de cambio pesos/dólar para el periodo del 3/01/1992 al 30/12/2005 y del cual damos cuenta a continuación:

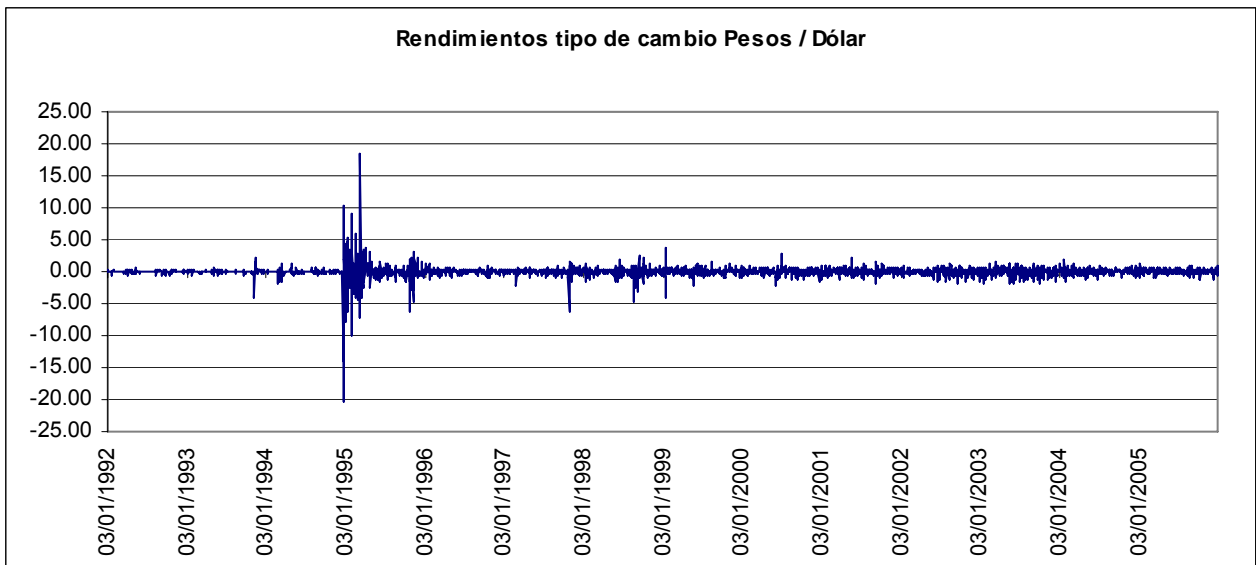
Un análisis de pre-estimación es necesario para ver el comportamiento de la serie y observar si los supuestos del modelo se cumplen. Para tal fin, se grafica la serie del tipo de cambio, para observar si existe evidencia de que la serie sea no estacionaria, lo que es necesario, toda vez que el modelo tiene un buen funcionamiento en condiciones estables del mercado, es decir en condiciones estacionarias.



Como el modelo se construye sobre los rendimientos y no sobre los precios es necesario construir la serie de rendimientos a partir de la serie anterior como sigue:

$$r_t = \ln(P_t) - \ln(P_{t-1})$$

que se grafica, para observar si existe evidencia de volatilidad agrupada, la cual explica parte del efecto de la cola pesada, toda vez que otra parte de ese efecto se puede explicar si la distribución de los rendimientos no es una distribución normal, sino, por ejemplo, una t de student, además de mostrarnos posible autocorrelación de la volatilidad y su variabilidad en el tiempo.



De manera general se observa en la gráfica anterior que en los periodos de 1995 y parte de 1998 existe volatilidad alta agrupada, mientras que en los periodos de mediados de 1992 y primer trimestre de 1993 existe volatilidad baja agrupada.

Los modelos RiskMetrics y GARCH, consideran que la volatilidad es cambiante en el tiempo y además autocorrelacionada, supuestos que se pueden detectar en la serie de tiempo de los rendimientos, toda vez que si los rendimientos fueran independientes e idénticamente distribuidos, ello implicaría que:

- En cada punto del tiempo, los rendimientos tuvieran una distribución con media = 0 y varianza = σ^2 (idénticamente distribuidos). Lo que implica que la media y varianza de los rendimientos son homocedásticos, es decir, no cambiantes en el tiempo
- También significaría que no existe relación entre si de cualesquiera de ellos, es decir, que cada rendimiento se generaría sin que los rendimientos anteriores tuvieran alguna implicación en él.

Existen diferentes pruebas para detectar la independencia de una muestra, aquí sólo mostraremos algunas de ellas. Una primera es ver el coeficiente de autocorrelación el cual mide la correlación de los rendimientos a través del tiempo. De manera general, el coeficiente de correlación entre dos variables aleatorias X e Y esta dado por la covarianza de X e Y dividido por sus respectivas desviaciones estándares, esto es:

$$\rho_{XY} = \text{Cov}(X,Y) / \sqrt{\text{Var}(X)\text{Var}(Y)}$$

Para una muestra de tamaño N de ambas variables, el estimador de la correlación de estas variables se calcula como:

$$\hat{\rho}_{XY} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

donde \bar{x} y \bar{y} son los estimadores de la media de cada variable respectivamente

Siguiendo esta definición, para la serie de rendimientos r_t $t=1,2,\dots,T$, se pueden definir T-1 pares de observaciones como $(r_1, r_2), \dots, (r_{T-1}, r_T)$, si consideramos que la primera observación de estos pares representa, una variable y la segunda observación representa otra variable, se tiene que el coeficiente de autocorrelación se puede estimar como:

$$\hat{\rho} = \frac{\sum_{i=1}^{T-1} (x_i - \bar{x}_1)(x_{i+1} - \bar{x}_2)}{\sqrt{\sum_{i=1}^{T-1} (x_i - \bar{x}_1)^2 \sum_{i=1}^{T-1} (x_{i+1} - \bar{x}_2)^2}}$$

donde \bar{x}_1 y \bar{x}_2 son los estimadores de la media para cada variable,

Para tamaños de muestra grandes, se puede observar que los estimadores de la media de ambas variables son semejantes, por lo que los podemos sustituir por el estimador de la media general, de donde se obtiene que el estimador se puede expresar como:

$$\hat{\rho} \cong \frac{\sum_{i=1}^{T-1} (x_i - \bar{x})(x_{i+1} - \bar{x})}{\sqrt{\sum_{i=1}^{T-1} (x_i - \bar{x})^2 \sum_{i=1}^{T-1} (x_{i+1} - \bar{x})^2}} \cong \frac{\sum_{i=1}^{T-1} (x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^{T-1} (x_i - \bar{x})^2} \cdot \frac{T-1}{T}$$

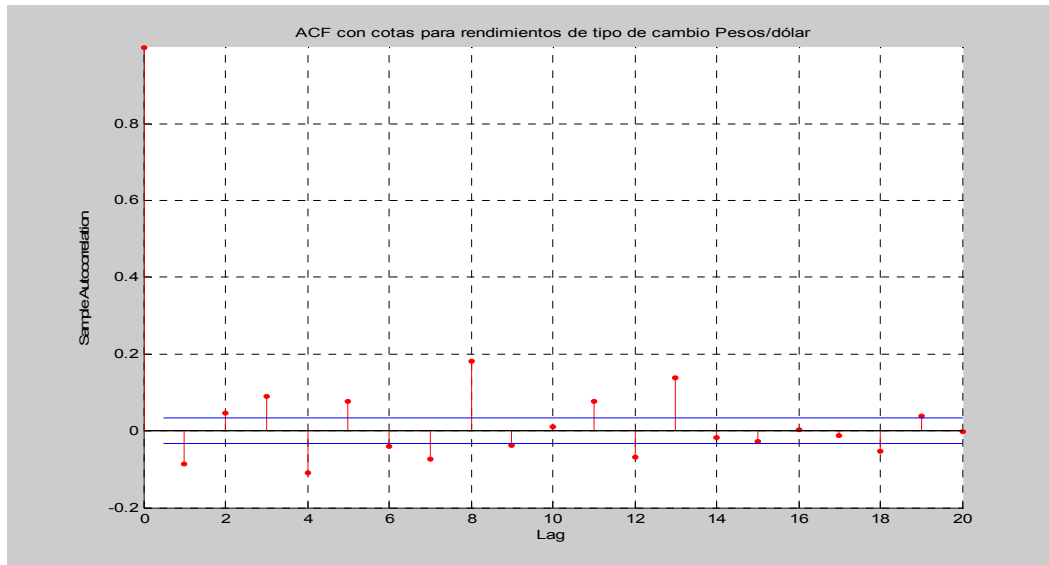
de acuerdo con lo anterior, la estimación de la autocorrelación para pares de observaciones con k periodos de distancia será igual a:

$$\hat{\rho}_k \cong \frac{\sum_{i=1}^{T-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{T-k} (x_i - \bar{x})^2} \cdot \frac{T-k}{T}$$

Si la serie de rendimientos no es autocorrelacionada, entonces el valor estimado de la autocorrelación (ρ_k) debe ser cercano a cero, además si el tamaño de muestra es mayor a 1000 observaciones, de acuerdo a varios autores, es posible calcular bandas de

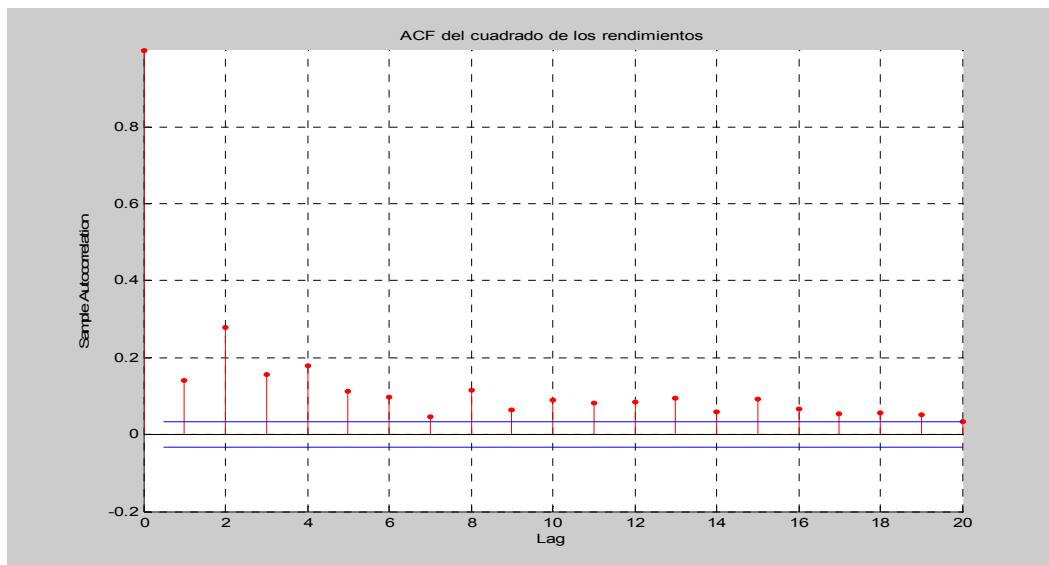
confianza al 95% alrededor de cero, para cada coeficiente de autocorrelación como $1.96/\sqrt{T}$ y $-1.96/\sqrt{T}$.

Regresando a nuestro ejemplo, se obtuvo la siguiente grafica de autocorrelación:



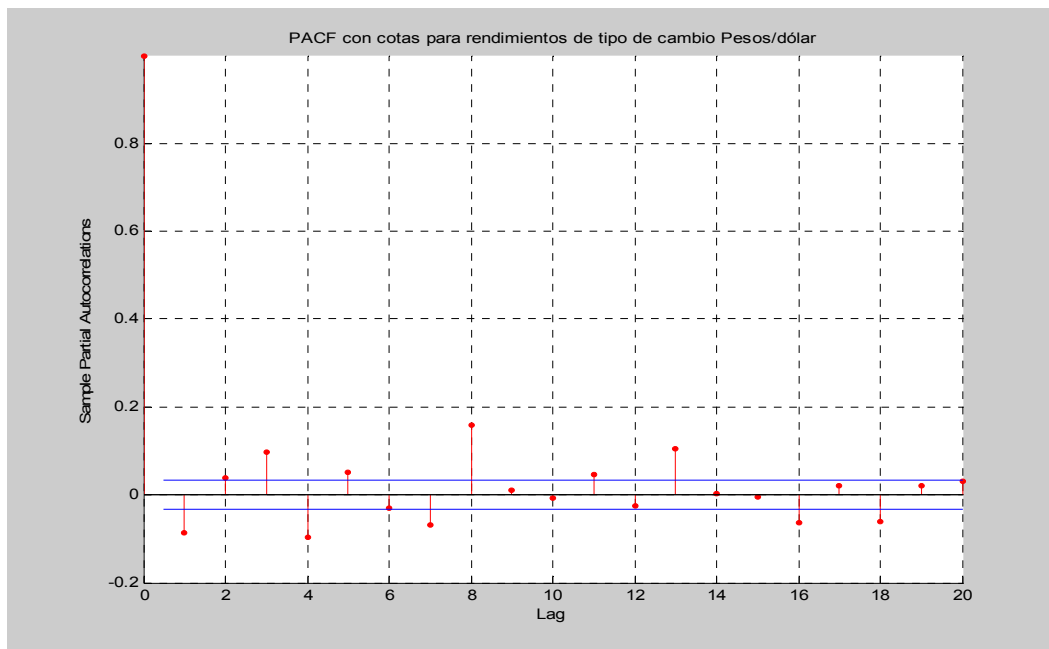
De la gráfica se desprende que la serie presenta autocorrelación

A fin de confirmar este hecho se puede realizar la misma prueba sobre la serie de rendimientos al cuadrado, de donde se obtuvo la siguiente gráfica:



Otro elemento gráfico y numérico para detectar autocorrelación la da el estimador de la autocorrelación parcial, el cual se define, para una distancia k de las observaciones, como

la correlación entre r_t y r_{t+k} quitando el efecto de las observaciones $r_{t+1}, r_{t+2}, \dots, r_{t+k-1}$ y cuya gráfica mostramos:



Si bien los métodos gráficos y numéricos anteriores nos dan una idea general de la autocorrelación de la serie de rendimientos, un método más formal para probar la autocorrelación es la prueba a partir de la estadística de Box_Ljung definida como:

$$BL(p) = T(T + 2) \sum_{k=1}^p \frac{\rho_k^2}{T - k}$$

Bajo la hipótesis nula de que la serie de rendimientos es no autocorrelacionada, $BL(p)$ se distribuye como una χ -cuadrada con p grados de libertad, donde p denota el número de autocorrelaciones usadas para estimar la estadística y T es el tamaño de la muestra, entonces aplicando esta prueba a nuestra serie de rendimientos pesos/dólar tenemos los siguientes resultados:

| Valor de p | Valor de la estadística | Valor de la χ -cuadrada |
|------------|-------------------------|------------------------------|
| 10 | 270.5518 | 18.3070 |
| 15 | 379.0075 | 24.9958 |
| 20 | 395.0613 | 31.4104 |

Resultados que confirman lo observado en las graficas de autocorrelación total y parcial, al ser rechazada la hipótesis nula.

Los efectos de asimetría e “Influencia”

El modelo GARCH supone que los shocks positivos y negativos de igual magnitud absoluta, influyen de manera idéntica en la varianza condicional futura. En contraste, se observa en la práctica, que la volatilidad de índices agregados de acciones, muestran una respuesta asimétrica a shocks negativos los cuales producen rendimientos negativos que resultan en volatilidades futuras más grandes. Esta asimetría generalmente es llamada el efecto “influencia”, la cual, no puede explicar por si sola la magnitud del efecto, ya que para acciones individuales los efectos son menos pronunciados. Las tres formulaciones GARCH, comúnmente más usadas para describir este tipo de asimetría son el GJR o GARCH umbral (TGARCH) modelados en Glosten, Jagannathan y Runkle [36] y Zakoïan [79], el GARCH Asimétrico (AGARCH) modelo de Engle y Ng [30], y el GARCH Exponencial (EGARCH) modelo de Nelson [60].

La varianza condicional en el modelo GJR(1,1), o TGARCH(1,1), simplemente introduce en el modelo GARCH(1,1) un término ARCH adicional de acuerdo al signo del rendimiento pasado como sigue;

$$\sigma_t^2 = w + \alpha r_{t-1}^2 + \gamma r_{t-1}^2 I(r_{t-1} < 0) + \beta \sigma_{t-1}^2$$

Donde $I(\cdot)$ denota la función indicadora. Se tiene que para $\gamma > 0$, los shocks de rendimientos negativos anteriores, tienen un impacto sobre la varianza condicional futura. Mecánicamente, el cálculo del pronóstico de la varianza multi-periodo trabaja exactamente como en el modelo GARCH. En particular, suponiendo que, $\Pr(r_t < 0) = 0.5$ se tiene que:

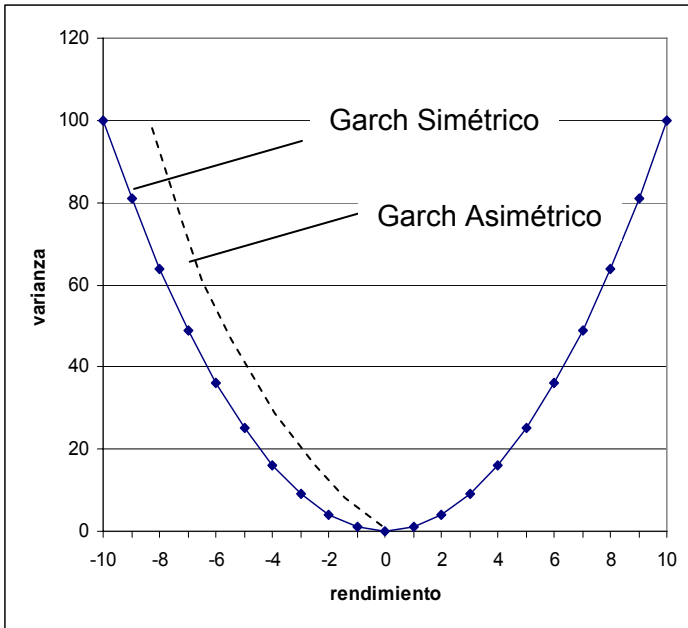
$$\sigma_{t+h}^2 = \sigma^2 + (\alpha + 0.5\gamma + \beta)^{h-1} (\sigma_{t+1}^2 - \sigma^2)$$

donde la varianza no-condicional, ahora es igual,

$$\sigma^2 = w(1 - \alpha - 0.5\gamma - \beta)^{-1}$$

Aunque la fórmula del pronóstico es casi idéntica a la del modelo GARCH(1,1) en la ecuación (9), la inclusión del término asimétrico puede afectar los pronósticos alterando el valor de la varianza condicional σ_{t+1}^2 .

La curva de impacto de noticias, definida por la relación funcional entre σ_t^2 y r_{t-1} manteniendo todas las otras variables constantes, proporciona una manera simple de caracterizar la influencia de los más recientes shocks sobre la varianza condicional de los próximos periodos. En el modelo GARCH esta curva es cuadrática alrededor de $r_{t-1} = 0$, mientras que en el modelo GJR, con $\gamma > 0$, la curva presenta valores mayores de la varianza para valores negativos de r_{t-1} (vea la siguiente gráfica).



De manera alterna, el modelo GARCH Asimétrico, o AGARCH(1,1) de la forma,

$$\sigma_t^2 = w + \alpha (r_{t-1} - \gamma)^2 + \beta \sigma_{t-1}^2$$

desplaza el centro de la curva de impacto de noticias de cero a γ , capturando de manera diferente los efectos asimétricos.

En lugar de modelar directamente la varianza condicional, el modelo EGARCH, formula la varianza en términos de su logaritmo, como muestra el modelo EGARCH(1,1),

$$\log(\sigma_t^2) = w + \alpha \left(|z_{t-1}| - E(|z_{t-1}|) \right) + \gamma z_{t-1} + \beta \sigma_{t-1}^2 \quad (10)$$

Donde $z_{t-1} = r_t / \sigma_t$, el modelo en términos de logaritmos tiene la ventaja de evitar las restricciones de no-negatividad en los parámetros, dado que la varianza implicada por la varianza logarítmica exponentiada del modelo garantiza ser positivo. Como en los modelos GJR y AGARCH, en el modelo EGARCH valores de $\gamma > 0$ capturan directamente la respuesta asimétrica, o el efecto “influencia”.

Sin embargo, dado que la derivada de este modelo en z_t igual a cero no está definida, el modelo EGARCH es más difícil de estimar y analizar numéricamente. Desde una perspectiva de pronóstico, la forma recursiva del modelo EGARCH (definida por la ecuación (10)), entrega de manera rápida el pronóstico óptimo -en un sentido de error cuadrático medio- para el futuro logaritmo de la varianza condicional $E(\log(\sigma_{t+k}^2) | \mathcal{I}_t)$.

Los modelos GARCH, TGARCH, AGARCH y EGARCH presentados, implican que los shocks sobre la volatilidad caen en una proporción exponencial. Para ilustrar esto,

considere el modelo GARCH(1,1). Se sigue de la ecuación (9) que el efecto impulso de un shock, sobre el pronóstico de la varianza en el periodo futuro h, esta dado por

$$\frac{\partial \sigma_{t+h}^2}{\partial r_t^2} = \alpha (\alpha + \beta)^{h-1}, \text{ o de manera general,}$$

$$\frac{\partial \sigma_{t+h}^2}{\partial r_t^2} = b(\delta)^h$$

donde $0 < \delta < 1$. Esta caída exponencial trabaja bien para pronósticos de horizontes cortos. Sin embargo, en los estudios de Ding, Granger y Engle [23] y Andersen y Bollerslev [2], se argumenta que la autocorrelación de los rendimientos al cuadrado y absolutos cae en proporción hiperbólica muy lenta, sobre retrasos más largos. En este contexto, mejores pronósticos de volatilidad a largo plazo, usando modelos GARCH modelados en términos de r_t^2 , podrían obtenerse al formular la varianza condicional de forma que el efecto impulso se comporte como,

$$\frac{\partial \sigma_{t+h}^2}{\partial r_t^2} \approx b(h)^\delta$$

para valores grandes de b.

Atendiendo esta observación, se han sugerido modelos de memoria larga, o fraccionarios integrados, tipo GARCH en la literatura para lograr esta meta.

Por ejemplo, en el modelo integrado fraccional FIGARCH(1,1) propuesto por Baillie, Bollerslev y Mikkelsen [7], la varianza condicional se define como:

$$\sigma_t^2 = w + \beta \sigma_{t-1}^2 + \left[1 - \beta L - (1 - \alpha L - \beta L)(1 - L)^d \right] r_t^2 \text{ donde } L \text{ es el operador de retraso}$$

Para $d = 0$ el modelo se reduce al modelo GARCH(1,1), pero para valores de $0 < d < 1$ el efecto de los shocks para los pronósticos de la volatilidad en el modelo disminuyen a una tasa lenta hiperbólica. Los pronósticos de este modelo se construyen por la forma recursiva,

$$\sigma_{t+k}^2 = w(1 - \beta)^{-1} + \lambda(L) \sigma_{t+k-1}^2$$

con $\sigma_{t+k}^2 = \varepsilon_t^2$ para $k < 0$, y el coeficiente

$$\lambda(L) = \left[1 - (1 - \beta L)^{-1} (1 - \alpha L - \beta L)(1 - L)^d \right] \text{ calculado recursivamente como}$$

$$\lambda_1 = \alpha + d \quad \lambda_j = \beta \lambda_{j-1} + \left[(j-1-d) j^{-1} - (\alpha + \beta) \right] \delta_{j-1} \quad j = 2, 3, \dots$$

donde $\delta_j = \delta_{j-1}(j-1-d)j^{-1}$ son los coeficientes en la serie de MaClaurin del operador fraccionario diferenciando $(1-L)^d$. En estos modelos se puede introducir asimetrías permitiendo influencias como en el modelo GJR o TGARCH.

Estimación del modelo

3.3.1.a La función de verosimilitud

Sea X_1, X_2, \dots, X_n una muestra aleatoria de una distribución que tiene como función de distribución a $f(x, \theta)$, $\theta \in \Omega$. La distribución conjunta de X_1, X_2, \dots, X_n es $f(x_1; \theta)f(x_2; \theta) \dots f(x_n; \theta)$, distribución que es considerada como una función de θ , y que es llamada la función de verosimilitud (L) de la muestra aleatoria, que se denota como sigue:

$$L(\theta, x_1, x_2, \dots, x_n) = f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta)$$

Suponga que podemos encontrar una función $u(x_1, x_2, \dots, x_n)$ de x_1, x_2, \dots, x_n , tal que cuando θ es remplazada por $u(x_1, x_2, \dots, x_n)$ la función de verosimilitud L se maximiza. Entonces dado que $L(u(x_1, x_2, \dots, x_n); x_1, x_2, \dots, x_n)$ es mayor o igual a $L(\theta; x_1, x_2, \dots, x_n)$ para todo $\theta \in \Theta$, si $u(x_1, x_2, \dots, x_n) \in \Theta$ para todo punto (x_1, x_2, \dots, x_n) positivo de la densidad de probabilidad, entonces la estadística $u(x_1, x_2, \dots, x_n)$ es llamada la estadística de máxima verosimilitud.

Del proceso definido en las ecuaciones (6) y (7) generalizado se tiene que todas las observaciones (ε_t) tienen distribución normal condicional a la información anterior (\mathfrak{F}) con media cero y varianza $h_t = \sigma_t^2$

Sea $f(\varepsilon | \mathfrak{F}) = \frac{1}{\sqrt{2h_t\pi}} \exp\left[-\frac{\varepsilon_t^2}{2h_t}\right]$ la función condicional de distribución de cada observación

Tomando el logaritmo natural de la anterior expresión se tiene

$$\begin{aligned} \ln f(\varepsilon) &= \ln(1) - \frac{1}{2} [\ln(2\pi) + \ln h_t] - \frac{1}{2} \varepsilon_t^2 h_t^{-1} \\ &= -\frac{1}{2} \ln h_t - \frac{1}{2} \varepsilon_t^2 h_t^{-1} - \frac{1}{2} \ln(2\pi) \end{aligned}$$

De acuerdo al método de máxima verosimilitud tenemos que la densidad conjunta es el producto de todas las funciones condicionales por lo que si aplicamos la función logaritmo natural (Ln) a dicho producto obtenemos que el Ln de la función de verosimilitud, es la suma de las correspondientes ln-normales.

Sea T el tamaño de muestra, entonces podemos definir la función de verosimilitud, salvo por una constante como

$$l = \frac{1}{T} \sum_{t=1}^T l_t$$

donde

$$l_t = -\frac{1}{2} \ln h_t - \frac{1}{2} \frac{\varepsilon_t^2}{h_t}$$

Sea $z_t' = (1, \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \dots, \varepsilon_{t-q}^2, h_{t-1}, h_{t-2}, \dots, h_{t-p})$, $w' = (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \dots, \beta_p)$ entonces el modelo GARCH(p,q) lo podemos escribir como sigue:

$$r_t = h_t \varepsilon_t$$

$$h_t = z_t' w$$

$$\varepsilon_t \sim N(0, h_t)$$

De acuerdo al método de máxima verosimilitud, el estimador se encuentra al resolver las condiciones de primer orden. La derivada de l_t con respecto a w es:

$$\begin{aligned} \frac{\partial l_t}{\partial w} &= \frac{\partial}{\partial w} \left[-\frac{1}{2} \ln(h_t) - \frac{1}{2} \frac{\varepsilon_t^2}{h_t} \right] \\ &= -\frac{1}{2} \frac{\partial \ln(h_t)}{\partial w} - \frac{1}{2} \frac{\partial}{\partial w} \frac{\varepsilon_t^2}{h_t} \\ &= -\frac{1}{2} \frac{\partial h_t}{h_t \partial w} + \frac{1}{2} \frac{\partial h_t}{\partial w} \frac{\varepsilon_t^2}{h_t^2} \\ &= \frac{1}{2} \frac{\partial h_t}{h_t \partial w} \left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \end{aligned}$$

y el Hessiano es:

$$\begin{aligned} \frac{\partial^2 l_t}{\partial w \partial w'} &= \frac{\partial}{\partial w'} \left[\frac{1}{2} \frac{\partial h_t}{h_t \partial w} \left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \right] \\ &= \left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial w'} \left[\frac{1}{2} \frac{\partial h_t}{h_t \partial w} \right] - \frac{1}{2} \frac{\partial h_t}{h_t^2} \frac{\partial h_t}{\partial w} \frac{\varepsilon_t^2}{h_t} \end{aligned}$$

cuando

$$\begin{aligned} \frac{\partial h_t}{\partial w} &= \frac{\partial}{\partial w} [z_t w] \\ &= z_t \frac{\partial w}{\partial w} + w \frac{\partial z_t}{\partial w} \\ &= z_t + \sum_{i=1}^p \beta_i \frac{\partial h_{t-i}}{\partial w} \end{aligned}$$

Las ecuaciones hasta aquí descritas son para el modelo GARCH, y sólo se diferencian del modelo ARCH en la inclusión de la parte recursiva representada por la ecuación anterior.

Dada la complejidad de las ecuaciones de primer orden para los diferentes parámetros, los estimadores de máxima verosimilitud son obtenidos al resolver por algún método numérico dichas ecuaciones.

Retomando nuestro ejemplo, se presentan los resultados de la estimación, utilizando la función garchfit del paquete MatLab 7.0

$$r_t = C + \varepsilon_t \quad \sigma_t^2 = k + \beta \sigma_{t-1}^2 + \alpha r_{t-1}^2$$

Number of variables: 4

Functions

| | |
|------------------------------------|---------------------------------------|
| Objective: | Garchlfn |
| Gradient: | finite-differencing |
| Hessian: | finite-differencing (or Quasi-Newton) |
| Nonlinear constraints: | Armanlc |
| Gradient of nonlinear constraints: | finite-differencing |

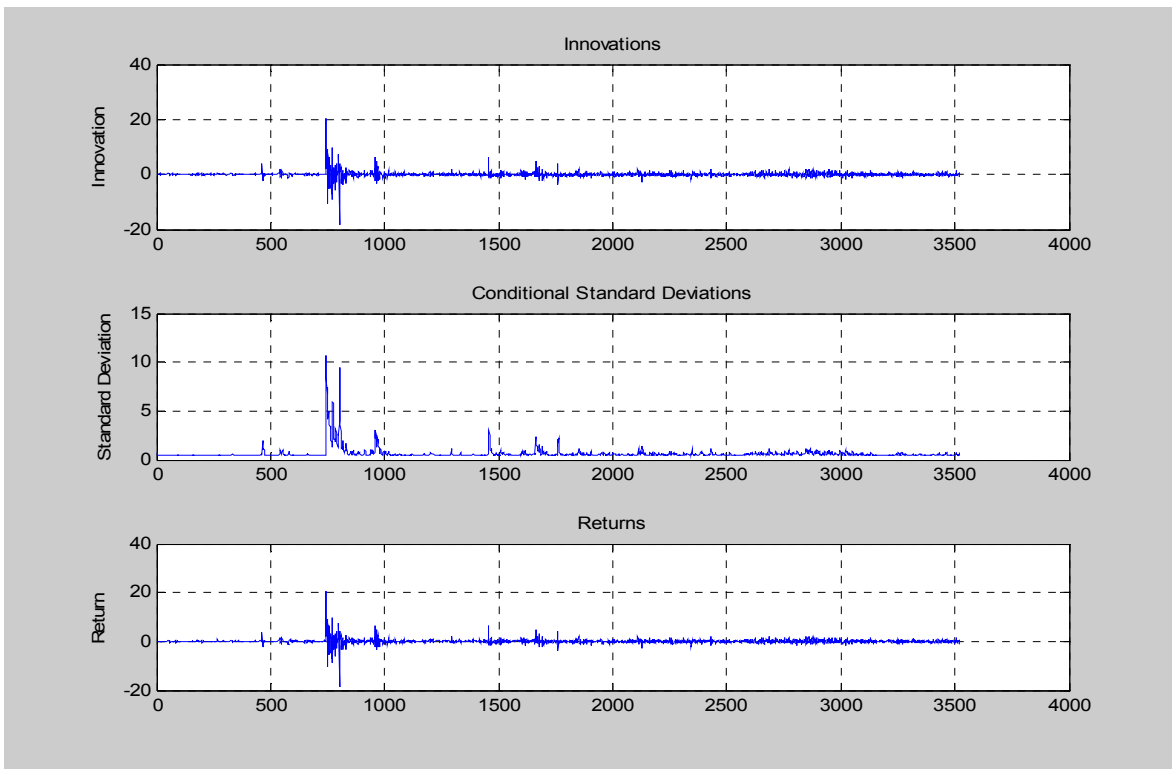
Variance: GARCH(1,1)

Conditional Probability Distribution: Gaussian

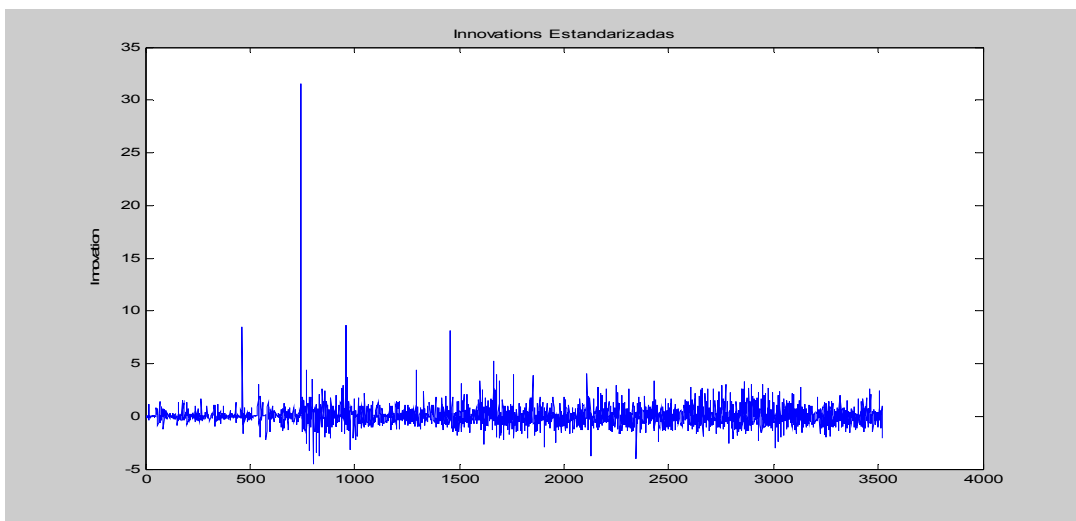
Number of Model Parameters Estimated: 4

| Parameter | Value | Standard Error | T Statistic |
|-----------|-----------|----------------|-------------|
| C | -0.011177 | 0.011565 | -0.9664 |
| K | 0.074059 | 0.0029306 | 25.2706 |
| GARCH(1) | 0.61208 | 0.014535 | 42.1120 |
| ARCH(1) | 0.23422 | 0.009727 | 24.0797 |

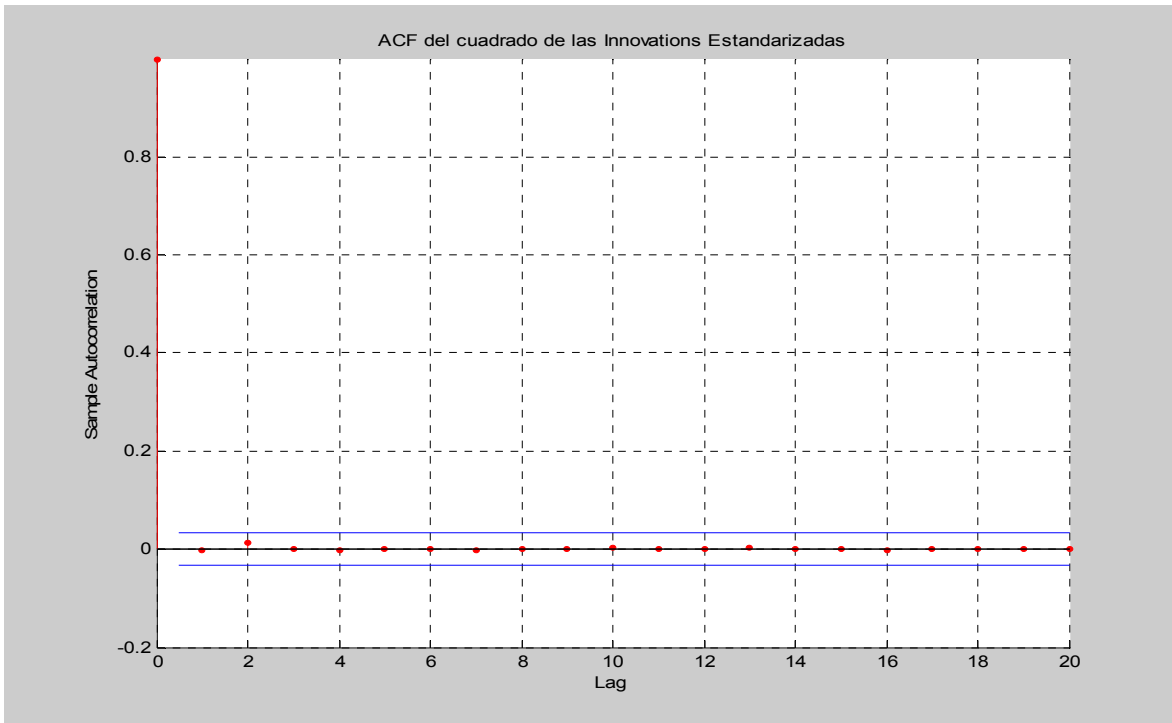
Un primer análisis es de tipo gráfico para comparar los residuales (innovaciones), las desviaciones estándar (sigmas) y los rendimientos.



Si bien, la gráfica anterior, muestra que los residuales ajustados exhiben volatilidad agrupada semejante a los rendimientos, en la grafica siguiente de los residuales estandarizados, es decir el residual dividido por su respectiva desviación estándar, se exhibe un comportamiento más estable, salvo el caso de una observación que puede considerarse como una observación aberrante (outliers).



Además no se observa autocorrelación, que puede ser confirmada con la gráfica de autocorrelación del cuadrado de los residuales estandarizados, como se muestra.



Gráfica que comparada con la gráfica de los rendimientos al cuadrado del pre-análisis, nos indica que el modelo explica la heterocedasticidad en los rendimientos.

A manera de confirmación del anterior resultado se realiza la prueba de Box-Ljung de los residuos estandarizados, con valores de $p = 10, 15$ y 20 que arrojan los siguientes valores al 5% de:

| Valor de p | Valor de la estadística | Valor de la χ -cuadrada |
|--------------|-------------------------|------------------------------|
| 10 | 0.6361 | 18.3070 |
| 15 | 0.6598 | 24.9958 |
| 20 | 0.6698 | 31.4104 |

De acuerdo a Manganelli y Engle [52], el suponer una distribución normal para los residuos, la cual además debe ser independiente y estable para toda la muestra, es un requerimiento para estimar los parámetros desconocidos. Sin embargo, una alternativa que mejora el ajuste, dada la característica leptokúrtica, consiste en utilizar una función de distribución del tipo t-student.

Distribuciones Estadísticas para Modelar Rendimientos

| | |
|------------------|--|
| Normal | $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$ |
| t-Student | $f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{k}{2}\right)\sqrt{k}} \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}$ $\Gamma(z) = \int_0^{\infty} \lambda^{z-1} e^{-\lambda} d\lambda$ |

Hemos señalado que la hipótesis de normalidad parece no ser validada desde un punto de vista empírico. Sin embargo, tal como señalan Bollerslev y Woolridge [13] y Manganelli y Engel [52], el supuesto de normalidad no es tan restrictivo como parece. Los autores demostraron que la maximización de un modelo GARCH utilizando una distribución normal, es capaz de entregar estimadores consistentes aun cuando los residuos estandarizados no se distribuyan como una normal. La restricción que se requiere para conseguir este resultado es que la ecuación que representa a la varianza debe ser correctamente especificada.

GARCH multivariados

Hasta ahora, los modelos GARCH descritos, son modelos univariados, esto es, sólo se han descrito para pronosticar la volatilidad de un recurso, y por tanto no se considera la estimación de las covarianzas, las cuales juegan un papel importante en los portafolios con N recursos.

En el contexto de pronósticos de la volatilidad, considere el caso que involucra la variación de un portafolio compuesto de N recursos individuales,

$$r_t = \sum_{i=1}^N w_i r_{i,t}$$

La varianza condicional del portafolio para un periodo la podemos escribir como,

$$\sigma_t^2 = \sum_{i=1}^N w_i^2 \sigma_{i,t}^2 + 2 \sum_{i=1}^N w_i w_j \sigma_{ij,t} \quad (11)$$

siendo $\sigma_{i,t}^2$ la varianza de $r_{i,t}$ y $\sigma_{ij,t}$ la covarianza entre $r_{i,t}$ y $r_{j,t}$ matricialmente, lo podemos escribir como

$$\sigma_t^2 = \mathbf{w}' \Omega_t \mathbf{w}$$

donde Ω_t es la matriz de covarianzas de los rendimientos.

Una estimación para la varianza del rendimiento del portafolio basada en el modelo anterior requiere la estimación de $\frac{1}{2}N(N+1)$ elementos en la matriz de covarianzas para los recursos en el portafolio.

Una manera de construir estimaciones de la varianza del portafolio, sería la de construir una serie de tiempo artificial univariada de los rendimientos del portafolio en base a las ponderaciones o participaciones de los recursos del actual portafolio como,

$$r_j = \sum_{i=1}^N w_i^t r_{i,j} \quad j=1,2,\dots,t \quad \text{donde las } w_i^t, \text{ son las ponderaciones actuales del portafolio}$$

Un pronóstico para la varianza de los rendimientos de este portafolio artificial, asegura indirectamente que las covarianzas entre los recursos individuales reciben exactamente la misma ponderación como en la ecuación (11). A menos que las ponderaciones en el portafolio sean re-balanceadas constantemente, los rendimientos en el portafolio construido artificialmente generalmente serán diferentes de los rendimientos del portafolio

actual, dado que $r_k^t = \sum_{i=1}^N w_i^t r_{i,k} \neq \sum_{i=1}^N w_i^j r_{i,k} = r_k^j$ para $j \neq t$. Como tal, la construcción

de los pronósticos de la varianza para r_j^t requiere la estimación de un nuevo modelo (univariado) en cada periodo para reflejar la composición del portafolio en el momento t . Sin embargo, los modelos univariados de volatilidad son generalmente más fáciles de instrumentar que los multivariados, los cuales son altamente demandantes de tiempo computacional especialmente para valores grandes de N . Además, de que los cambios en las ponderaciones del portafolio actual de un periodo al próximo son probablemente pequeños, los valores de inicio para los parámetros en los modelos univariados periodo por periodo son posibles de los estimadores obtenidos en el periodo anterior. Esta simplificación requiere los rendimientos históricos para los diferentes recursos en el portafolio. Si eso no es posible, se podrían construir los precios históricos de manera artificial con un modelo de precios, o tomando los rendimientos de otros recursos con características similares; vea, por ejemplo, Andersen, Bollerslev, Christoffersen y Diebold [3].

Sin embargo, existen situaciones donde los pronósticos de las varianzas y covarianzas juegan un papel importante en comparaciones de riesgo, por lo que es necesario tener un modelo multivariado que tenga esta cualidad.

La clase de modelos GARCH multivariados fue introducida y estimada empíricamente por Bollerslev, Engle y Wooldridge [14]. Denotando el vector de medias y la matriz de covarianzas condicionales de R_t por $M_t = E(R_t|\mathfrak{F}_{t-1})$, $\Omega_t = \text{Var}(R_t|\mathfrak{F}_{t-1})$ respectivamente, la versión multivariada de la descomposición de los rendimientos se puede expresar como,

$$R_t = M_t + \Omega_t^{1/2} Z_t \quad Z_t \text{ i.i.d } E(Z_t) = 0 \quad \text{Var}(Z_t) = I \quad (12)$$

donde Z_t denota un vector con media cero y varianza unitaria. La raíz cuadrada de la matriz Ω_t no es única, pero cualquier operador que satisfaga la condición que $\Omega_t^{1/2} \Omega_t^{1/2} = \Omega_t$ deberá dar la misma matriz de covarianzas condicional.

El modelo multivariado GARCH(1,1) se puede definir de manera natural como,

$$Vr(\Omega_t) = C + A Vr(\varepsilon_{t-1}\varepsilon'_{t-1}) + B Vr(\Omega_{t-1}) \quad (13)$$

Donde $\varepsilon_t = \Omega_t^{1/2} Z_t$, y $Vr()$ denota el operador que convierte los $\frac{1}{2}N(N+1)$ elementos del triangulo inferior de la matriz, en un vector de dimensión $\frac{1}{2}N(N+1)$, y las matrices de los parámetros C , A , y B , son de dimensiones $\frac{1}{2}N(N+1) \times 1$, $\frac{1}{2}N(N+1) \times \frac{1}{2}N(N+1)$, y $\frac{1}{2}N(N+1) \times \frac{1}{2}N(N+1)$, respectivamente. Si $N=1$ el modelo en (13) es idéntico a la formulación en (8), pero para $N>1$ cada uno de los elementos en la matriz de covarianzas depende (linealmente) de todos los elementos retrasados en la matriz de covarianzas condicional, así como de los productos cruzados de todas las innovaciones retrasadas.

La formulación en (13) puede extenderse fácilmente para permitir influencias asimétricas, como en el modelo GJR o TGARCH, incluyendo los productos cruzados de los residuales. La generalizaron sería simplemente incluir $Vr(\min(\varepsilon_{t-1}, 0)\min(\varepsilon_{t-1}, 0)')$. Asimismo, otras variables exógenas podrían ser incluidas.

De su parte los pronósticos multi periodos para las varianzas y covarianzas condicionales del modelo lineal en (13) se generan con el siguiente proceso recursivo:

$$Vr(\Omega_{t+k}) = C + A Vr(F_{t+k-1}) + B Vr(\Omega_{t+k-1}) \quad (14)$$

Donde

$$F_{t+k-1} = \varepsilon_{t+k}\varepsilon'_{t+k} \quad k \leq 0$$

$$y F_{t+k-1} = \Omega_{t+k} \quad k \geq 1$$

Este proceso recursivo, asegura que la norma de todos los eigenvalores de $A+B$ sean menores a uno, y que los pronósticos de largo plazo para Ω_{t+k} converjan a “la matriz de covarianzas no-condicional” implicada por el modelo $(I-A-B)^{-1}C$ la cual cae a una tasa exponencial dictada por $(A+B)^k$. Estos resultados reflejan los resultados del modelo univariado.

Sin embargo, nada garantiza que la matriz de covarianzas no-condicional $(I-A-B)^{-1}C$ sea definida positiva, ni que del proceso recursivo se obtengan resultados para que la matriz de covarianzas sea definida positiva en el periodo k del pronóstico. De hecho, sin imponer restricciones adicionales a las matrices de los parámetros C , A , y B , los pronósticos para la matrices de covarianzas, probablemente, no sean definidas positivas.

Así mismo, el modelo multivariado GARCH(1,1) no restringido involucra un total de $\frac{1}{2}N^4 + N^3 + N^2 + \frac{1}{2}N$ parámetros que deben ser estimados, para $N=5$ el modelo tiene 465 parámetros, mientras que para $N=100$ hay un total de 51,010,050 parámetros, por lo que

la estimación de los parámetros es impráctica. Así, diversas simplificaciones se han diseñado para asegurar el que las matrices sean definidas positivas y se tenga un número manejable de parámetros.

Para que un elemento particular en la matriz de covarianza condicional sólo dependa del valor retrasado y el producto cruzado correspondiente de las innovaciones. El modelo (12) puede escribirse alternativamente en términos de los productos de Hadamard, o multiplicación de elemento por elemento, como

$$\Omega_t = C + A (\varepsilon_{t-1} \varepsilon'_{t-1}) + B (\Omega_{t-1}), \quad (15)$$

donde C, A, y B denotan matrices simétricas definidas positivas de dimensión $N \times N$. Este modelo reduce el número de parámetros a $3(N^2+N)/2$, donde, los pronósticos de la matrices de covarianza generadas para este modelo a través del proceso recursivo definido en (14), si garantizan que sean definidas positivas. Sin embargo, el modelo aún permanece con un número importante de parámetros a estimar, para $N=100$ hay 15,150 parámetros.

Una simplificación más fuerte, se obtiene restringiendo a todos los elementos de las matrices A y B en (15) como sigue,

$$\Omega_t = C + \alpha (\varepsilon_{t-1} \varepsilon'_{t-1}) + \beta (\Omega_{t-1}) \quad (16)$$

Esta representación del GARCH multivariado imita el proceso RiskMetrics excepto por la matriz C definida positiva, lo cual reduce el número de parámetros a $(N^2+N)/2 + 2$. Sin embargo, hay que probar que $\alpha + \beta < 1$, y que la matriz de covarianza no-condicional implicada por el modelo en (16) hace que $\Omega = (1 - \alpha - \beta)^{-1} C$, y en forma paralela la expresión para el modelo GARCH(1,1) univariado de la ecuación (9), del pronóstico para el k-periodo de Ω sea acorde a la siguiente fórmula,

$$\Omega_{t+k|t} = C + (\alpha + \beta)^{k-1} (\Omega_{t+1|t} - \Omega),$$

Esto contrasta grandemente con los pronósticos de RiskMetrics que no muestran ninguna reversión a la media, con la implicación de que el pronóstico multi-periodo de la covarianza, para (aproximadamente) rendimientos no correlacionados se escala con el horizonte de pronóstico.

Limitaciones de los modelos GARCH

- Los modelos GARCH son parte de la solución. Toda vez que los modelos GARCH aplicados a las series de rendimientos son sólo una herramienta para explicar la volatilidad, que apoya las decisiones financieras, pero que no deben constituirse como la función de decisión de las inversiones.
- Los modelos GARCH son especificaciones paramétricas que operan mejor bajo condiciones de mercado relativamente estables. Si bien GARCH es un modelo diseñado para modelar la varianza condicional variante en el tiempo, los modelos GARCH no capturan fenómenos altamente irregulares, tales como desplomes y

sus subsecuentes rebotes, así como otros eventos no anticipados que pueden llevar cambios estructurales significantes.

- Los modelos GARCH no capturan lo pesado de las colas de las series de rendimientos. La heterocedasticidad explica parte del comportamiento de las colas pesadas, pero no todo. Para compensar esta limitación, se pueden utilizar distribuciones de colas pesadas tales como la t de student

Finalmente, se encontró que en la literatura especializada, si bien la forma funcional del modelo GARCH(p,q) esta estandarizada, las restricciones sobre los parámetros a estimar, regularmente no son expresadas como parte del modelo o se expresan como desigualdades no lineales, mismas que son difíciles de aplicar en la practica.

Además también existen discrepancias en torno a cuando sí o cuando no los efectos positivos y negativos de las innovaciones (ε) afectan el proceso de la varianza condicional.

4. LOS MÉTODOS NO-PARAMÉTRICOS

4.1 Simulación Histórica

Uno de los métodos más comunes para la estimación de VaR, es el de Simulación Histórica. Este método simplifica el procedimiento para calcular el Valor en Riesgo, dado que no hace ningún supuesto de la distribución de los rendimientos del portafolio. La simulación histórica se basa en el concepto de ventanas rodantes (rolling windows), que generalmente contemplan un rango de 6 meses a dos años de observaciones.

El método de Simulación Histórica hace uso de la distribución empírica de los rendimientos. La premisa detrás de la Simulación Histórica es que los cambios potenciales en los factores de riesgo subyacentes son idénticos a los cambios observados en esos factores de riesgo sobre un periodo definido históricamente. Esto significa que realizar una Simulación Histórica es realizar un “muestreo” sobre los rendimientos del pasado, y aplicarlos al nivel actual de los factores de riesgo para obtener escenarios evaluados de riesgo. Finalmente se usan estos escenarios para obtener los escenarios de pérdidas y ganancias del portafolio. Este método tiene la ventaja de reflejar la distribución histórica multivariada de los factores de riesgo en los rendimientos, además de que incorpora la información sobre los rendimientos extremos siempre que estos sean incluidos en el periodo de muestra.

Para formalizar estas ideas, suponga que tenemos n factores de riesgo, por ejemplo, los tipos de cambio, las tasas de interés, los índices accionarios etc. y que se tiene un banco de datos que contiene los rendimientos de m+1 días de esos factores. Definimos la matriz (m+1)xn de rendimientos históricos como:

$$R = \begin{pmatrix} r_{t,1} & r_{t,2} & \dots & r_{t,n} \\ r_{t-1,1} & r_{t-1,2} & \dots & r_{t-1,n} \\ \vdots & \vdots & & \vdots \\ r_{t-m,1} & r_{t-m,2} & \dots & r_{t-m,n} \end{pmatrix}$$

Entonces, como cada escenario corresponde a un día de rendimientos históricos, podemos pensar que un escenario específico r corresponde a una fila de R.

Ahora, si tenemos M recursos en un portafolio donde el valor presente de cada recurso es una función de los n factores de riesgo $V_j(P)$, con $j = 1, \dots, M$ y $P = (P_1, P_2, \dots, P_n)$, nosotros podemos obtener el Valor en Riesgo para el portafolio para T días como sigue:

1. Tome un renglón r de R correspondiente a un escenario de rendimientos de los factores de riesgo.
2. Calcular el precio de cada factor de riesgo a T días con la fórmula $P_T = P_0 \exp(r\sqrt{T})$.
3. Evalué cada instrumento usando los actuales precios P_0 y los precios P_T del escenario a T días.
4. Encuentre las pérdidas y ganancias del portafolio como:

$$\sum_j V_j(P_T) - V_j(P_0)$$

5. Ordene la serie de pérdidas y ganancias de menor a mayor y de acuerdo al nivel de confianza, obtenga el VaR correspondiente.

Cabe hacer notar que en el paso 2, se multiplica el escenario de un día de rendimiento (r) por \sqrt{T} para obtener un escenario de T días de los rendimientos. Esto garantiza que la volatilidad del rendimiento es múltiplo de la raíz cuadrada del horizonte de tiempo, lo cual es una regla práctica.

Ejemplo

Suponga que un inversionista tiene un portafolio que consiste en una posición de dinero en efectivo de un millón de dólares y 13,000 acciones de Telmex. El tipo de cambio actual es 10.5395 pesos por dólar y el precio actual de cada acción de Telmex es 377.517533 pesos. El valor actual del portafolio es entonces de 15,447,227 de pesos.

El Valor del portafolio (pesos)

| | |
|----------|------------|
| Efectivo | 10,539,500 |
| Acciones | 4,907,727 |
| Suma | 15,447,227 |

Si se tiene la matriz de rendimientos del tipo de cambio y de las acciones (factores de riesgo) en el periodo del 1/12/05 al 29/12/05, estos serán los escenarios que se aplican a la composición actual en pesos del portafolio como se muestra en el cuadro siguiente.

| Fecha | Rendimiento Pesos (%) | Rendimiento Telmex (%) | Valor pesos | valor acciones | Valor del Portafolio |
|------------|-----------------------|------------------------|-------------|----------------|----------------------|
| 01/12/2005 | -0.4652 | 0.11906 | -48,912 | 5,847 | -43,065 |
| 02/12/2005 | -0.4339 | 0.11910 | -45,630 | 5,849 | -39,781 |
| 05/12/2005 | 0.1051 | 0.13289 | 11,079 | 6,526 | 17,606 |
| 06/12/2005 | -0.4641 | 0.04429 | -48,799 | 2,174 | -46,625 |
| 07/12/2005 | 0.3781 | 0.49164 | 39,928 | 24,188 | 64,116 |
| 08/12/2005 | 0.4386 | 0.00133 | 46,323 | 65 | 46,388 |
| 09/12/2005 | 1.2902 | 0.23030 | 136,857 | 11,316 | 148,173 |
| 13/12/2005 | -0.2868 | -0.09064 | -30,188 | -4,446 | -34,634 |
| 14/12/2005 | 0.8861 | 0.13155 | 93,801 | 6,460 | 100,262 |
| 15/12/2005 | 0.1819 | 0.01684 | 19,185 | 827 | 20,012 |
| 16/12/2005 | 0.0838 | 0.02751 | 8,839 | 1,350 | 10,189 |
| 19/12/2005 | 0.0884 | 0.16206 | 9,322 | 7,960 | 17,282 |
| 20/12/2005 | -0.6205 | -0.02077 | -65,195 | -1,019 | -66,214 |
| 21/12/2005 | -0.5349 | 0.00275 | -56,229 | 135 | -56,094 |
| 22/12/2005 | 0.3100 | 0.02084 | 32,728 | 1,023 | 33,751 |
| 23/12/2005 | 0.2810 | 0.09296 | 29,661 | 4,565 | 34,225 |
| 26/12/2005 | 0.0748 | 0.06779 | 7,887 | 3,328 | 11,215 |
| 27/12/2005 | 0.4988 | 0.10333 | 52,707 | 5,074 | 57,781 |
| 28/12/2005 | 0.2462 | 0.02551 | 25,977 | 1,252 | 27,230 |
| 29/12/2005 | -1.0540 | 0.24822 | -110,499 | 12,197 | -98,303 |

Los valores de la columna de la posición en pesos se calculan, por ejemplo, tomando los rendimientos históricos del 01-dic-2005, la posición en pesos habrá ganado -0.4652% y sus pérdidas y ganancias habrán sido $10,539,500 * \left[e^{-0.004652} - 1 \right] = -48,912$ pesos de igual manera, se calculan las pérdidas y ganancias para la posición de acciones como $4,907,727 \left[e^{0.0011906} - 1 \right] = 5,847$. Para calcular las pérdidas y ganancias de portafolio total, sumamos simplemente las pérdidas y ganancias individuales de cada posición.

Entonces, para encontrar el VaR por Simulación Histórica, se ordenan estos 20 valores de menor a mayor y con un nivel de confianza del 5% el valor seleccionado es de -98,303, esto es, el 5% de 20 que es igual a primer resultado.

| | | | | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|--------|--------|
| -98,303 | -66,214 | -56,094 | -46,625 | -43,065 | -39,781 | -34,634 | 10,189 | 11,215 | 17,282 | 17,606 |
| 20,012 | 27,230 | 33,751 | 34,225 | 46,388 | 57,781 | 64,116 | 100,262 | 148,173 | | |

En este método, el conjunto de valores finales, da lugar a la distribución de pérdidas y ganancias del portafolio, que se utiliza para obtener el VaR del portafolio para un determinado nivel de confianza. Por ejemplo, si el período histórico utilizado es de 100 días, obtendremos 100 posibles valores finales del portafolio y por lo tanto 100 posibles valores de las ganancias y pérdidas del portafolio. El VaR con un nivel de confianza del 95% será simplemente el quinto peor resultado de la distribución de pérdidas y ganancias observado.

Finalmente, las estimaciones del VaR basadas en la Simulación Histórica pueden presentar saltos predecibles, debido a la discrecionalidad de rendimientos extremos. Para ver el por qué, supongamos que se calcula el VaR de un portafolio con una ventana rodante de 180 días y que el rendimiento de hoy es un número negativo grande. Es fácil predecir que los VaR estimados serán mayores al VaR de hoy. El mismo efecto (invertido) reaparecerá después de 180 días, cuando la observación grande salga de la ventana.

Así mismo, se observa que aun cuando este procedimiento no hace ningún supuesto explícito sobre la distribución de los rendimientos del portafolio, un supuesto implícito está oculto detrás de este procedimiento: la distribución de los rendimientos del portafolio no cambia dentro de las ventanas. Este supuesto implica varios problemas.

Primero, este método es lógicamente inconsistente. Si se asume que todos los rendimientos dentro de la ventana deben tener la misma distribución, entonces la lógica consecuencia es que todos los rendimientos de la serie de tiempo deben tener la misma distribución: si r_{t-v}, \dots, r_t y $r_{t+1-v}, \dots, r_{t+1}$ son i.i.d, entonces, por la propiedad transitiva, también r_{t+1} y r_{t-v} deben ser i.i.d., Segundo, el estimador empírico del cuantil, es consistente solo si v , el tamaño de la ventana, tiende a infinito. El tercer problema involucra la longitud de la ventana. Éste es un problema muy delicado, dado que el pronóstico de VaR bajo este procedimiento sólo es significativo si los datos históricos usados en el cálculo tienen (aproximadamente) la misma distribución. Sin embargo, de las descripciones de Fama y Mandelbrot, se tiene que existe volatilidad agrupada, la cual no siempre es fácil de identificar.

Por lo anterior, se tiene que la longitud de la ventana debe satisfacer dos propiedades contradictorias: debe ser bastante grande para hacer la inferencia estadística consistente, y no debe ser demasiado grande, para evitar el riesgo de tomar las observaciones fuera del actual grupo de volatilidad. Es claro que no es fácil solucionar este problema.

Una variación interesante del método de la Simulación Histórica, es el método híbrido propuesto por Boudoukh, Richardson y Whitelaw [15]. Este híbrido combina la parte de la metodología de RiskMetrics y la metodología de la Simulación Histórica, aplicando ponderaciones exponenciales a los rendimientos pasados del portafolio. Para ello se siguen los siguientes tres pasos. Primero, a cada uno de los k más recientes escenarios del portafolio, $e_t, e_{t-1}, \dots, e_{t-k+1}$, se le asocia la ponderación $\frac{1-\lambda}{1-\lambda^k}, \left(\frac{1-\lambda}{1-\lambda^k}\right)\lambda, \dots, \left(\frac{1-\lambda}{1-\lambda^k}\right)\lambda^{k-1}$,

respectivamente.⁴ Segundo, los escenarios son ordenados en forma ascendente, junto con su ponderación. Tercero, para encontrar el c% VaR del portafolio, sume las ponderaciones correspondientes hasta que el c% se alcanza, empezando por el escenario más bajo. El VaR del portafolio es el escenario que corresponde a la última ponderación usada en la suma anterior.⁵

Ejemplo, tomando los mismos datos del ejemplo anterior y con $k = 20$ se tiene lo siguiente:

| Fecha | Rendimiento pesos | Rendimiento Telmex | Valor pesos | Valor Telmex | Total | $\lambda = 0.95$ | $\lambda = 0.99$ |
|------------|-------------------|--------------------|-------------|--------------|---------|------------------|------------------|
| 01/12/2005 | -0.4652 | 0.11906 | -48,912 | 5,847 | -43,065 | 0.0294 | 0.0454 |
| 02/12/2005 | -0.4339 | 0.11910 | -45,630 | 5,849 | -39,781 | 0.0310 | 0.0458 |
| 05/12/2005 | 0.1051 | 0.13289 | 11,079 | 6,526 | 17,606 | 0.0326 | 0.0463 |
| 06/12/2005 | -0.4641 | 0.04429 | -48,799 | 2,174 | -46,625 | 0.0343 | 0.0468 |
| 07/12/2005 | 0.3781 | 0.49164 | 39,928 | 24,188 | 64,116 | 0.0361 | 0.0472 |
| 08/12/2005 | 0.4386 | 0.00133 | 46,323 | 65 | 46,388 | 0.0380 | 0.0477 |
| 09/12/2005 | 1.2902 | 0.23030 | 136,857 | 11,316 | 148,173 | 0.0400 | 0.0482 |
| 13/12/2005 | -0.2868 | -0.09064 | -30,188 | -4,446 | -34,634 | 0.0421 | 0.0487 |
| 14/12/2005 | 0.8861 | 0.13155 | 93,801 | 6,460 | 100,262 | 0.0443 | 0.0492 |
| 15/12/2005 | 0.1819 | 0.01684 | 19,185 | 827 | 20,012 | 0.0467 | 0.0497 |
| 16/12/2005 | 0.0838 | 0.02751 | 8,839 | 1,350 | 10,189 | 0.0491 | 0.0502 |
| 19/12/2005 | 0.0884 | 0.16206 | 9,322 | 7,960 | 17,282 | 0.0517 | 0.0507 |
| 20/12/2005 | -0.6205 | -0.02077 | -65,195 | -1,019 | -66,214 | 0.0544 | 0.0512 |
| 21/12/2005 | -0.5349 | 0.00275 | -56,229 | 135 | -56,094 | 0.0573 | 0.0517 |
| 22/12/2005 | 0.3100 | 0.02084 | 32,728 | 1,023 | 33,751 | 0.0603 | 0.0522 |
| 23/12/2005 | 0.2810 | 0.09296 | 29,661 | 4,565 | 34,225 | 0.0635 | 0.0528 |
| 26/12/2005 | 0.0748 | 0.06779 | 7,887 | 3,328 | 11,215 | 0.0668 | 0.0533 |
| 27/12/2005 | 0.4988 | 0.10333 | 52,707 | 5,074 | 57,781 | 0.0703 | 0.0538 |
| 28/12/2005 | 0.2462 | 0.02551 | 25,977 | 1,252 | 27,230 | 0.0740 | 0.0544 |
| 29/12/2005 | -1.0540 | 0.24822 | -110,499 | 12,197 | -98,303 | 0.0779 | 0.0549 |

⁴ El papel del término $(1-\lambda)/(1-\lambda^k)$ es simplemente el asegurar que las ponderaciones sumen 1

⁵ Para lograr exactamente el c% de la distribución, se interpolan linealmente los puntos adyacentes

Los valores ordenados para $\lambda = 0.95$ son

| | | | | | | | | | | |
|-------------|---------|---------|---------|---------|---------|---------|---------|--------|---------|---------|
| Escenario | -98,303 | -66,214 | -56,094 | -46,625 | -43,065 | -39,781 | -34,634 | 10,189 | 11,215 | 17,282 |
| Ponderación | 0.0779 | 0.0544 | 0.0573 | 0.0343 | 0.0294 | 0.0310 | 0.0421 | 0.0491 | 0.0668 | 0.0517 |
| Escenario | 17,606 | 20,012 | 27,230 | 33,751 | 34,225 | 46,388 | 57,781 | 64,116 | 100,262 | 148,173 |
| Ponderación | 0.0326 | 0.0467 | 0.0740 | 0.0603 | 0.0635 | 0.0380 | 0.0703 | 0.0361 | 0.0443 | 0.0400 |

Donde el VaR sería igual a -98,303

Los valores ordenados para $\lambda = 0.99$ son

| | | | | | | | | | | |
|-------------|---------|---------|---------|---------|---------|---------|---------|--------|---------|---------|
| Escenario | -98,303 | -66,214 | -56,094 | -46,625 | -43,065 | -39,781 | -34,634 | 10,189 | 11,215 | 17,282 |
| Ponderación | 0.0549 | 0.0512 | 0.0517 | 0.0468 | 0.0454 | 0.0458 | 0.0487 | 0.0502 | 0.0533 | 0.0507 |
| Escenario | 17,606 | 20,012 | 27,230 | 33,751 | 34,225 | 46,388 | 57,781 | 64,116 | 100,262 | 148,173 |
| Ponderación | 0.0463 | 0.0497 | 0.0544 | 0.0522 | 0.0528 | 0.0477 | 0.0538 | 0.0472 | 0.0492 | 0.0482 |

Donde el valor de VaR es -98,303

Los valores obtenidos para el VaR en los tres casos son iguales, en este ejemplo, dado que el tamaño de la ventana es muy pequeño.

Esta aportación constituye una mejora significativa de la metodología, dado que simplifica los supuestos en el método de Riskmetrics e incorpora una especificación más flexible que el proceso de la Simulación Histórica. Para entender mejor los supuestos detrás de este método, note lo siguiente:

$$\hat{q}_{t+1,c} = \sum_{j=t-k+1}^t e_j I\left(\sum_{i=1}^k f_i(\lambda; k) I(e_{t+1-i} \leq e_j)\right) = c$$

dónde $I(\bullet)$ es la función indicadora y las $f_i(\lambda; K)$ son las ponderaciones asociadas con el escenario e_i . Si $f_i(\lambda; K) = 1/K$ se tendría el estimador de la Simulación Histórica. La principal diferencia entre estos métodos está en la especificación del cuantil. En el método de Simulación Histórica, a cada escenario se le da la misma ponderación, mientras que en el método híbrido, escenarios diferentes tienen ponderaciones diferentes.

Boudoukh, Richardson y Whitelaw [15] dan a λ el valor de 0.97 y 0.99, pero en su argumentación ningún método es propuesto para los valores de este parámetro desconocido, por lo que se puede concluir que dada la combinación de los métodos, se tomó el procedimiento de Riskmetrics para calcular el valor de λ . Los modelos de Simulación Histórica e Híbrido, son simplemente un caso especial del modelo CAViaR, por lo que se puede usar la regresión de cuantiles para estimar el parámetro λ y la longitud de la ventana.⁶

4.2 Simulación por Monte Carlo.

Una metodología más sofisticada e intensiva en tiempo computacional es la Simulación de Monte Carlo debida a Glasserman, Heidelberger y Shahabuddin [33], [34] y [35]. Esta

⁶ Aunque podría haber algunos problemas teóricos, debido a la discontinuidad de la función indicadora

consiste, en la generación de múltiples escenarios para los rendimientos de un activo (o activos) con un horizonte predefinido, como por ejemplo, un día, una semana o un mes.

Antes de pasar a la metodología, durante el desarrollo de la misma se considerará lo siguiente:

S = el vector de factores de riesgo

Δt = el horizonte de VaR

ΔS = el cambio en los factores de riesgo sobre Δt

L = la pérdida en valor del portafolio es el resultado del cambio ΔS sobre Δt .

La pérdida L es la diferencia entre el valor actual del portafolio y el valor del portafolio al final del horizonte Δt , cuando el vector de factores de riesgo se mueve de S a $S+\Delta S$.

Monte Carlo para VaR

Los principales pasos en el método básico de Monte Carlo para estimar la probabilidad de pérdida son:

1. Generar N escenarios muestreando cambios en los factores de riesgo $\Delta S(1), \dots, \Delta S(N)$ sobre el horizonte Δt ;
2. Evaluar el portafolio al final de horizonte Δt en los escenarios $S+\Delta S(1), \dots, S+\Delta S(N)$; y determinar las pérdidas $L(1), \dots, L(N)$ substrayendo de la evaluación el valor actual del portafolio, en cada escenario;
3. Calcular la fracción de escenarios en que las pérdidas exceden x: como

$$N^{-1} \sum_{i=1}^N I(L(i) < x) \text{ donde } I(L(i) < x) = 1 \text{ si } L(i) < x \text{ y } 0 \text{ en otro caso.}$$

En otras palabras el método de Monte Carlo requiere de

1. La generación de escenarios

Usando la volatilidad y la correlación estimada para los recursos subyacentes en el portafolio, se producen un número grande de escenarios de precios futuros.

2. valuación del portafolio

Para cada escenario, se calcula el valor del portafolio.

3. El resumen

Se obtiene como resultado de la simulación, la distribución de pérdidas y ganancias del portafolio de donde se obtiene el VaR.

El primer paso requiere algunos supuestos sobre los datos del mercado. En la Simulación Histórica, los $\Delta S(i)$ son los cambios observados en los datos del mercado sobre los n periodos pasados de longitud Δt . Esto implícitamente supone que los cambios futuros en los factores de riesgo se parecerán a los cambios del pasado. Alternativamente, se puede utilizar un modelo estadístico de los datos históricos para seleccionar una distribución con parámetros estimados para describir los cambios futuros. Un supuesto simple y ampliamente usado es que los cambios ΔS sobre un horizonte corto Δt son descritos

como una distribución normal multivariada $N(0;\Sigma_s)$. La matriz de covarianzas Σ_s normalmente se estima de los rendimientos, usando una matriz de covarianzas muestral, a través del método de promedios móviles ponderados exponencialmente, o usando una estimación GARCH; vea, por ejemplo, (Alexander, C. (1999) Volatility and correlation: measurement, models, and applications" in Risk Management and Analysis, Vol 1 pp125-172)

Dada una matriz de covarianzas Σ_s y el supuesto de cambios en los factores de riesgo normalmente distribuidos, es simple la manera de generar las muestras de ΔS requeridas en la simulación. Factorizando la matriz de covarianza encontrando una matriz C tal que $CC' = \Sigma_s$ y entonces:

$$\Delta S = CZ \tag{17}$$

donde Z es un vector de variables aleatorias independientes de una normal estándar. Por ejemplo, suponga que Σ_s , es positiva definida, la factorización de Cholesky produce una matriz triangular inferior C para la cual $CC' = \Sigma_s$.

Para ilustrar la metodología, considere un portafolio que contiene dos recursos: una posición de dinero en efectivo de un millón de dólares y 13,000 acciones de Telmex. El tipo de cambio actual es 10.5395 pesos por dólar y el precio actual de cada acción de Telmex es 377.517533 pesos, de donde el valor actual del portafolio es de 15,447,227.

Entonces, el primer paso es la generación de escenarios, por lo que para obtener los cambios en los factores de riesgo, para ello recordemos que en el capítulo 2, se supone para todos los recursos subyacentes el modelo de precios como un modelo lognormal, en donde para un horizonte de pronóstico de t días, con un precio inicial P_0 y un estimado de la volatilidad de un día (σ), el precio en el horizonte de tiempo se calcula con la siguiente formula:

$$P_t = P_0 \exp(\sigma Z \sqrt{T})$$

Donde Z es una variable aleatoria normal $(0,1)$, entonces el primer paso es generar variables aleatorias normales $(0,1)$. Para nuestro ejemplo se generan mil números aleatorios que siguen una normal $(0,1)$ para cada factor de riesgo, en el siguiente cuadro se muestran los primeros 35 datos de cada serie.

Siguiendo nuestro ejemplo, diremos que las volatilidades diarias de esos dos factores de riesgo son de 0.0042 y de 0.008 para el tipo de cambio y acciones respectivamente y que se tiene una correlación de -0.17 entre ambos factores.

Entonces siguiendo con la metodología, tenemos que calcular variables aleatorias que tengan las características anteriores, a partir de los números aleatorios normales generados, esto es: $\Delta S=Y = CZ$, donde C es una matriz tal que $C'C = \Sigma_s$, en este ejemplo usaremos la descomposición de Cholesky, por que es más práctica, ya que deja una matriz triangular inferior que facilita el cálculo de las nuevas variables. Cabe también señalar que en nuestro ejemplo usaremos la matriz de correlación en lugar de la de covarianza.

Se tiene que

$$\Sigma_S = \begin{pmatrix} 1, \rho \\ \rho, 1 \end{pmatrix} = \begin{pmatrix} 1, -0.17 \\ -0.17, 1 \end{pmatrix} \text{ entonces } C = \begin{pmatrix} 1, 0 \\ \rho, \sqrt{1-\rho^2} \end{pmatrix} = \begin{pmatrix} 1, 0 \\ -0.17, 0.9854 \end{pmatrix}$$

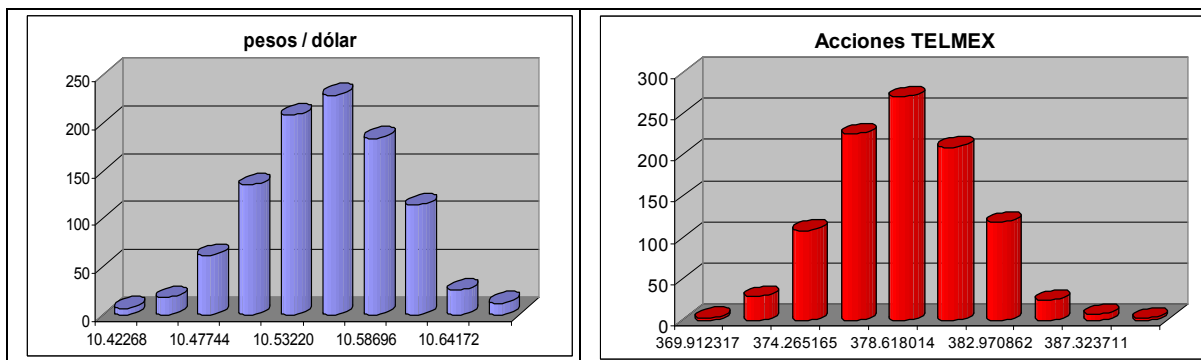
$$\text{De donde } \Delta S = Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = C \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 1, 0 \\ -0.17, 0.9854 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} Z_1 \\ -0.17Z_1 + 0.9854Z_2 \end{pmatrix}$$

Entonces los precios de los factores en riesgo, para un día (T=1) son calculados como:

$$\begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = \begin{pmatrix} P_{0,1} e^{\sigma_1 Y_1} \\ P_{0,2} e^{\sigma_2 Y_2} \end{pmatrix} = \begin{pmatrix} 10.5395 e^{0.0042 Y_1} \\ 377.517533 e^{0.008 Y_2} \end{pmatrix}$$

| Z ₁ | Z ₂ | Y ₁ | Y ₂ | Pesos/ dólar | Precio Acción | Valor efectivo | Valor acciones | Valor Portafolio | Valor portafolio -Valor Actual |
|----------------|----------------|----------------|----------------|-----------------|------------------|-------------------|-------------------|---------------------|-----------------------------------|
| -3.1 | -1.67 | -3.1 | -1.11862 | 10.40317 | 374.154221 | 10,403,165.18 | 4,864,004.88 | 15,267,170.05 | -180,056.95 |
| -1.93 | -2.66 | -1.93 | -2.29306 | 10.45441 | 370.655293 | 10,454,412.14 | 4,818,518.81 | 15,272,930.95 | -174,296.05 |
| -2.57 | -1.84 | -2.57 | -1.37624 | 10.42635 | 373.383904 | 10,426,348.41 | 4,853,990.76 | 15,280,339.17 | -166,887.83 |
| -3.28 | -0.88 | -3.28 | -0.30955 | 10.39530 | 376.583799 | 10,395,303.35 | 4,895,589.39 | 15,290,892.74 | -156,334.26 |
| -1.27 | -2.6 | -1.27 | -2.34614 | 10.48343 | 370.497943 | 10,483,431.97 | 4,816,473.26 | 15,299,905.23 | -147,321.77 |
| -1.02 | -2.54 | -1.02 | -2.32952 | 10.49445 | 370.547219 | 10,494,445.36 | 4,817,113.85 | 15,311,559.21 | -135,667.79 |
| -2.22 | -1.25 | -2.22 | -0.85435 | 10.44169 | 374.946074 | 10,441,686.42 | 4,874,298.96 | 15,315,985.38 | -131,241.62 |
| -1.64 | -1.77 | -1.64 | -1.46536 | 10.46715 | 373.117786 | 10,467,153.37 | 4,850,531.21 | 15,317,684.58 | -129,542.42 |
| -2.6 | -0.84 | -2.6 | -0.38574 | 10.42503 | 376.354352 | 10,425,034.78 | 4,892,606.57 | 15,317,641.35 | -129,585.65 |
| -1.77 | -1.61 | -1.77 | -1.28559 | 10.46144 | 373.654757 | 10,461,439.87 | 4,857,511.84 | 15,318,951.70 | -128,275.30 |
| -1.25 | -2.07 | -1.25 | -1.82728 | 10.48431 | 372.039038 | 10,484,312.62 | 4,836,507.49 | 15,320,820.11 | -126,406.89 |
| -1.82 | -1.45 | -1.82 | -1.11943 | 10.45924 | 374.151791 | 10,459,243.19 | 4,863,973.28 | 15,323,216.47 | -124,010.53 |
| -1.26 | -1.98 | -1.26 | -1.73689 | 10.48387 | 372.308152 | 10,483,872.29 | 4,840,005.97 | 15,323,878.26 | -123,348.74 |
| -3.17 | -0.14 | -3.17 | 0.40094 | 10.40011 | 378.730384 | 10,400,107.09 | 4,923,494.99 | 15,323,602.09 | -123,624.91 |
| -1.09 | -2.05 | -1.09 | -1.83477 | 10.49136 | 372.016740 | 10,491,360.44 | 4,836,217.62 | 15,327,578.06 | -119,648.94 |
| -1.87 | -1.28 | -1.87 | -0.94341 | 10.45705 | 374.679021 | 10,457,046.98 | 4,870,827.28 | 15,327,874.26 | -119,352.74 |
| -1.62 | -1.52 | -1.62 | -1.22241 | 10.46803 | 373.843682 | 10,468,032.65 | 4,859,967.87 | 15,328,000.52 | -119,226.48 |
| -2.02 | -1.06 | -2.02 | -0.70112 | 10.45046 | 375.405968 | 10,450,461.12 | 4,880,277.58 | 15,330,738.70 | -116,488.30 |
| -2.41 | -0.68 | -2.41 | -0.26037 | 10.43336 | 376.731991 | 10,433,357.28 | 4,897,515.89 | 15,330,873.16 | -116,353.84 |
| -1.15 | -1.87 | -1.15 | -1.64720 | 10.48872 | 372.575398 | 10,488,716.95 | 4,843,480.18 | 15,332,197.13 | -115,029.87 |
| -2.78 | -0.27 | -2.78 | 0.20654 | 10.41716 | 378.141834 | 10,417,156.43 | 4,915,843.85 | 15,333,000.28 | -114,226.72 |
| -0.69 | -2.29 | -0.69 | -2.13927 | 10.50900 | 371.111622 | 10,509,000.74 | 4,824,451.08 | 15,333,451.83 | -113,775.17 |
| -2.26 | -0.7 | -2.26 | -0.30558 | 10.43993 | 376.595766 | 10,439,932.36 | 4,895,744.95 | 15,335,677.32 | -111,549.68 |
| -1.05 | -1.87 | -1.05 | -1.66420 | 10.49312 | 372.524731 | 10,493,123.14 | 4,842,821.51 | 15,335,944.65 | -111,282.35 |
| -0.22 | -2.7 | -0.22 | -2.62318 | 10.52977 | 369.677710 | 10,529,766.00 | 4,805,810.23 | 15,335,576.23 | -111,650.77 |
| -1.51 | -1.4 | -1.51 | -1.12286 | 10.47287 | 374.141524 | 10,472,870.00 | 4,863,839.81 | 15,336,709.81 | -110,517.19 |
| -0.64 | -2.22 | -0.64 | -2.07879 | 10.51121 | 371.291218 | 10,511,207.87 | 4,826,785.83 | 15,337,993.70 | -109,233.30 |
| -1.35 | -1.51 | -1.35 | -1.25845 | 10.47991 | 373.735893 | 10,479,910.13 | 4,858,566.61 | 15,338,476.75 | -108,750.25 |
| -1.52 | -1.34 | -1.52 | -1.06204 | 10.47243 | 374.323623 | 10,472,430.15 | 4,866,207.10 | 15,338,637.24 | -108,589.76 |
| -1.58 | -1.28 | -1.58 | -0.99271 | 10.46979 | 374.531277 | 10,469,791.43 | 4,868,906.60 | 15,338,698.03 | -108,528.97 |

| | | | | | | | | | |
|-------|-------|-------|----------|----------|------------|---------------|--------------|---------------|-------------|
| -2.85 | -0.05 | -2.85 | 0.43523 | 10.41409 | 378.834280 | 10,414,094.24 | 4,924,845.64 | 15,338,939.87 | -108,287.13 |
| -0.7 | -2.12 | -0.7 | -1.97005 | 10.50856 | 371.614352 | 10,508,559.38 | 4,830,986.58 | 15,339,545.95 | -107,681.05 |
| -1.86 | -0.97 | -1.86 | -0.63964 | 10.45749 | 375.590671 | 10,457,486.19 | 4,882,678.72 | 15,340,164.91 | -107,062.09 |
| -1.22 | -1.59 | -1.22 | -1.35939 | 10.48563 | 373.434240 | 10,485,633.73 | 4,854,645.12 | 15,340,278.84 | -106,948.16 |
| -1.51 | -1.3 | -1.51 | -1.02432 | 10.47287 | 374.436584 | 10,472,870.00 | 4,867,675.59 | 15,340,545.59 | -106,681.41 |



A diferencia del ejemplo que nos ocupa, la evaluación del portafolio en cada escenario, puede no ser lineal, para un portafolio de productos derivados complejos, y por tanto puede ser muy consumidora de tiempo computacional, ya que los instrumentos individuales requieren ejecutar rutinas numéricas para el cálculo de precios o incluso procesos adicionales de Monte Carlo para las estimaciones de precios, como podría ser el caso de que en lugar de tener acciones de Telmex o de efectivo se tuvieran opciones sobre bonos gubernamentales o de otro tipo, lo que significa utilizar fórmulas como la de Black and Scholes para calcular el precio de la opción.

Continuando con nuestro ejemplo y una vez calculado el valor del portafolio para cada escenario, así como su diferencia con el valor presente, tenemos que ordenar estos últimos valores de menor a mayor y obtener la distribución de pérdidas y ganancias, de donde se puede obtener el VaR.

| Cuantil | VaR |
|-----------|-------------------|
| 1% | -128,275.30 |
| 2.5% | -111,650.77 |
| 5% | -98,183.00 |
| 10% | -71,578.13 |
| 25% | -34,755.78 |
| 50% | 2,622.04 |
| 75% | 38,658.26 |
| 90% | 69,719.37 |
| 95% | 87,751.78 |
| 97.5% | 105,447.61 |
| 99% | 116,325.40 |

La Aproximación Delta-gamma

Una alternativa para abatir el número de evaluaciones cuando se tienen recursos financieros complicados en el portafolio, como opciones o derivados, es usar una aproximación de como los cambios en los factores de riesgo determinan los cambios en el valor del portafolio.

La aproximación delta-gamma supone la disponibilidad de: 1) el vector δ de primeras derivadas parciales del valor del portafolio con respecto a los componentes del vector S de factores de riesgo; 2) la matriz Γ de segundas derivadas parciales, y 3) un escalar Θ dada la derivada parcial del valor del portafolio con respecto al tiempo.

De manera específica, podemos suponer que el cambio en el valor del portafolio del periodo t al periodo Δt , es $\Delta V = V(t+\Delta t) - V(t)$, de donde la pérdida se define como $L = \Delta V$ y el VaR (x), asociado con un nivel de confianza c esta dado por la relación $P(L < x) = 1-c$.

Como los cambios en el valor del portafolio son determinados por los cambios en los factores de riesgo, a través de una relación no lineal que supone tiene derivadas parciales de primer y segundo orden, entonces podemos aproximar estos cambios en el portafolio de acuerdo al teorema de Taylor, de la siguiente forma:

$$\Delta V = L \approx a_0 + \delta' \Delta S + \frac{1}{2} \Delta S' \Gamma \Delta S \quad \text{donde} \quad a_0 = \Theta \Delta t, \quad \Theta = \partial V / \partial t, \quad \delta_i = \partial V / \partial S_i \text{ y}$$

$\Gamma_{ij} = \partial^2 V / \partial S_i \partial S_j$, entonces dado que $\Delta S = CZ$, se tiene que

$$\begin{aligned} L &\approx a_0 + \delta' CZ + \frac{1}{2} (CZ)' \Gamma CZ \\ &= a_0 + b' Z + \frac{1}{2} Z' C' \Gamma CZ = a_0 + b' Z + Z' \Lambda Z \equiv a_0 + Q \end{aligned} \quad (18)$$

donde Z es un vector de variables aleatorias independientes distribuidas como una normal 0,1 y Λ es una matriz diagonal de la forma,

$$\Lambda = \begin{pmatrix} \lambda_1 & 0, \dots, 0 \\ 0, & \lambda_2, \dots, 0 \\ \vdots & \vdots \\ 0, & \dots, & \lambda_m \end{pmatrix}$$

con $\lambda_1 \geq \lambda_2 \dots \geq \lambda_m$ los eigenvalores de $1/2 \Gamma \Sigma$ y seleccionando C en (17) que satisfaga.

$$CC' = \Sigma \text{ y } 1/2 C' \Gamma C = \Lambda \quad (19)$$

El vector b' el término lineal de (18) está dado por : $b' = \delta' C$

Esta transformación logra dos importantes simplificaciones: reemplaza la correlación en los factores de riesgo ΔS con los elementos no-correlacionados de Z , y diagonaliza el término cuadrático en la aproximación. El vector ΔS se recupera de Z a través de (17), entonces se puede pensar en los elementos de Z como (hipotéticamente) el impulso subyacente de los factores de riesgo en los cambios en el mercado ΔS . Note que la matriz diagonal Λ captura la información sobre el portafolio (a través de Γ) y de la distribución de los factores de riesgo (a través de Σ_S).

Con estas simplificaciones es relativamente sencillo encontrar la función característica (transformada de Fourier) de la aproximación delta-gamma, o más precisamente, de la forma cuadrática Q en (18).

La cual se define como:

$$\psi(\theta) = \sum_{i=1}^m \frac{1}{2} \left(\frac{(\theta b_i)^2}{1 - 2\theta \lambda_i} - \ln(1 - 2\theta \lambda_i) \right) \quad (20)$$

La inversión transformada puede ahora usarse para calcular los valores de la distribución $P(Q < x)$. En (18), la distribución de pérdida puede aproximarse usando la siguiente expresión:

$$P(L < x) \approx P(Q < x - a_0).$$

4.2.1 Muestreo de Importancia Basado en la Aproximación Delta-gamma

La principal virtud de la aproximación delta-gamma es que puede calcularse rápidamente. Sin embargo, la exactitud de la aproximación no siempre es satisfactoria. La simulación de Monte Carlo es más exacta pero consume mucho más tiempo. Nuestro objetivo es usar la información contenida en la aproximación delta-gamma para acelerar la simulación de Monte Carlo y así aprovechar las aportaciones de los dos métodos.

La manera más simple de usar la aproximación delta-gamma en una simulación es instrumentarla como una variable de control. Estimando la probabilidad de pérdida $P(L < x)$, se produce un estimador de la forma:

$$\frac{1}{N} \sum_{i=1}^N I(L^{(i)} < x) - \beta \left[\frac{1}{N} \sum_{i=1}^N I(Q^{(i)} < x - a_0) - P(Q < x - a_0) \right]$$

Donde, $L^{(i)}$ son las pérdidas actuales calculadas en los N escenarios simulados y las $Q^{(i)}$ son las aproximaciones cuadráticas (vea (18)) calculadas en los mismos escenarios. La verdadera probabilidad $P(Q < x - a_0)$ se calcula a través de la inversión transformada (20). El término en los paréntesis cuadrados es el error observado en la simulación dentro de la aproximación delta-gamma; este error observado se usa para ajustar la estimación en la simulación de la verdadera pérdida del portafolio. El coeficiente β puede elegirse para minimizar la varianza del estimador combinado.

Este método se propuso independientemente en (Cárdenas y otros [17]) y puede probarse que existe una razonable reducción en la varianza. Note que el método de control variante usa la aproximación delta-gamma para ajustar el estimador estándar en particular, los escenarios usados se generan de la manera usual (es decir, como se presentó en el método básico de Monte Carlo). En contraste, el método descrito usa la aproximación delta-gamma antes que cualquier escenario sea generado y usa la aproximación para guiar el muestreo de escenarios.

El muestreo de Importancia

A través de (17), el problema de muestreo de los cambios en los factores de riesgo de mercado (ΔS), se transforma en un problema de muestrear el vector Z de las subyacentes variables aleatorias normales(0,1). En el muestreo de importancia (IS), se cambia la distribución de la cual se generan las variables subyacentes, para generar más muestras de regiones “importantes”. El presente trabajo se enfocara sobre métodos que cambian la distribución de Z de $N(0;1)$ (la normal multivariada estándar) a $N(\mu;\Sigma)$ (la normal multivariada con vector de medias μ y matriz de covarianza Σ).

La igualdad clave necesaria para el muestreo de importancia es

$$P(L < x) = E_{\mu, \Sigma} [I(Z)I(L < x)] \quad (21)$$

Con el subíndice μ, Σ en la expresión del lado derecho, se esta indicando que la esperanza es tomada con Z muestreado de $N(\mu; \Sigma)$ en lugar de su distribución original $N(0;1)$. Para corregir este cambio de distribución, debemos ponderar el indicador de pérdida $I(L < x)$ por la razón de verosimilitud.

$$I(Z) = \frac{e^{-\frac{1}{2}Z'Z}}{|\Sigma|^{1/2} e^{-\frac{1}{2}[(Z-\mu)'\Sigma^{-1}(Z-\mu)]}} \quad (22)$$

qué simplemente es la razón de las densidades $N(0;1)$ y $N(\mu; \Sigma)$ evaluadas en Z . En ambos lados de (21), la pérdida L se calcula de los cambios de mercado ΔS , los cuales son calculados de Z en la ecuación (17). De la ecuación (21) con una muestra Z de tamaño n de cualquier $N(\mu; \Sigma)$, es posible obtener un estimador insesgado de la probabilidad de pérdida, como:

$$\hat{P}_x = \frac{1}{n} \sum_{j=1}^n I(Z_j)I(L_j < x) \quad (23)$$

¿Cómo deben seleccionarse μ y Σ para producir un estimador con baja varianza (y de mayor precisión)? Al cambiar μ y Σ no cambian los resultados esperados, ya que comparar las varianzas es equivalente a comparar segundos momentos.

El segundo momento de (23) es

$$E_{\mu,\Sigma} \left[\left(I(Z)I(L < x) \right)^2 \right] = E \left[I(Z)I(L < x) \right] \quad (24)$$

la esperanza en el lado derecho de la ecuación es tomada con respecto a la original distribución $N(0;I)$, de lo cual se observa que la clave para reducir la varianza es hacer la razón de verosimilitud más pequeña cuando $L < x$. Equivalentemente, es deseable seleccionar μ y Σ para hacer escenarios con $L < x$ más probables bajo $N(\mu;\Sigma)$ que bajo $N(0;I)$.

Desafortunadamente, la expresión (22), aporta muy poco en la elección de μ y Σ para lograr este objetivo. Sin embargo, podemos usar la aproximación delta-gamma para conseguir que los escenarios tiendan a producir pérdidas grandes y usar esta información en la selección de μ y Σ .

Para ello, es necesario reescribir (18) más explícitamente como

$$L = a_0 + \sum_i b_i Z_i + \sum_i \lambda_i Z_i^2$$

y preguntarse, ¿Qué valores de Z tenderán a hacer (aproximadamente) la expresión de pérdida grande? La inspección de esta fórmula sugiere que las pérdidas grandes resultan de

- Valores positivos grandes de Z_i con $b_i > 0$;
- Valores negativos grandes de Z_i con $b_i < 0$;
- Valores grandes de $(Z_i)^2$ con $\lambda_i > 0$.

Lo anterior describe las regiones que dan mayor probabilidad bajo la distribución IS que bajo la distribución original. Lo que sugiere que se debe

- Aumentar la media de Z_i con $b_i > 0$;
- Disminuir la media de Z_i con $b_i < 0$;
- Aumentar la varianza de Z_i con $\lambda_i > 0$;

y además

- Disminuir la varianza de Z_i con $\lambda_i < 0$.

Lo anterior se logra en dos pasos. Primero se reduce la elección de μ y Σ a la elección de un escalar θ , y se especifica el valor de ese parámetro como sigue. Para cualquier $\theta > 0$ (y $\theta < 1/(2\lambda_1)$ si $\lambda_1 > 0$), se define,

$$\Sigma(\theta) = (I - 2\theta\Lambda)^{-1} \quad \mu(\theta) = \theta \Sigma(\theta) b \quad (25)$$

Con estos parámetros, Z_i es normal con media y varianza

$$\mu_i(\theta) = \frac{\theta b_i}{1 - 2\theta \lambda_i} \quad \sigma_i^2(\theta) = \frac{1}{1 - 2\theta \lambda_i} \quad (26)$$

y los Z_i permanecen independientes entre sí. Note que, con este tipo de IS, el muestreo sobre la distribución de Z_i es como se sugirió; es decir, si $\lambda_i > 0$, entonces la varianza de Z_i se incrementa resultando más muestras con valores grandes de Z_i^2 . La observación importante es que con ese cambio de distribución, la razón de verosimilitud se expresa como.

$$l(Z) = e^{-\theta Q + \psi(\theta)} \quad (27)$$

Aquí, ψ es la función introducida en (20) y puede interpretarse como una normalización constante. Lo notable de esta expresión es que la razón de verosimilitud -que en general podría depender del vector Z , como en (22)- ahora tiene el escalar Q como su único elemento estocástico.

El estimador asociado con esto la distribución IS es

$$e^{-\theta Q + \psi(\theta)} I(L < x)$$

donde Z es usado para calcular L y Q , generada ahora usando (21). Debe enfatizarse que este estimador es insesgado (en (21)) para la probabilidad exacta de pérdida $P(L < x)$, aunque involucra la aproximación delta-gamma.

Recordando de la presentación alrededor de (24) que un muestreo de importancia eficaz de la distribución, hace que la proporción de verosimilitud sea pequeña en los escenarios para los cuales $L < x$, se puede observar, basado en (18), que podemos esperar que cuando $L < x$ tendremos a menudo que $Q < x - a_0$; en particular, Q será típicamente grande cuando L lo sea y en este caso la razón de verosimilitud (21) tenderá a ser pequeña cuando $L < x$.

De la especificación del parámetro θ , y de las expresiones para la media y varianza se tiene que:

$$\frac{d}{d\theta} \psi(\theta) = E_{\mu(\theta), \Sigma(\theta)} [Q] \quad (28)$$

La expresión (25), define una familia de distribuciones exponencial con función generadora acumulada ψ ; entonces la ecuación (28) es un caso especial de una propiedad estándar de familias exponenciales. De donde la ecuación (28), señala que la derivada de ψ en θ da la esperanza de pérdida delta-gamma, cuando Z es tomada de $N(\mu(\theta); \Sigma(\theta))$. Entonces el objetivo de estimar $P(L < x) \approx P(Q < x - a_0)$, es seleccionar θ como θ_x , la solución de la siguiente ecuación:

$$\frac{d}{d\theta}\psi(\theta) = E_{\mu(\theta_x), \Sigma(\theta_x)}[Q] = x - a_0 \quad (29)$$

Si se muestrean los escenarios Z de $N(\mu(\theta); \Sigma(\theta))$, donde $L < x$ que eran casos raros anteriormente, ahora deberán ser "típicos", entonces el valor esperado de la pérdida aproximada $a_0 + Q$ es ahora x .

El Procedimiento

Resumiendo el proceso de muestreo. Suponemos la disponibilidad del vector delta del portafolio (δ) y la matriz gamma (Γ), qué también se requerirá para la aproximación delta-gamma.

1. Calcular C , satisfaciendo (17):

- a) Encontrar cualquier matriz A que cumpla con que $AA' = \Sigma$ s (por ejemplo, la factorización de Cholesky)
- b) Encontrar V , una matriz ortogonal ($VV' = I$) cuyas columnas son los eigenvectores de $1/2(A'\Gamma A)$ y Λ la matriz diagonal de eigenvalores asociados (dado que $1/2(A'\Gamma A) = V\Lambda V'$)
- c) Calcular $C = AV$ y $b = \delta'C$

2. Sea $\theta = \theta x$ la solución a (29)

3. Calcular $\Sigma(\theta) = (I - 2\theta\Lambda)^{-1}$ y $\mu(\theta) = \theta\Sigma(\theta)b$

4. Simulación:

- a) Genere $Z(1), \dots, Z(N)$ de $N(\mu(\theta); \Sigma(\theta))$
- b) Sea $\Delta S(i) = CZ(i)$, $i = 1, \dots, N$
- c) Calcule las pérdidas del portafolio $L(i)$ como resultado de los escenarios $\Delta S(i)$, $i = 1, \dots, N$
- d) Calcule $Q(i)$ para cada $Z(i)$, $i = 1, \dots, N$, usando (18)
- e) calcule la estimación del rendimiento como

$$\frac{1}{N} \sum_{i=1}^N e^{-\theta Q(i) + \psi(\theta)} I(L(i) < x) \quad (30)$$

Un rasgo importante de este método es que puede ser instrumentado sobre una versión existente de Monte Carlo. El centro del algoritmo -el cálculo de pérdidas del portafolio en cada escenario- es exactamente el mismo aquí como en el método básico de Monte

Carlo. Después de algún preproceso de los pasos 1-3, el algoritmo de muestreo de importancia difiere sólo en cómo se generan los escenarios y en cómo pondera los mismos en (30). Como con el método básico de Monte Carlo, la expresión en (30) podría calcularse fácilmente para múltiples valores del umbral de pérdida x , todos basados en un solo valor de θ .

El método presentado es probablemente eficaz, en el sentido de reducir la varianza substancialmente, entre el umbral de pérdida o incrementando el número de factores de riesgo. Estos resultados se establecen bajo la hipótesis de que la relación $L = a_0 + Q$ es una aproximación. Interpretamos estos resultados como la evidencia de que el método debe ser eficaz siempre que $a_0 + Q$ proporcione una aproximación razonable de L , aun cuando no sea suficientemente exacto como para reemplazar la simulación completa. La importancia de reducir la varianza en la simulación estimada es que reduce el número de escenarios requeridos para lograr la precisión deseada. Esto puede producir reducciones sustanciales de tiempo de cómputo, porque evaluar un portafolio en cada escenario es el paso que más tiempo consume en la estimación de la probabilidad de pérdida en el método de Monte Carlo.

Resultados prácticos de esta metodología, se pueden consultar en la documentación del tercer trimestre de 1996 de Riskmetrics Monitor, donde se observa que las diferencias en el cálculo entre los modelos Monte Carlo completo y la simplificación a través del acercamiento delta-gama, son aceptables, para cierto tipo de opciones.

4.2.2 Muestreo Estratificado

La ecuación (30) sugiere que para reducir la varianza se debe reducir la variabilidad en la muestra de la aproximación cuadrática Q . De hecho, si $L = a_0 + Q$, entonces eliminando la varianza debida a Q se elimina toda la varianza en (30), pero como a_0+Q sólo aproxima L , entonces reduciendo la variabilidad de Q debe producirse la reducción de la varianza.

Siguiendo sobre esta idea, se instrumenta el muestreo estratificado, en lugar de tomar $Q(i)$ aleatoria e independientemente, se debe asegurar que fracciones fijas de la muestra caigan en bloques específicos, toda vez que si se generan muestras $Q(i)$ como en el muestreo de importancia, no se puede esperar que exactamente fracciones iguales de la muestra caigan en cada uno de los bloques; porque el mecanismo de muestreo es aleatorio, y algunos bloques terminarán con demasiadas observaciones, y otros con pocas. En contraste, usando el muestreo estratificado aseguramos que exactamente en cada uno de los bloques haya igual número de observaciones generadas. En la práctica, se usan comúnmente 40 bloques asegurando que $1/40$ de las observaciones cae dentro de cada bloque. Con 40 bloques, se elimina mucha de la varianza en el estimado de la probabilidad de pérdida debida a la variabilidad del muestreo en Q .

El primer paso para instrumentar este método es definir los bloques. Para definir los k bloques, se necesita encontrar puntos y_1, \dots, y_{k-1} tal que

$$P_{\theta}(Q < y_i) = \frac{i}{k} \quad i=1, 2, \dots, k-1$$

El subíndice θ sobre la probabilidad enfatiza que este método considera la distribución IS, donde además Q sigue siendo una función cuadrática en las variables aleatorias normales, dado que la aproximación delta-gamma todavía es aplicable. Usando este método se encuentran los y_i . Los intervalos $(y_i; y_{i+1})$, $i = 0, \dots, k-1$, ($y_0 \equiv -\infty$, $y_k \equiv \infty$), son los k bloques. Se pueden definir los bloques con cualquier otro conjunto fijo de probabilidades.

Una vez definidos los bloques, hace falta definir el mecanismo de muestreo bajo el cual una fracción igual de las $Q(i)$ generadas caiga en cada bloque. Suponga que se desea generar n muestras para cada bloque lo que significa un tamaño de muestra total igual a nk . Entonces se genera un número grande de muestras independientes Z de $N(\mu(\theta); \Sigma(\theta))$; para cada Z generada se evalúa Q y se chequea en que bloque cae; si aun no se tienen las n muestras en ese bloque, se guarda el Z generado, en otro caso se desecha. Este proceso se repite hasta que se tenga el número requerido de muestras para cada bloque.

Sea $Q(ij)$ la j -ésima muestra del bloque i y $Z(ij)$ la salida de $N(\mu(\theta); \Sigma(\theta))$ que produce esa muestra. Para esa $Z(ij)$ se tiene $\Delta S(ij) = CZ(ij)$ como antes y se calcula la pérdida del portafolio correspondiente $L(ij)$. El estimador resultante es

$$\sum_{i=1}^k \frac{1}{nk} \sum_{j=1}^n e^{-\theta Q(ij) + \psi(\theta)} I(L(ij) < x)$$

De manera más general, si se definen los bloques con probabilidades p_1, \dots, p_k y se asignan n_i muestras para cada bloque i , $i = 1, \dots, k$, el estimador será

$$\sum_{i=1}^k \frac{p_i}{n_i} \sum_{j=1}^{n_i} e^{-\theta Q(ij) + \psi(\theta)} I(L(ij) < x)$$

Esto no requiere que las n_i asignaciones sean proporcionales a las probabilidades p_i de los estratos. Las diferentes estrategias para seleccionar las n_i asignaciones son investigadas en (Glasserman [34]).

5. Modelos Alternativos

5.1 El Déficit Esperado

Hasta el año 2000, se había presentado al VaR, como la mejor opción para el manejo y administración del riesgo de mercado, sin embargo a partir de 1997, cuando Artzner, Delbaen, Eber y Heath, publican su documento "Pensado coherentemente" [5] y dos años más tarde publican su teoría sobre "las medidas coherentes de riesgo" [6] es que se desatan una serie de críticas en contra del VaR y de las cuales surge como medida alternativa el Déficit Esperado.

En la administración del riesgo financiero, VaR ha representado un paso significativo respecto a las medidas más tradicionales basadas principalmente en la sensibilidad de las variables de mercado. La fuerza de VaR reside en que:

1. VaR aplica a cualquier instrumento financiero y se expresa en la misma unidad de medida, a saber el dinero; normalmente en dinero perdido. Anteriormente, por el contrario se tenían medidas creadas para cada instrumento financiero o variable de riesgo, específico que comúnmente se expresaban en unidades diferentes, complicando la comparación de riesgo entre dos portafolios, en tanto que para la metodología VaR la comparación es en términos de la misma medida.
2. VaR incluye una estimación de eventos futuros y convierte el riesgo de un portafolio en un solo número. Por el contrario, anteriormente la medida se daba en términos esencialmente de posibles escenarios. Por ejemplo: se decía se pierde 1 peso si las tasas de interés se incrementan 1 punto, pero sin dar respuesta o explicación a la pregunta sobre ¿cuan probable es esto?. VaR por el contrario hace exactamente este trabajo, es decir, da respuesta a los cambios en los factores de riesgo.

Por su parte, las dificultades cuando se calcula el VaR de portafolios financieros vienen exactamente de las ambiciosas tareas que a continuación se señalan:

1. Incluir en el análisis el conjunto completo de variables de riesgo que afectan el portafolio.
2. Estimar correctamente la probabilidad de los eventos futuros del mercado

En casos reales, debido a la complejidad de los aspectos computacionales y de la estimación de las probabilidades del mercado para calcular VaR se tienen que asumir supuestos (a veces fuertes) sobre la dependencia funcional de los instrumentos financieros respecto a los factores de riesgo (aproximaciones de primero y segundo orden, etc.) y sobre la distribución de probabilidad (VaR histórico, VaR paramétrico, etc.).

Asimismo, en el caso de portafolios complejos expuestos a muchas variables de riesgo como en las instituciones financieras, a veces es necesario calcular el VaR sobre alguno o algunos de estos factores, de acuerdo a las necesidades de los administradores de riesgo. En esos casos se habla de VaR's parciales como VaR de la tasa de Interés

(IRVaR), el VaR de tipo de cambio (FXVaR), el VaR Accionario (EQVaR), el VaR Crediticio. (CVaR) y así sucesivamente.

Sin embargo, existe el problema de que el cálculo del portafolio completo, no puede dividirse en cálculos separados, debido a la no aditividad del VaR, que se manifiesta de dos maneras:

1. La no aditividad por instrumento financiero: dado un portafolio que se divide en dos sub-portafolios, el VaR total no es la suma de los dos VaR's, consecuentemente al agregar un nuevo instrumento financiero al portafolio es necesario volver a calcular el VaR para el portafolio con ese nuevo instrumento.
2. La no aditividad en las variables de riesgo: Para un portafolio que depende de múltiples factores de riesgo, VaR no es la suma de VaR's por factor de riesgo, esto es, VaR no es la suma del IRVaR más el FXVaR, por ejemplo.

En ambos casos, cuando los rendimientos de un portafolio son distribuidos normalmente, se puede mostrar que existe sub-aditividad: el VaR total siempre es menor o igual que la suma de VaR's parciales (por instrumento o por manejo del riesgo).

Aun cuando la no aditividad trae consigo serias dificultades computacionales, sin embargo, es uno de los aspectos más interesantes de VaR como instrumento para el análisis de riesgo, a saber su habilidad de exhibir las ventajas debidas a la diversificación de instrumentos financieros y manejo de riesgo. VaR es de hecho sensible al efecto cercano de diferentes posiciones y al efecto de correlaciones mutuas sobre el manejo del riesgo.

Además de lo anterior, existen otros factores que extienden las críticas al VaR, entre esos factores, que pueden hacer la estimación particularmente difícil, se pueden identificar los siguientes:

- 1) La complejidad de los instrumentos financieros
- 2) La dimensión del portafolio
- 3) La valoración de probabilidades del mercado
- 4) Las aproximaciones introducidas para acelerar el cálculo
- 5) El error estadístico en la estimación de VaR. El VaR, como cuantil, puede ser estimado con muy poca precisión, en particular cuando el nivel de confianza seleccionado es muy pequeño.

El conjunto de supuestos, las aproximaciones y errores estadísticos en el cálculo de VaR pueden ser tan grandes que el resultado final no tenga sentido. En otros términos, detrás de un razonable VaR de 100 pesos, un $VaR = 100 \pm 10,000$ pesos puede estar escondido.

Estas observaciones, realizadas por los críticos del VaR, pueden parecer pesimistas. Pero como en cualquier otra estimación estadística, es necesario dejarlas planteadas: uno no puede confiar en que todas las aproximaciones y supuestos incorporados para obtener la estimación estén bajo control y menos que el usuario del resultado final este consciente de los supuestos que se han adoptado.

Del VaR podría decirse, lo mismo que para la clase entera de estadísticas de la cola izquierda de la distribución de los rendimientos del portafolio, que el esfuerzo computacional necesario para modelar la distribución de probabilidad y crear una muestra de eventos es el mismo. Una vez hecho esto, cualquier variable estadística de la cola izquierda puede determinarse sin esfuerzo adicional. Entonces surge la interrogante ¿por qué el mercado financiero elige VaR y no otras estadísticas?

¿Realmente VaR es el resumen más representativo para caracterizar los riesgos asociados con eventos extremos de un portafolio? ¿Por qué VaR ha sido elegido mundialmente por administradores de riesgo financieros y reguladores de bancos como la medida apropiada por atender los riesgos financieros de mercado de un portafolio?

Empecemos con algunas consideraciones elementales sobre la propia definición de VaR. Como se ha comentado anteriormente, para calcular el VaR se tiene que seleccionar un nivel de confianza y un horizonte de tiempo sobre el que se requiere la estimación de pérdidas y ganancias. Habitualmente se considera un nivel de confianza del 99% y un horizonte de 10 días, de acuerdo con el Comité de Basilea.

Una definición que se usa a menudo para VaR es la siguiente.

VaR es la pérdida potencial máxima que un portafolio puede sufrir en el 1% de los peores casos en 10 días.

Es claro que esta definición no es correcta, toda vez que, si consideramos la cola izquierda de la distribución de pérdidas y ganancias, el mayor punto a la izquierda es infinito. Sin embargo, la versión correcta que se da a continuación se usa raramente porque parece extraña y quizá complicada:

VaR es la mínima pérdida potencial que un portafolio puede sufrir en el 1% de los peores casos en 10 días.

La cual es reemplazada frecuentemente por

VaR es la máxima pérdida potencial que un portafolio puede sufrir en el 99% de los mejores casos en 10 días.

Un análisis cuidadoso de las definiciones anteriores vierte algunas dudas sobre el hecho que VaR sea realmente lo más apropiado para describir el riesgo asociado con el 1% de los peores casos de la distribución. De hecho, es importante la mínima pérdida, pero cómo son el resto de las pérdidas, en este sentido VaR, es “lo mejor, en un escenario, del peor de los casos”, además de subestimar sistemáticamente las pérdidas potenciales asociadas con el nivel especificado de confiabilidad.

Se puede concluir que entre las estadísticas del 1% en la cola izquierda VaR es quizás la menos apropiada para caracterizar los riesgos, toda vez que no da mayor información sobre el resto de las pérdidas.

Después de tener en cuenta lo hasta aquí presentado, parece natural acudir a alguna estimación central del 1% de la cola izquierda, dado que es prioritario construir otro

resumen que describa mejor los riesgos. En particular se puede considerar la posibilidad de introducir la media del 1% de los peores casos, la cual se define como el Déficit Esperado (Expected Shortfall) y de la que puede decirse que:

ES es el valor esperado de la pérdida del portafolio en el 1% de los peores casos en 10 días.

Esta variable se ha estudiado recientemente y ha sido considerada por algunos autores. Incluso fue considerado en finanzas antes que el VaR madurara.

En su documento Artzner Delbaen Ebert y Heath [6], definen un conjunto de axiomas que debe cumplir una medida de riesgo, en un sentido general. Una medida que satisface dichos axiomas se define como una Medida Coherente de Riesgo. En general estas medidas son tales que siempre que un portafolio sea más riesgoso que otro, se tendrá un valor más alto de la medida de riesgo. Al contrario, cualquier medida que no satisface algunos de los axiomas producirá resultados paradójicos. En el documento señalado, se identifica, se caracteriza y se da una prueba de coherencia de la clase de medidas coherentes.

En el segundo axioma para una medida $\rho(\cdot)$, se dice que para cualesquiera dos variables aleatorias X y Y, $\rho(\cdot)$ es sub-aditiva si

$$\rho(X + Y) \leq \rho(X) + \rho(Y).$$

Anteriormente se señaló que VaR no es sub-aditiva, salvo en el caso Gaussiano, propiedad que es uno de los soportes principales de VaR. La sub-aditividad involucra en términos matemáticos la reducción del riesgo asociado con el concepto de diversificación. En varios libros de texto de Administración de Riesgos se dan ejemplos de VaR bajo el supuesto Gaussiano, con el cual se demuestra la sub-aditividad, de tal manera que el lector se lleva la idea de que esta propiedad es de validez general. El problema es que no es verdad. El VaR no es sub-aditivo fuera del caso Gaussiano y algunos otros especiales..

Por todo lo anterior, se propone una medida alternativa de riesgo, el Déficit Esperado, la cual mide el valor esperado de los rendimientos de un portafolio que exceden algún umbral y que además satisface los axiomas de medida de riesgo propuestos por Artzner, Delbaen, Ebert y Heath, la cual se puede expresar analíticamente como:

$$E_t(r_t | r_t < q_{t,\theta}) \tag{31}$$

Entonces, si σ_t es una función de variables conocidas en el momento t, es posible reescribir la ecuación anterior, en términos de los rendimientos estandarizados, como sigue:

$$E_t(r_t | r_t < q_{t,\theta}) = \sigma_t E_t \left(\frac{r_t}{\sigma_t} \mid \frac{r_t}{\sigma_t} < \frac{q_{t,\theta}}{\sigma_t} \right) \tag{32}$$

Si se supone que los rendimientos estandarizados siguen una distribución EVT, esta distribución proporciona una fórmula muy simple para calcular el valor en (32). En general, si X tiene una distribución de Pareto Generalizada con parámetros $\xi < 1$ y β , entonces para un umbral u menor al valor máximo de las observaciones (esta parte será presentada con mayor detalle en la siguiente sección):

$$E(X - u | X > u) = \frac{\beta + \xi u}{1 - \xi} \beta + u \xi > 0$$

Puede mostrarse (vea, por ejemplo, McNeil y Frey [56]) que para cualquier $x_p > u$,

$$E(X | X > x_p) = x_p \left(\frac{1}{1 - \xi} + \frac{\beta - \xi u}{(1 - \xi)x_p} \right) \quad (33)$$

Suponga que ajustamos un GARCH aumentado con EVT y conseguimos las estimaciones para $q_{t,\theta}$ y σ_t en (32). Entonces, aplicando la fórmula (33) para calcular la esperanza en (31), tenemos

$$\hat{E}_t(r_t | r_t < q_{t,\theta}) = \hat{q}_{t,\theta} \left(\frac{1}{1 - \hat{\xi}} - \frac{\hat{\beta} - \hat{\xi} \hat{u}}{1 - \hat{\xi}} \frac{\hat{\sigma}_t}{\hat{q}_{t,\theta}} \right)$$

Un resultado similar puede obtenerse para el modelo CAViaR. Asumiendo que $q_{t,\theta}$ es un número negativo, de (31) se tiene:

$$E_t(r_t | r_t < q_{t,\theta}) = q_{t,\theta} E_t \left(\frac{r_t}{q_{t,\theta}} \middle| \frac{r_t}{q_{t,\theta}} > 1 \right) \quad (34)$$

Dado el supuesto de que los cuantiles residuales estandarizados son i.i.d. más allá de un cierto umbral, podemos aplicar EVT de nuevo para calcular el déficit esperado:

$$\hat{E}_t(r_t | r_t < q_{t,\theta}) = \hat{q}_{t,\theta} \left(\frac{1}{1 - \hat{\xi}} + \frac{\hat{\beta} - \hat{\xi} \hat{u}}{1 - \hat{\xi}} \right)$$

Las formulas (31) y (34) sugieren una manera alternativa de calcular el déficit esperado, puesto que la esperanza condicional en ambas formulas debe ser constante (dado el supuesto de i.i.d.), uno puede calcular el déficit esperado haciendo una regresión de los rendimientos que exceden el cuantil contra el correspondiente estimador de la desviación estándar (o cuantil):

$$r_t = \beta \sigma_t + \eta_t \quad \text{para} \quad r_t < q_t \quad (35)$$

$$r_t = \delta q_t + \eta_t \quad \text{para} \quad r_t < q_t \quad (36)$$

Donde el déficit esperado esta dado, para los modelos GARCH y CAViaR respectivamente por:

$$\hat{E}_t(r_t | r_t < q_{t,\theta}) = \hat{\beta} \hat{\sigma}_t$$

$$\hat{E}_t(r_t | r_t < q_{t,\theta}) = \hat{\delta} \hat{q}_t$$

Dónde $\hat{\beta}$ y $\hat{\delta}$ son los coeficientes de la regresión de (35) y (36).

Este segundo método de estimación puede ser visto como complementario al método de EVT. De hecho, este planteamiento puede ser útil para calcular el déficit esperado en cuantiles relativamente altos (como 5%) donde la aproximación proporcionada por el EVT puede ser muy pobre. Por otro lado, el método de regresión puede dar estimaciones muy pobres en cuantiles muy bajos (como 0.01%), dado que se tendrían pocas observaciones disponibles. En el caso extremo, cuando el cuantil está más allá del límite de los datos observados, ninguna observación estaría disponible, y el método no sería aplicable.

5.2 Medidas de Riesgo y Teoría del Valor Extremo (EVT)

Recientemente, se han propuesto métodos alternativos para estimar el Valor en Riesgo, como una aplicación de la Teoría de Valor Extremo (vea, por ejemplo, Danielsson y De Vries [20] o Gouriéroux y Jasak [37]) y aplicaciones de la técnica de regresión de cuantiles en Chernozhukov y Umantsev [19] y Engle y Manganelli [29]. Otras aportaciones que pueden ser incluidas bajo esta sección son aquellas basadas en Bollerslev y Woolridge [13] llamadas cuasi-máxima verosimilitud GARCH, como sugirió Diebold, Schuermann y Stroughair [24], y que independientemente instrumentaron McNeil y Frey [56] y Engle y Manganelli [29]. Una idea similar puede encontrarse en Hull y White [43], aunque su modelo no es tan robusto, dado que la varianza no es estimada usando QML GARCH.

Para la administración de riesgo, como para cualquier problema de inferencia, es de gran importancia tener la mayor cantidad de información disponible. Parte de esta información se relaciona con el tamaño de las pérdidas en que puede incurrir un portafolio. Sin embargo, cuando se está interesado en las máximas pérdidas de un activo, pueden existir inconvenientes por la poca disponibilidad de este tipo de datos.

Adicionalmente, no es posible realizar inferencias sobre pérdidas muy grandes con metodologías como las descritas en las secciones previas, ya que en algunos casos puede que no existan pérdidas históricas de esta magnitud y por lo tanto no son consideradas por dichos modelos de inferencia.

En este caso interesa la distribución del máximo valor de la serie en lugar de la distribución de los rendimientos. Problema que es estudiado en detalle por la teoría de valores extremos, *EVT* por sus siglas en inglés, la cual presenta una estructura adecuada para la medición de riesgos, debido a que ésta se concentra en la estimación de la probabilidad de eventos que son más extremos que los observados. Análogo al teorema central del límite, que estudia el comportamiento de la medida de tendencia central de la distribución, la *EVT* se concentra en resultados asintóticos con relación al valor máximo de una variable aleatoria.

La *EVT* ha sido ampliamente usada en áreas como la hidrología y la ingeniería, y en los últimos años en finanzas. De acuerdo con Coles [18], es importante tener en cuenta algunas limitaciones de esta metodología: *i)* los resultados obtenidos se basan en derivaciones asintóticas, por lo que estos no se pueden tomar como exactos en muestras finitas, *ii)* los modelos usados en la *EVT* están contruidos bajo ciertos supuestos, los cuales no necesariamente se mantienen en la práctica y *iii)* estos modelos implican el uso parcial de la información, ya que únicamente se concentran en las observaciones extremas.

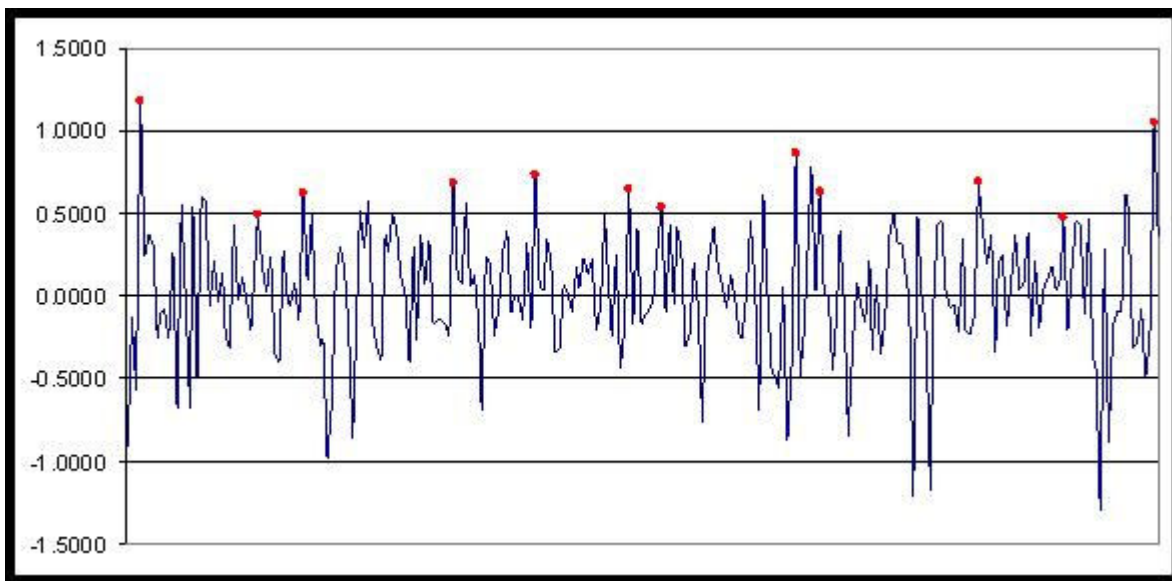
Cómo obtener valores extremos:

La implementación de la *EVT* requiere, en una etapa inicial, la obtención de una serie de valores extremos a partir de los datos observados. Como ya se ha señalado anteriormente, debido a que en finanzas generalmente se trabaja con el logaritmo de los

precios (rendimientos), los eventos con mayor valor corresponden a pérdidas muy grandes, es decir a los negativos de los rendimientos.

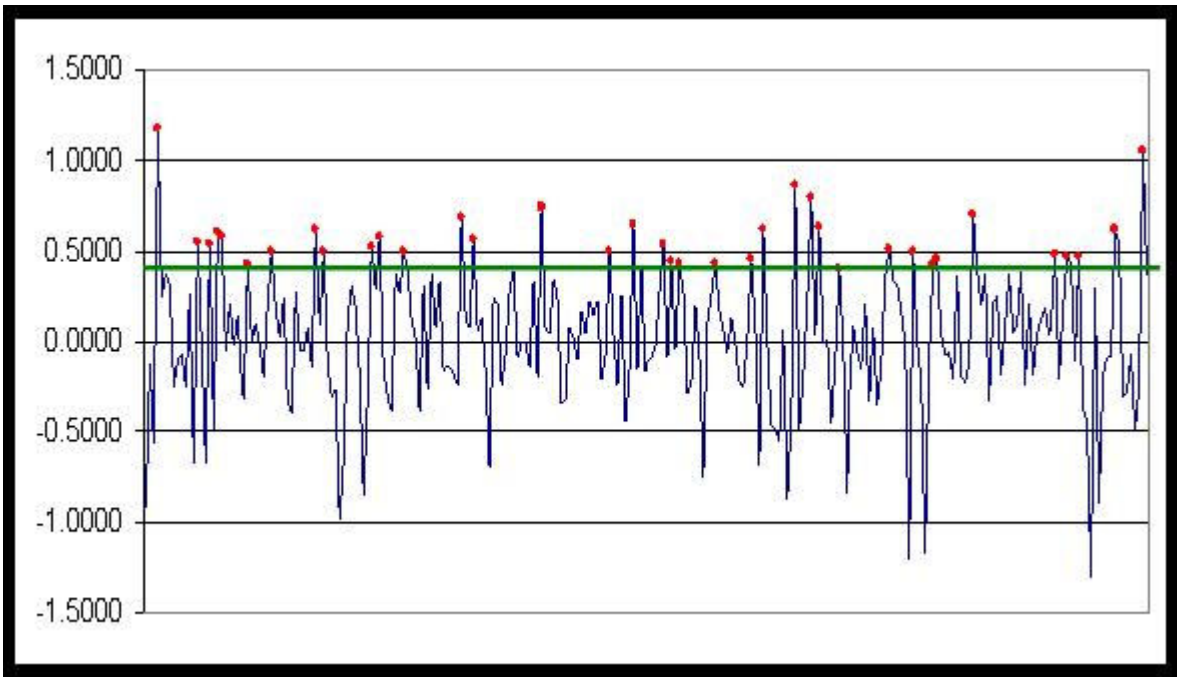
Existen dos metodologías para definir un valor extremo, el método de máximo por bloques y el método de picos sobre un umbral.

En la metodología de máximos por bloques se identifican los eventos que tengan el valor más alto durante un período de tiempo específico. Este periodo de tiempo es calculado a través de la división de la muestra en j bloques contiguos disjuntos de tamaño n . La serie de valores extremos, $\{M_{n,i}\}$, $i=1,2,\dots,j$ se obtiene a partir del valor máximo de la serie dentro de cada bloque, $M_{n,i}$. Por ejemplo, en la siguiente gráfica se presentan los máximos mensuales para los rendimientos diarios del tipo de cambio pesos/dólar, marcados por círculos rojos, entre enero de 2005 y diciembre de 2005.



Al identificar los máximos de la serie a través de esta metodología, es posible reducir la dependencia entre los eventos extremos y garantizar que se están tomando observaciones a lo largo de toda la muestra. Sin embargo, este método desconoce la información asociada a otras observaciones extremas que no correspondan al máximo de su respectivo bloque. Adicionalmente, la información obtenida se reduce a medida que se aumenta el tamaño del bloque, generando posibles pérdidas de información.

El método de picos sobre el umbral, *POT* por sus siglas en inglés, identifica los valores extremos de la serie de rendimientos como aquellos que excedan un umbral u , estos valores son conocidos como *excesos de rendimiento*. Un ejemplo de esto, aplicado al tipo de cambio pesos/dólar, en el período de enero 2005 hasta diciembre del 2005, se muestra en la siguiente gráfica. En esta gráfica los excesos de rendimiento están marcados por círculos rojos. El umbral utilizado es $u = 0.4$



Un problema que presentan ambas metodologías esta asociado a la definición de la longitud del bloque y a la definición del umbral, temas que veremos más adelante.

De acuerdo con cada una de las definiciones de máximo existen dos resultados asintóticos para modelar el comportamiento de los extremos de una variable aleatoria.

5.1.1 Distribución de los máximos por bloques

Sea $\{R_t\} t=1,2,\dots, n$ una serie de variables aleatorias *iid* las cuales siguen una distribución F_R . El máximo de esta serie, M_n es una variable aleatoria definida por:

$$M_n = \max\{R_1, R_2, \dots, R_n\} \quad (37)$$

entonces, la distribución de M_n , corresponde a la distribución del estadístico de orden n , como sigue:

$$P(M_n \leq z) = P(R_1 \leq z, R_2 \leq z, \dots, R_n \leq z) = \prod_{i=1}^n P(R_i \leq z) = (F_R(z))^n \quad (38)$$

Dado que F_R es desconocida, existen dos procedimientos para obtener la distribución de M_n . Un primer proceso consiste en encontrar F_R a partir de los datos observados, y posteriormente utilizar (38) para obtener la distribución de M_n . Un segundo procedimiento estima en forma directa F_R^n a partir, de técnicas utilizadas en *EVT*, como se explica a continuación.

Para obtener la distribución de M_n es necesario encontrar una serie de constantes $\{a_n > 0\}$ y $\{b_n\}$ que establezcan la serie de máximos tanto en localización como en escala⁷. Dicha transformación es de la forma:

$$M_n^* = \frac{M_n - b_n}{a_n} \quad (39)$$

de esta manera se pueden definir familias de distribuciones límite, que satisfagan la siguiente condición, para variables aleatorias *iid* R_1, R_2, \dots, R_n y constantes apropiadas $a_n > 0, b_n \in \mathfrak{R}$

$$F_R^n(a_n x + b_n) \xrightarrow{d} F_R(x) \quad (40)$$

Si R satisface la condición (40) se dice que la función de distribución de R, F_R , es una *distribución estable en los máximos* de R_1, R_2, \dots, R_n . Es decir, una distribución para la cual el máximo muestral, re-escalado, posee la misma distribución que la variable aleatoria inicial. Fisher y Tippet y Gnedenko demostraron que las únicas distribuciones límite, no degeneradas, a las que puede converger F_R^n están dadas por el siguiente teorema.

Teorema 1. (Fisher-Tippett, Gnedenko). Sea $\{R_n\}$ una serie de variables aleatorias *iid*. Si existen constantes estabilizadoras $a_n > 0, b_n \in \mathfrak{R}$ y alguna función no degenerada $G(x)$ tal que a medida que $n \rightarrow \infty$ entonces

$$P\left(\frac{M_n - b_n}{a_n} < z\right) = P(M_n \leq a_n z + b_n) = F_R^n(x) \xrightarrow{d} G(x)$$

donde G es una función de distribución no degenerada, con las siguientes características:

- I). Gumbel: $\Lambda(x) = \exp(-e^{-x}) \quad x \in \mathfrak{R}$
- II). Fréchet: $\Phi_\alpha(x) = \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases}$
- III). Weibull: $\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^{-\alpha}) & x \leq 0 \\ 1 & x > 0 \end{cases}$

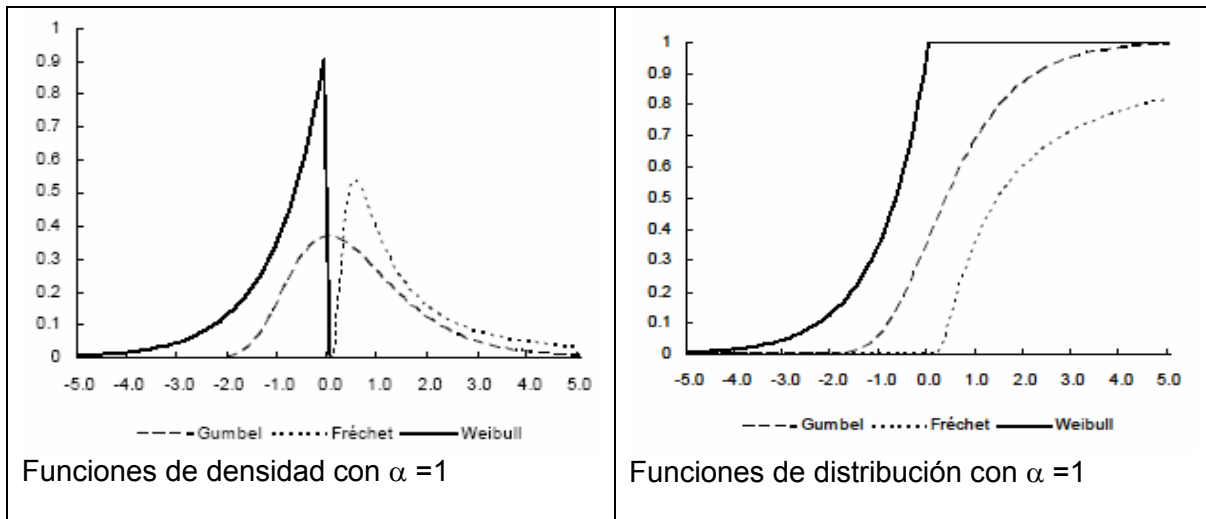
y $x = a_n z + b_n$ y $\alpha > 0$ se denomina el parámetro de forma para las familias Fréchet y Weibull.

Estas distribuciones son conocidas como *distribuciones de valor extremo* (EV por sus siglas en Ingles).

⁷ Para el caso del teorema central del límite $a_n = \sigma / \sqrt{n}$ y $b_n = \mu$.

El teorema 1 implica que el máximo de la variable aleatoria puede ser estabilizado a través de secuencias de parámetros de escala, $\{a_n > 0\}$, y de localización $\{b_n\}$ de tal manera que M_n^* converja en distribución a alguna de las familias mencionadas.

En las siguientes gráficas se presentan las funciones de densidad y distribución para cada una de las familias de distribución de valor extremo. La función de distribución Weibull posee un extremo superior finito, mientras la distribución Fréchet presenta un extremo inferior finito, además, la distribución Fréchet presenta un crecimiento más lento que la distribución Gumbel y Weibull, siendo la distribución Weibull aquella que posee la tasa de crecimiento más rápida.



Un concepto asociado a la función de distribución de los máximos es el *dominio de atracción de los máximos*, *MDA* por sus siglas en inglés, el cual establece que si la condición del Teorema 1, se tiene para alguna distribución no degenerada, G se dice que la función de distribución F_R pertenece al dominio de atracción de los máximos de G . Por lo tanto, el *MDA* relaciona la distribución, F_R , de una variable aleatoria R con la distribución límite asociada con sus máximos.

En términos prácticos, la estimación de los máximos en bloques implica varias etapas. En la primera fase se obtiene la serie de máximos, posteriormente se selecciona la distribución que se ajuste apropiadamente a esta serie y finalmente, se estiman los parámetros de la distribución seleccionada a través del método de máxima verosimilitud.

Para evitar la selección de una función de distribución a priori, es posible utilizar la representación de Jenkinson y Von Mises de las distribuciones involucradas en el teorema 1. En esta reformulación se representan las distribuciones de valor extremo en una sola función de distribución, la *distribución generalizada de valor extremo*, *GEV* por sus siglas en inglés, la cual se define como:

$$H_{\xi}(x) = \begin{cases} \exp\left(-\left(1 + \xi x\right)^{-1/\xi}\right) & \text{si } \xi \neq 0 \\ \exp\left(-e^{-x}\right) & \text{si } \xi = 0 \end{cases}$$

donde $1 + \xi x > 0$. De una manera más general, es posible obtener la distribución *GEV* con parámetros de localización, $\mu \in \Re$ y $\sigma > 0$, tal que $H_{\xi, \mu, \sigma}(x) = H_{\xi}((x - \mu) / \sigma)$

De esta manera es posible realizar la estimación en la metodología de máximos por bloques usando la distribución *GEV* y el método de máxima verosimilitud. Dependiendo del resultado de la estimación del parámetro ξ , si es mayor, menor o igual a cero, la *GEV* toma alguna de las tres formas funcionales anteriormente descritas.

La función a maximizar por el método de máxima verosimilitud es:

$$L(\xi, \mu, \sigma, x) = \prod_i \log(h(\xi, \mu, \sigma, x_i)) \quad x_i \in M$$

$$\text{donde } h(\xi, \mu, \sigma; x) = \frac{1}{\sigma} \left(1 + \xi \left(\frac{x - \mu}{\sigma}\right)\right)^{-1/\xi - 1} \exp\left(-\left(1 + \left(\frac{x - \mu}{\sigma}\right)\right)^{-1/\xi}\right)$$

Es la función de densidad si $\xi \neq 0$ y $1 + (x - \mu) / \sigma > 0$; si $\xi = 0$ entonces la función de densidad será

$$h(\xi, \mu, \sigma; x) = \frac{1}{\sigma} \exp\left(-\left(\frac{x - \mu}{\sigma}\right)\right) \exp\left(-\exp\left(-\left(\frac{x - \mu}{\sigma}\right)\right)\right)$$

En general, el *MDA* para las distribuciones de valor extremo depende de las características de las distribuciones que se estén analizando. El *MDA* para las distribuciones que presentan colas pesadas, usuales en aplicaciones financieras, se encuentran en las distribuciones Gumbel y Fréchet, aunque principalmente se concentran en esta última. Para la distribución *GEV* con $\xi > 0$, distribución Fréchet, el *MDA* esta definido como:

$$F_R \in MDA(H_{\xi}) \Leftrightarrow \bar{F}_R(x) = x^{-1/\xi} L(x) \quad (41)$$

donde $\xi > 0$, $F_R(x) = 1 - \bar{F}_R(x)$ es la función de sobrevivencia de R y $L(x)$ corresponde a una función que varía lentamente. Las funciones que pertenecen a este *MDA* son funciones en las cuales sus colas decaen a una tasa constante⁸, el parámetro $\alpha = 1/\xi$ es conocido

⁸ En estos casos, se dice que las distribuciones poseen colas que varían regularmente o *power tails*. Estas a su vez se encuentran relacionadas con un concepto conocido como leyes de potencia. Para una discusión amplia acerca de las leyes de potencia y sus implicaciones en administración de riesgo véase Farmer J. Doyne y John Geanakoplos (2004) "Power Laws in Finance and their Implications for Economic Theory". Working Paper. Yale University.

como el *índice de colas*. Formalmente, se dice que $L(x)$ es una función que *varía lentamente* si es una función que se puede valorar en el intervalo $(0, \infty)$ y satisface la condición:

$$\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1$$

para $t > 0$. Ejemplos de esta función son $L(x) = k$ donde k es una constante y $L(x) = \ln(x)$. Embrechts (1997) muestran que para variables aleatorias no negativas que pertenecen al *MDA* de la distribución Fréchet, $E(X^k)$, su k -ésimo momento, es igual a infinito para $k > 1/\xi$.

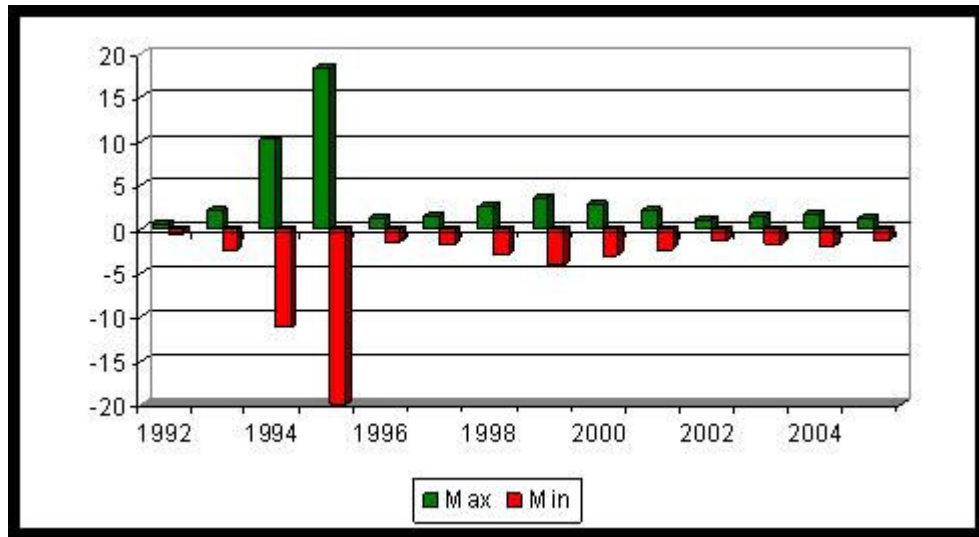
Para el caso Gumbel, $\xi = 0$, aunque su *MDA* se encuentra especialmente en distribuciones de colas livianas, también incluye algunas distribuciones con colas más pesadas que las de la normal. A diferencia del caso Fréchet donde sus colas decaen a una tasa constante, en general las colas de las distribuciones que pertenecen al *MDA* Gumbel decrecen de forma exponencial, un ejemplo de este caso es la distribución lognormal. Para $\xi < 0$, el *MDA* de la distribución Weibull resulta poco atractivo para la medición de riesgos de mercado, ya que a éste pertenecen distribuciones truncadas en la cola derecha, como es el caso de las distribuciones uniforme, triangular y beta.

La decisión acerca del tamaño del bloque n , tiene grandes implicaciones en esta metodología. Si se utilizan bloques muy pequeños se tiene una serie de máximos con gran cantidad de observaciones, aunque puede producir sesgos en la estimación debido a la mala aproximación del resultado límite presentado en el teorema 1. Por otra parte, tamaños de bloque grandes reducen este problema, pero generan una serie de máximos con pocas observaciones, lo cual produce estimaciones con varianzas grandes.

Ejemplo

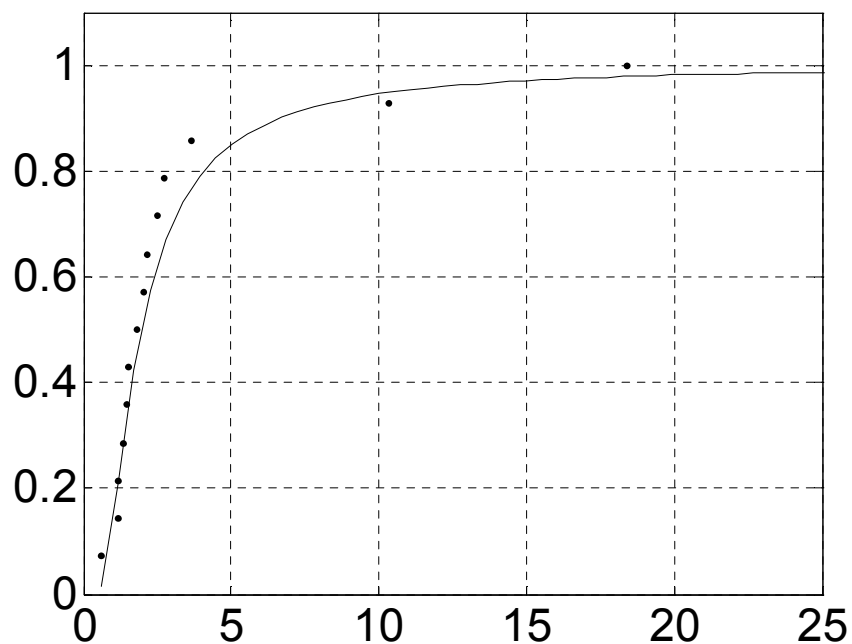
Aplicaremos el método de máximos por bloque a los rendimientos del tipo de cambio pesos/dólar del periodo de 1 de enero de 1992 al 30 de diciembre de 2005 mediante la programación en MatLab 7.0. Tal y como lo hemos mencionado el punto crítico es la elección del tamaño de bloque, a fin de ejemplificar la metodología se seleccionaran bloques anuales, que si bien no representan una muestra suficiente, en términos del teorema 1, si nos dan idea del método.

De acuerdo a lo anterior un primer paso que se realiza es cambiar el signo de los rendimientos para que los valores negativos sean ahora positivos y por ende podamos seleccionar los máximos, que son los que nos interesan, seguidamente seleccionamos los máximos en los 14 bloques anuales no traslapados, los cuales constituyen la muestra para estimar la distribución generalizada de valores extremos (GEV), y los cuales se observan en la siguiente gráfica.



De acuerdo con la teoría descrita anteriormente, la GEV es la distribución límite de extremos normalizados. Dado que en la práctica se desconoce la verdadera distribución de los rendimientos y como resultado, no se tiene idea acerca de las constantes a_n y b_n , utilizaremos la especificación de los tres parámetros. $H_{\xi, \mu, \sigma}(x) = H_{\xi}((x - \mu) / \sigma)$ con $x \in (-\infty, \mu - \sigma/\xi)$ si $\xi < 0$; $x \in (-\infty, \infty)$ si $\xi = 0$; $x \in (\mu - \sigma/\xi, \infty)$ si $\xi > 0$, donde ahora se tienen tres parámetros desconocidos μ , σ y ξ para su estimación.

De la estimación máximo verosímil se obtuvieron los estimadores $\xi = 0.644$; $\sigma = 1.004$ y $\mu = 3.638$, en la siguiente gráfica se tienen los datos maestres y la correspondiente distribución GEV ajustada



Medidas de riesgo

A partir de la teoría del valor extremo es posible introducir las medidas de riesgo que se concentran en la frecuencia y la magnitud de la realización de eventos extremos. Estas medidas son conocidas como nivel de rendimiento (*Return level*) y periodo de rendimiento (*Return period*).

Sea $M_n = \{ M_{n,1}, M_{n,2}, \dots, M_{n,j} \}$ el conjunto de los máximos obtenidos en j bloques no traslapados de tamaño n generados a partir de la realización de una variable aleatoria iid cuya función de distribución pertenece al MDA de $H_{\xi, \mu, \sigma}(x)$. El nivel de rendimiento, $r_{n,k}$ indica cual es el nivel de pérdida que se espera sea excedido, en promedio, una vez cada k períodos de tamaño n . Por ejemplo, para una serie mensual con bloques anuales, $r_{12,10}$ corresponde al nivel de pérdida anual que será excedido una vez cada diez años. Por consiguiente, la probabilidad de que una pérdida sea mayor que $r_{n,k}$ es $p = P(R > r_{n,k}) = 1/k$. El nivel de rendimiento se puede definir como el cuantil de la distribución de los máximos asociado a un nivel de confianza $1 - p$ formalmente:

$$r_{n,k} = H_{\xi, \mu, \sigma}^{-1}(1 - P)$$

$$r_{n,k} = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left(1 - [-\log(1 - p)]^{-\hat{\xi}} \right) & \text{para } \hat{\xi} \neq 0 \\ \hat{\mu} - \hat{\sigma} \log[-\log(1 - p)] & \text{para } \hat{\xi} = 0 \end{cases}$$

Siguiendo con nuestro ejemplo, $r_{360,10}$, es decir $n = 360$ representa bloques anuales, $k = 10$ y los estimadores $\hat{\xi} = 0.644$; $\hat{\sigma} = 1.004$ y $\hat{\mu} = 3.638$ tiene un valor de $r_{360,10} = 6.636$

Por otro lado, el periodo de rendimiento $k_{n,v}$, indica cada cuantos períodos de tamaño n una pérdida excederá, en promedio, un nivel de pérdidas v , en el conjunto $M_n = \{ M_{n,1}, M_{n,2}, \dots, M_{n,j} \}$ definido de igual manera que en los párrafos anteriores, entonces la probabilidad del evento $\{ M_n > v \}$ esta dada por:

$$P(M_n > v) = H_{\xi, \mu, \sigma}(v) = p$$

donde $H_{\xi, \mu, \sigma}(v)$ es la función de sobrevivencia de M_n . Sea $L(u)$ la variable aleatoria que cuenta el número de períodos hasta que un máximo presente una pérdida superior a v , es decir⁹:

$$\tilde{L}(v) = \min\{i > 1 : M_{n,i} > v\}$$

$L(v)$ sigue una distribución geométrica con probabilidad p . El periodo de rendimiento se define como:

$$k_{n,v} = E[\tilde{L}(v)] = \frac{1}{p} = (H_{\xi, \mu, \sigma}(v))^{-1}$$

⁹ Una demostración completa de este resultado se encuentra en [26]. Pág. 305.

Por ejemplo, para una serie mensual con bloques anuales, $k_{12,0.05}$ indica cada cuantos años, en promedio, se presentará una pérdida mensual superior al 5%.

5.1.2 Distribución de máximos sobre un umbral

Un segundo tipo de metodología desarrollado en teoría del valor extremo es la asociada a la distribución de los máximos sobre un umbral. El procedimiento para obtener los valores extremos según el método *POT*, se ejemplificó anteriormente. Como lo menciona Coles (2001), esta metodología tiene una gran aceptación debido al mejor aprovechamiento de los datos frente a la metodología de máximos en bloques.

Sea R una variable aleatoria la cual tiene una función de distribución, F_R la función de distribución de R condicionada a que esta variable sea mayor que un umbral dado u , es conocida como la distribución de *excesos de* R , $F_{R,u}$ y esta dada por:

$$F_{R,u}(y) = P(R - u \leq y | R > u) \quad (42)$$

donde $0 < y < R_U - u$ donde R_U , es el extremo superior de la variable aleatoria.

Si $F_R \in \text{MDA}(H_{\xi,\mu,\sigma})$ es posible mostrar de acuerdo al siguiente teorema:

Teorema. (Pickands-Balkema-de Haan). Para una clase grande de funciones de distribución F , la distribución condicional de los excesos $F_{R,u}$, para u grande se puede aproximar por

$$F_{R,u} \approx G_{\xi,\sigma} = \begin{cases} 1 - (1 + \xi y / \sigma)^{1/\xi} & \text{si } \xi \neq 0 \\ 1 - \exp(y / \sigma) & \text{si } \xi = 0 \end{cases}$$

Para $y \in [0, (R_u - u)]$ si $\xi \geq 0$ e $y \in [0, -\sigma / \xi]$ si $\xi < 0$ $G_{\xi,\sigma}$. Los parámetros ξ y σ son conocidos como los parámetros de forma y escala, respectivamente y $G_{\xi,\sigma}$, es conocida como la *Distribución Generalizada de Pareto*, *GPD* por sus siglas en inglés. Balkema y De Hann (1974) y Pickands [62], muestran que la *GPD* es la única función de distribución no degenerada la cual aproxima el comportamiento de la función de excesos.

sí x esta definida como $x = u + y$, la *GPD* puede también expresarse como una función de x , es decir, $G_{\xi,\sigma} = 1 - (1 + \xi(x-u)/\sigma)^{-1/\xi}$

En otras palabras, para aquellas distribuciones en las que sus máximos muestrales convergen en distribución a la distribución *GEV*, su distribución de excesos converge a una *GPD*. Además, ya que el parámetro ξ de la *GPD* corresponde al mismo parámetro de la distribución *GEV*, el valor de ξ sigue describiendo el tipo de función de distribución, tal como se mostró anteriormente.

Dependiendo del valor del parámetro ξ de la *GPD* se obtienen tres tipos de funciones de distribución. En el caso que $\xi > 0$, la *GPD* es una distribución de Pareto con parámetros $\alpha = 1/\xi$, $k = \sigma/\xi$, para valores no negativos de la variable aleatoria, $y \geq 0$. Para $\xi = 0$, la *GPD*

corresponde a una distribución exponencial con parámetro $1/\sigma$, $y \geq 0$, Cuando $\xi < 0$, las *GPD* toma la forma de una distribución del tipo Pareto II, la cual esta definida para el rango $0 \leq y \leq -\sigma/\xi$

Una función relacionada con los modelos de excesos sobre un umbral es la media de los excesos, $e(u)$, la cual se define como:

$$e(u) = E(X - u | X > u)$$

para el caso de la *GPD*, con parámetros $\sigma = \sigma_0 > 0$ y $\xi < 1$, la media de los excesos esta dada por:

$$e(u) = \frac{\sigma(u)}{1-\xi} = \frac{\sigma_0 + \xi u}{1-\xi} \quad (43)$$

en este caso en particular, la media de los excesos es lineal con respecto a u con pendiente $\xi/(1-\xi)$.

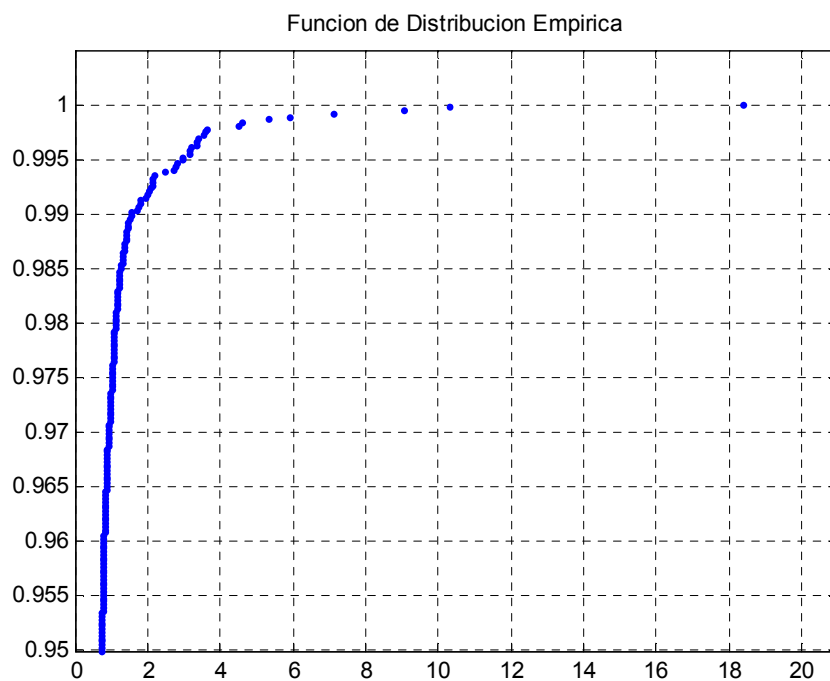
El problema de la selección del umbral.

De manera análoga al problema de la selección del tamaño del bloque, en la metodología *POT* la selección del umbral u es complicada. Un umbral muy bajo garantiza una serie de máximos con un número grande de observaciones, así mismo, un umbral demasiado alto genera pocos excesos.

Antes de pasar a ejemplificar la metodología *POT*, es conveniente hacer un poco de análisis exploratorio de los datos, el cual puede proporcionar cierta ayuda sobre el problema de la selección del umbral, lo que se hará primero es ordenar los datos a fin de graficar una muestra de la función de distribución sobre un conjunto de n observaciones como sigue:

$$F_n(X_i) = i/n \quad i = 1, 2, \dots, n$$

La función de distribución empírica corresponde a la cola derecha de los datos, porción de la distribución que se quiere estimar, y que se observa en la siguiente gráfica.



Sea $\{r_t\}$, una serie de tiempo de tamaño n , la cual representa los rendimientos de un activo. La serie de excesos sobre un umbral dado u , esta definido por:

$$Z = \{ r_t, : r_t > u \}$$

y el número de excesos sobre u , N_u esta dado por:

$$N_u = \text{card}\{Z\}$$

La selección del umbral en la metodología *POT* puede llevarse a cabo identificando u o seleccionando N_u . Existen diferentes metodologías para identificar el umbral las cuales se basan en métodos gráficos y técnicas paramétricas, los cuales mencionamos a continuación.

Gráfico de la media de los excesos

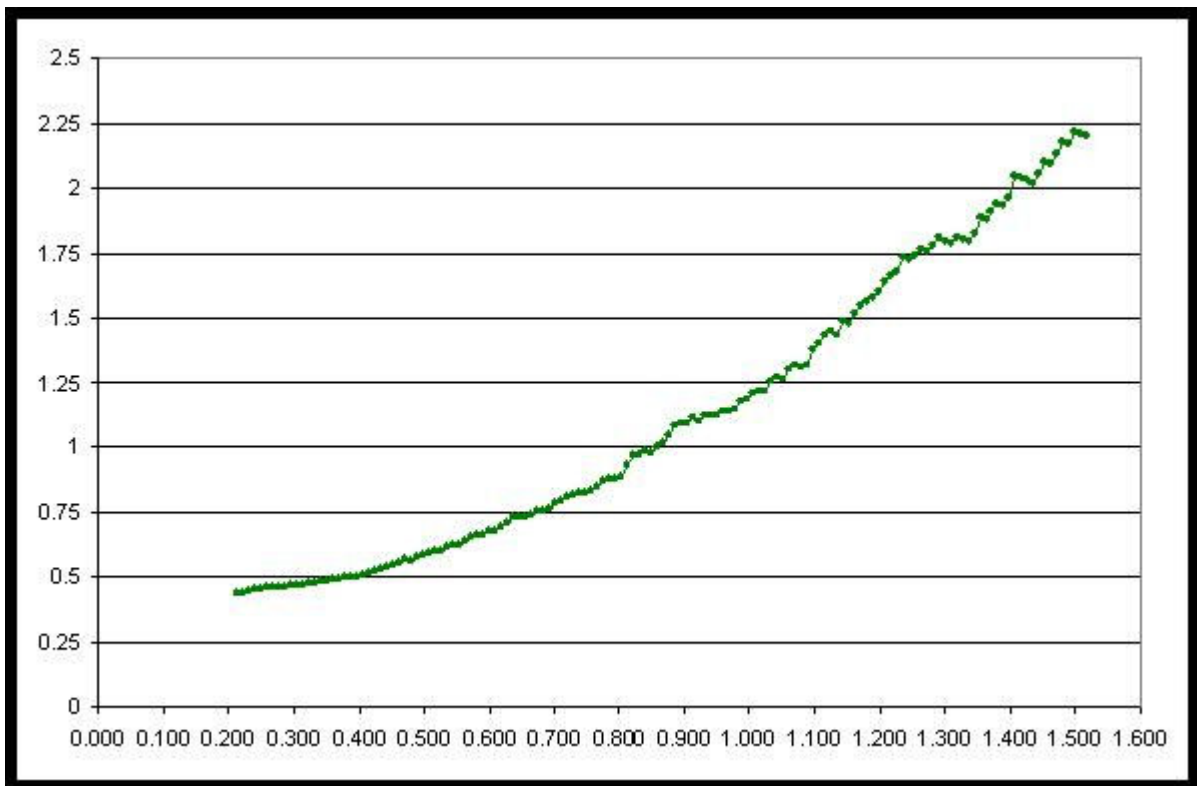
Esta metodología se basa en el análisis gráfico de la relación existente entre u y $e(u)$, de acuerdo con el resultado presentado en la ecuación (43). Para ello se utiliza el promedio muestral de la serie de excesos sobre un umbral w , como estimador del valor esperado de los excesos, definido como:

$$e(w) = \frac{1}{N_w} \sum_{i=1}^{N_w} (r_i - w)$$

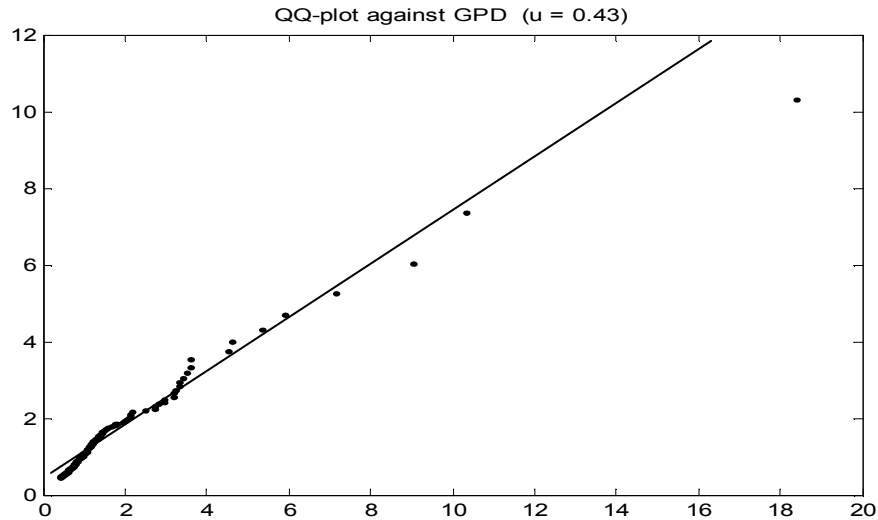
Donde r_1, r_2, \dots, r_{N_u} son las N_u observaciones mayores a $w < r_{MAX}$, donde r_{MAX} corresponde a la observación más grande de la muestra. A partir de este estimador de la media de los

excesos, se analiza la tendencia de la gráfica de las parejas ordenadas $(u, e(u))$. Si $G_{\xi, \beta}$, es una aproximación válida de $F_{R, u}$ para un umbral dado u^* , el gráfico de la media de los excesos debe ser aproximadamente lineal alrededor de u^* de acuerdo con la ecuación (43), lo cual permite establecer intervalos a partir de los cuales se puede seleccionar el umbral. Además, si la distribución de la serie bajo análisis tiene colas pesadas y su primer momento es finito, se tiene que $0 < \xi < 1$, y la pendiente $\xi / 1 - \xi$, es positiva.

En la siguiente gráfica se presenta la estimación de media de los excesos para los rendimientos pesos/dólar. En el eje (X) se encuentran diferentes umbrales, mientras que en el eje (Y) se encuentra el valor estimado de la media de los excesos ($e(u)$). Se puede observar que en los intervalos $(0.4, 0.46)$ y $(0.47, 0.55)$ de los umbrales, la función muestra una tendencia lineal alrededor de u^* , por lo que en estos intervalos podemos seleccionar el valor del umbral, de acuerdo a lo señalado en el párrafo anterior. El número de observaciones asociado a umbral extremo de los intervalos son $(505, 403)$ y $(389, 307)$ respectivamente.



Dada la relación en la gráfica anterior podemos seleccionar un valor en cualesquiera de los intervalos que tomaremos como el valor del umbral, para nuestro ejemplo seleccionamos el valor de 0.43, con este valor podemos realizar la grafica de cuantiles de la muestra contra los cuantiles de la distribución GPD, que podemos observar en la siguiente gráfica.



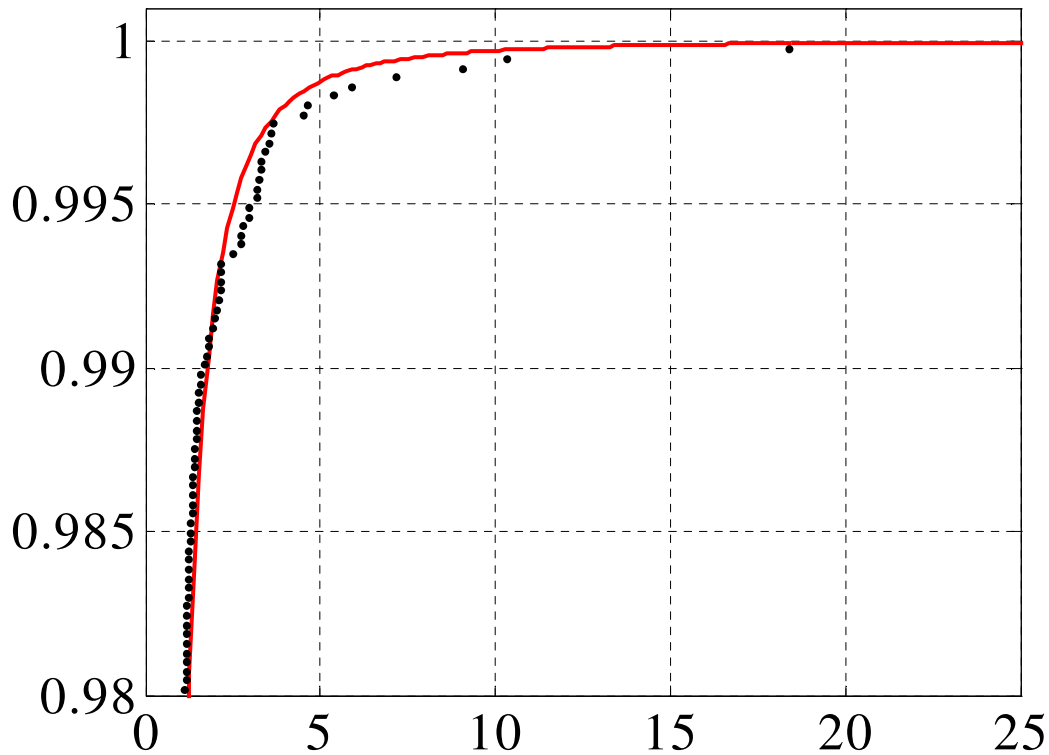
Dados los resultados teóricos presentados previamente, se conoce que la distribución de las observaciones sobre el umbral en la cola debería ser una distribución generalizada de Pareto (GPD), lo cual se confirma con la gráfica de cuantiles presentada anteriormente, donde se observa que el último punto, podría ser analizado de manera separada como un outlier.

Con esta elección del umbral y se tiene la muestra con la que se realiza la estimación de los parámetros de la GPD, que se realizará mediante el método de máxima verosimilitud, programado en MatLab 7.0, aunque se pueden utilizar otras alternativas de estimación.

Para nuestra muestra la función log-versomilitud que debemos maximizar tiene la siguiente forma:

$$L(\xi, \sigma | y) = \begin{cases} -n \ln \sigma - \left(\frac{1}{\xi} + 1 \right) \sum_{i=1}^n \ln \left(1 + \frac{\xi}{\sigma} y_i \right) & si \quad \xi \neq 0 \\ -n \ln \sigma - \frac{1}{\sigma} \sum_{i=1}^n y_i & si \quad \xi = 0 \end{cases}$$

Al realizar la optimización de esta función mediante el paquete MatLab 7.0, se obtuvieron los siguientes estimadores de $\xi = 0.478$ y $\sigma = 0.269$, Con estos valores se obtiene la siguiente gráfica del ajuste de la distribución estimada a la distribución empírica:



Medidas de riesgo en modelos sobre un umbral:

La función de *excesos de pérdida* de R_t , definida en la ecuación (42), puede ser expresada como:

$$F_{R,u}(y) = \frac{F_R(y+u) - F_R(u)}{1 - F_R(u)} \quad (44)$$

Sea $x = y + u$, si la condición del teorema 2 se cumple para un umbral u lo suficientemente grande, entonces:

$$F_{R,u}(x-u) \approx G_{\xi,\beta}(x-u) \quad (45)$$

Reemplazando la condición (45) en la (44):

$$F_R(x) = (1 - F_R(u)) G_{\xi,\beta}(x-u) + F_R(u) \quad (46)$$

Si F_R es conocida, el cálculo de las medidas de riesgo asociadas solo requieren de la estimación de los parámetros de la distribución $G_{\xi,\beta}$.

Sin embargo, en la práctica dicha función (F_R) es desconocida, por lo que es necesario en la ecuación (40), un estimador de $F_R(u)$, que puede ser la función de distribución empírica,

igual a $(n-k)/n$, donde n es el número de observaciones y k es el número de observaciones que son mayores al umbral.

Reemplazando $G_{\xi,\beta}(x-u)$ con sus parámetros estimados, así como el estimador de F_R en la ecuación (46) se obtiene que:

$$\hat{F}_R(x) = 1 - \frac{k}{n} \left(1 + \frac{\hat{\xi}}{\hat{\sigma}} \frac{x-u}{\hat{\sigma}} \right)^{-1/\hat{\xi}} \quad (47)$$

A partir de esta ecuación es posible construir medidas de riesgo como el VaR y el Déficit Esperado (ES). Para un nivel de probabilidad α , un estimador del VaR , como cuantil, esta dado por:

$$VaR_R = u + \frac{\hat{\sigma}}{\hat{\xi}} \left(\left(\frac{1-\alpha}{k/n} \right)^{-\hat{\xi}} - 1 \right)$$

Siguiendo con nuestro ejemplo y sustituyendo los valores estimados $\xi = 0.478$ y $\sigma = 0.269$ con $\alpha = 0.01$, y el valor del umbral igual a 0.43, con $k = 446$ y $n = 3519$, se tiene que el valor de $VaR = 1.3705$

Si definimos ES como:

$$ES\alpha = E(X | X > VaR\alpha)$$

Al sumar y restar $VaR\alpha$ de la expresión anterior se obtiene:

$$ES\alpha = VaR\alpha + E(X - VaR\alpha | X > VaR\alpha) \quad (48)$$

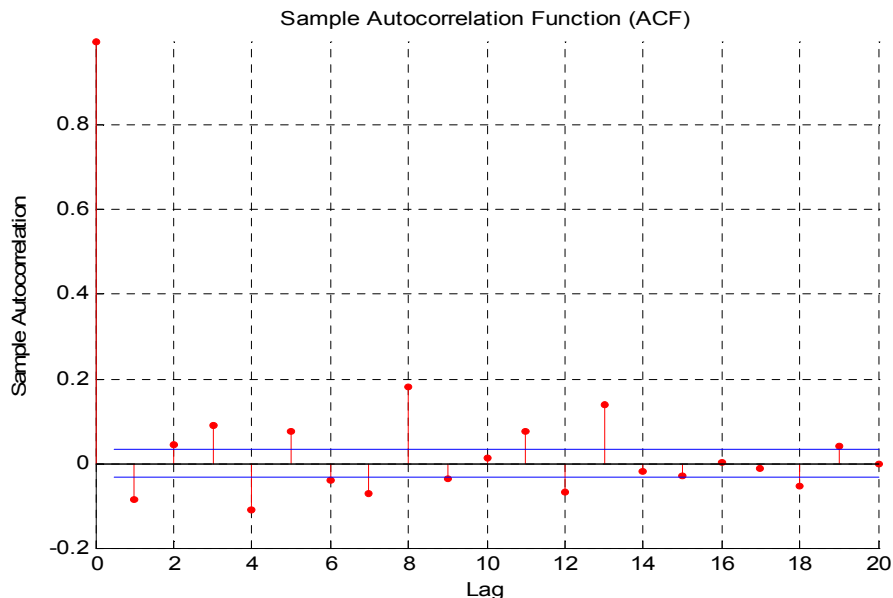
Combinando la ecuación (48) con los resultados obtenidos en (43), un estimador del ES , como cuantil, esta dado por:

$$\hat{ES}\alpha = \frac{\hat{VaR}\alpha}{1 - \hat{\xi}} + \frac{\hat{\sigma} + \hat{\xi}u}{1 - \hat{\xi}}$$

Siguiendo con nuestro ejemplo y sustituyendo los respectivos valores se tiene que $ES = 3.53455$

Para resumir, EVT parece un acercamiento muy general para la estimación de Var y ES , cuya fuerza principal es el empleo de la distribución GEV para estimar la cola de las pérdidas. Este procedimiento no parece tan restrictivo, en contraste con la mayoría de las distribuciones normalmente usadas. De otra parte hay varios problemas que necesitan ser considerados.

Primero, el supuesto de observaciones i.i.d. las cuales en la práctica no siempre se tienen, dadas las características de los datos financieros, tal y como se observa en la siguiente grafica donde se muestra la correlación de nuestros datos de ejemplo.



Aunque se han propuesto generalizaciones para el caso de las observaciones dependientes (vea, por ejemplo, Leadbetter, Lindgren, y Rootzen [50] o Embrechts, Kluppelber y Mikosch [26]), ellos estiman la distribución incondicional marginal o imponen condiciones que gobiernan el comportamiento de la volatilidad agrupada en los típicos datos financieros.

Segundo, EVT sólo trabaja para niveles de probabilidad muy bajos, estudios de Monte Carlo han revelado que el desempeño de los estimadores EVT se deteriora cuando nos movemos fuera de la cola.

Estrechamente relacionado a este problema es la selección del punto de terminación que determina el número de estadísticas de orden a ser usado en el procedimiento de estimación máximo verosímil. La elección del umbral u , presenta los mismos problemas encontrados en la selección del número k de estadísticas de orden superior que entran en el estimador de Hill. Si el umbral es demasiado alto, hay pocas excepciones y el resultado es un estimador alto de la varianza. Por otro lado, un umbral demasiado bajo produce un estimador sesgado, porque la aproximación asintótica puede ser muy pobre. Desafortunadamente hoy en día no hay ningún método estadístico, para seleccionar el umbral y uno confía en los elementos gráficos. Recientemente Danielsson y de Vries [21] y Danielsson, de Haan, Peng y de Vries [22] han propuesto un método de bootstrap de dos pasos para seleccionar el fragmento de la muestra sobre la cual el estimador de Hill puede basarse, contribución importante que puede complementar el usual análisis gráfico y estimular la investigación extensa de este problema ya que sería interesante saber si su resultado también puede extenderse al planteamiento de GPD.

Otro problema es que la elección de k del estimador de Hill, se basa en la minimización del estimador sesgado del error cuadrático medio. Como el cuantil es una función no-lineal del estimador de la cola, es importante cuantificar el tamaño de este sesgo. Muy poca atención ha recibido este problema, subsecuentemente el Valor en Riesgo es el último objeto de interés, y una estimación mala de él podría tener consecuencias sobre la rentabilidad de una institución financiera, en la literatura se piensa que esto es un tema que merece una exploración más extensa.

De acuerdo con Embrechts, la teoría de valores extremos tiene más ventajas que desventajas para su uso y por ello puede convertirse a la larga en la metodología con la que la administración de riesgos haga frente a los riesgos de mercado: Dichos pros y contras pueden resumirse como sigue:

Lo que EVT tiene que ofrecer a la administración de riesgo global.

- Complementa un modelo VaR existente.
- Es Consistente: entre los factores de riesgo y de mercado, con un análisis diferente de escenarios, es decir, cuando se seleccionan diferentes tamaños de bloque o umbral.
- Puede producir intervalos de confianza más allá del 99%.
- Proporciona un idioma de Riesgo útil, como el de “un día en un millar” o “uno en cada 7 años”.
- Se tiene la parte teórica, así como con las herramientas computacionales disponibles.
- EVT no es una manera ciega de generar VaRs más grandes que los que pueden producirse con pruebas de extremos (stress tests).

Las debilidades:

- EVT no es la solución completa para la administración de riesgo, hay varios problemas teóricos, como por ejemplo la dependencia múltiple de los factores de riesgo.
- EVT proporciona nuevas herramientas en la administración del riesgo, pero no es una nueva visión de la misma.
- EVT asume, implícitamente, que el comportamiento de la distribución de pérdidas es igual al comportamiento en la cola de la distribución de los extremos.

El cálculo

- El cálculo es rápido, en una PC pentium III, con 256 MB de memoria ram, y utilizando MatLab 7.0, los ejemplos descritos consumieron menos de 10 segundos de procesamiento.
- La convergencia (Máxima Verosimilitud) de los parámetros estimados no esta garantizada. Sin embargo existen métodos alternativos de estimación como, el método de momentos

6. Optimización del Portafolio

En general, la mayoría de los inversionistas se enfoca en los rendimientos, descuidando o prestando poca atención al riesgo que se toma para generar esos rendimientos. La información de los rendimientos se encuentra disponible en muchos medios como los periódicos, sitios de Internet, y en los informes de las instituciones financieras, en tanto que la información de riesgo es difícil de obtener. Los inversionistas ven clasificaciones jerárquicas en revistas de fondos de inversión, en publicidad de las instituciones financieras y otros medios con el supuesto implícito de que estos fondos representan lo mejor. Sin embargo, para juzgar el desarrollo de la inversión, se debe entender cuánto riesgo es tomado para generar esos rendimientos.

Además de los adjetivos o calificaciones que ofrecen las instituciones sobre el riesgo, los inversionistas deberían de recibir información cuantitativa del riesgo. Los fondos de inversión proporcionan información acerca de los rendimientos y la composición de los fondos, sin embargo, la información sobre el riesgo es vaga y general (por ejemplo, se dice cosas como: los fondo presenta un riesgo “agresivo”, “balanceado” o “moderado”, o en el mejor de los casos con números o letras de una clasificación que no es clara o de donde no se puede desprender cual es la pérdida que se puede obtener). El riesgo debe cuantificarse tan rigurosamente como los rendimientos. Por ejemplo, los inversionistas se sentirían informados si en sus estados de cuenta de sus fondos de inversión les dijeran que en el último trimestre los rendimientos fueron “equilibrados” o “moderados” en lugar de los rendimientos. La información de riesgo es tan importante como el valor del recurso y los rendimientos, y debe ponerse al día y estar accesible en una base diaria.

En el presente trabajo se han presentado diferentes metodologías para obtener el valor en riesgo (VaR), metodologías que con sus ventajas y desventajas, han resultado útiles en la administración de riesgos y permiten tener información sobre las posibles pérdidas. En esta parte del trabajo y relacionado con el VaR, veremos una metodología sobre la optimización del portafolio en el sentido de maximizar los rendimientos, minimizando el riesgo, la cual puede modelar el problema de dos maneras: la primera maximizando los rendimientos con un nivel de riesgo predeterminado y segundo minimizando el riesgo con un nivel de rendimiento fijado de antemano.

En ambos casos el riesgo se medirá en términos del déficit esperado (ES) que se definió como la esperanza condicional de las perdidas que superan un umbral (u) con una probabilidad $(1-c)$. Definición que, en primera instancia, nos dice que el VaR siempre será menor que el ES. Es decir, que para portafolios con ES pequeño se tendrá también un VaR pequeño.

Así mismo, de la definición de ambas medidas se desprende que al minimizar el ES de un portafolio se estará minimizando el VaR, de ese mismo portafolio, por ello, la inclusión de esta metodología en el presente trabajo.

Cabe señalar que en la literatura se encontraron diversos estudios que permiten aplicar esta metodología para evaluar:

- Derivados (opciones, futuros) lineales y no lineales
- Riesgo de mercado, crediticio y operacional
- Otros aspectos en las que se tenga o se este expuesto a riesgo financiero.

Descripción de la metodología

La optimización de un portafolio, ha recorrido un largo camino, desde Markowitz con sus trabajos pioneros en 1952 [53]. Camino que ha sido estimulado por dos aspectos básicos: a) el modelado de funciones de utilidad, riesgo y sus restricciones y b) la habilidad para manejar un número grande de instrumentos financieros y de posibles escenarios. Ambos factores, han influenciado en el desarrollo de herramientas, como el método de la desviación absoluta sobre la media, el método de optimización regret y el método minimax. Recientemente, otras metodologías eficientes, han recurrido a la Programación Lineal, en la solución del problema que nos ocupa en esta parte del trabajo.

Sea $f(\mathbf{x}; \mathbf{y})$ la función de pérdida asociada con el vector de decisión \mathbf{x} , elegido de un cierto subconjunto X de \mathfrak{R}^n , y sea \mathbf{y} un vector aleatorio en \mathfrak{R}^m . El vector \mathbf{x} puede interpretarse como un portafolio, donde X es el conjunto de portafolios disponibles, en tanto que el vector \mathbf{y} podría representar la incertidumbre, como por ejemplo, los factores de mercado, que afectan los rendimientos.

Para cada \mathbf{x} , la función de pérdida $f(\mathbf{x};\mathbf{y})$ es una variable aleatoria que tiene una distribución en \mathfrak{R} inducida por \mathbf{y} . La distribución de probabilidad subyacente de \mathbf{y} en \mathfrak{R}^m se supondrá, por conveniencia, que tiene densidad, la cual se denotará por $p(\mathbf{y})$, aunque para la aplicación del modelo no es necesaria.

La probabilidad de que $f(\mathbf{x}; \mathbf{y})$ no exceda un umbral α esta dada por:

$$\Psi(x, \alpha) = \int_{f(x,y) \leq \alpha} p(y) dy \quad (49)$$

Como una función de α para \mathbf{x} fijo, $\Psi(\mathbf{x};\alpha)$ es la función de distribución para la pérdida asociada con \mathbf{x} . Esta definición determina el comportamiento de esta variable aleatoria y es fundamental en la definición de VaR y ES.

En general, $\Psi(\mathbf{x};\alpha)$ es no decreciente con respecto a α y se supondrá que es continua con respecto a α . Este supuesto, al igual que el supuesto sobre la densidad de \mathbf{y} , se realiza por simplicidad, de acuerdo con Stanislav Uryasev [75], que señala que en situaciones prácticas, la continuidad requerida se sigue de las propiedades de la función de pérdida $f(\mathbf{x};\mathbf{y})$ y de la densidad $p(\mathbf{y})$.

Los valores β VaR y β ES para la variable aleatoria de pérdida asociada con \mathbf{x} y cualquier valor de probabilidad β en $(0,1)$, se denotarán por $\alpha_\beta(x)$ y $\phi_\beta(x)$, como sigue:

$$\alpha_\beta(x) = \min \{ \alpha \in \mathfrak{R} : \Psi(x, \alpha) \geq \beta \} \quad y$$

$$\phi_{\beta}(x) = (1 - \beta)^{-1} \int_{f(x,y) \geq \alpha_{\beta}(x)} f(x,y) p(y) dy$$

En la primera fórmula, $\alpha_{\beta}(x)$ representa el punto final izquierdo, del conjunto no vacío, que consiste de los valores α en los que $\Psi(\mathbf{x}, \alpha) = \beta$ (Esto se sigue de que $\Psi(\mathbf{x}, \alpha)$ es continua y no decreciente con respecto a α . El conjunto podría contener más de un punto si Ψ tiene alguna parte lineal). En la segunda fórmula, la probabilidad que $f(\mathbf{x}; \mathbf{y}) \geq \alpha_{\beta}(x)$ es igual a $1 - \beta$. De su parte $\phi_{\beta}(x)$ se calcula como la esperanza condicional de la pérdida asociada con \mathbf{x} , relativa a $\alpha_{\beta}(x)$.

La clave del método para maximizar los rendimientos, sujeto a un nivel del valor esperado, es la caracterización de $\phi_{\beta}(x)$ y de $\alpha_{\beta}(x)$ en términos de la función F_{β} en $X \times \Re$ que se define como:

$$F_{\beta}(x, \alpha) = \alpha + (1 - \beta)^{-1} \int_{y \in \Re^m} [f(x, y) - \alpha]^+ p(y) dy \quad (50)$$

donde $[t]^+ = t$ cuando $t > 0$ y $[t]^+ = 0$ cuando $t \leq 0$. De la definición de F_{β} , y con los supuestos de que $\Psi(\mathbf{x}, \alpha)$ es continua y no decreciente con respecto a α . Se puede demostrar que F_{β} es convexa, propiedad clave en la optimización, ya que elimina la posibilidad de tener mínimos locales vea Rockafellar [69], Shor [72].

Teorema 1. Como una función de α , $F_{\beta}(\mathbf{x}; \alpha)$ es convexa y continuamente diferenciable. El β ES de la pérdida asociada con cualquier $\mathbf{x} \in X$, puede determinarse como

$$\phi_{\beta}(x) = \min_{\alpha \in \Re} F_{\beta}(x, \alpha)$$

En esta fórmula el conjunto que contiene los valores α para el cual el mínimo se obtiene es:

$$A_{\beta}(x) = \arg \min_{\alpha \in \Re} F_{\beta}(x, \alpha)$$

el cual es no vacío, cerrado y acotado (reducido, quizás a un solo punto).

El β VaR de la pérdida esta dado por

$$\alpha_{\beta}(x) = \text{punto final izquierdo de } A_{\beta}(x)$$

en particular, siempre se tiene que

$$\alpha_\beta(x) \in \arg \min_{\alpha \in \mathfrak{R}} F_\beta(x, \alpha) \quad \text{y} \quad \phi_\beta(x) = F_\beta(x, \alpha_\beta(x))$$

Los interesados en la demostración, la pueden encontrar en R. Tyrrell Rockafellar y Stanislav Uryasev [70].

El Teorema 1, facilita la solución del problema planteado ya que es evidente que las funciones convexas continuamente diferenciables son más fáciles de minimizar. También revelan el hecho de que β ES puede calcularse sin tener que calcular primero el β VaR del cual depende su definición. El β VaR puede obtenerse, como un subproducto, determinando el conjunto $A_\beta(\mathbf{x})$ y extrayendo su punto final izquierdo, si contiene más de un punto.

Además, $F_\beta(\mathbf{x}; \alpha)$ puede aproximarse de varias maneras. Por ejemplo, muestreando la distribución de probabilidad de \mathbf{y} según su densidad $p(\mathbf{y})$. Este muestreo genera una colección de vectores $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k$ entonces la aproximación correspondiente a $F_\beta(\mathbf{x}; \alpha)$ se puede calcular como:

$$\tilde{F}_\beta(x, \alpha) = \alpha + \frac{1}{k(1-\beta)} \sum_{k=1}^k [f(x, y_k) - \alpha]^+ \quad (51)$$

Esta aproximación de $F_\beta(\mathbf{x}; \alpha)$ es convexa y lineal por tramos con respecto a α . Aunque no es diferenciable con respecto a α , pero puede minimizarse rápidamente, con técnicas de búsqueda lineal o representándola como un problema de programación lineal, como veremos más adelante.

Otras ventajas importantes de calcular el VaR y el ES a través de las fórmulas del Teorema 1, se siguen del siguiente teorema.

Teorema 2. Minimizar el β ES de la pérdida asociada con \mathbf{x} para toda $\mathbf{x} \in X$ es equivalente a minimizar $F_\beta(\mathbf{x}; \alpha)$ para toda $(\mathbf{x}; \alpha) \in X \times \mathfrak{R}$, esto es,

$$\min_{x \in X} \phi_\beta(x) = \min_{(x, \alpha) \in X \times \mathfrak{R}} F_\beta(x, \alpha) \quad (52)$$

donde además, $(\mathbf{x}^*; \alpha^*)$ es el mínimo del lado derecho de la ecuación (51) si y sólo si \mathbf{x}^* minimiza el lado izquierdo y $\alpha^* \in A_\beta(\mathbf{x}^*)$. En particular, si el conjunto $A_\beta(\mathbf{x}^*)$, se reduce a un solo punto, minimizar $F_\beta(\mathbf{x}; \alpha)$ sobre $(\mathbf{x}; \alpha) \in X \times \mathfrak{R}$, produce un par $(\mathbf{x}^*; \alpha^*)$, tal que \mathbf{x}^* minimiza el β ES y α^* el correspondiente β VaR.

Además, $F_\beta(\mathbf{x}; \alpha)$ es convexa con respecto a $(\mathbf{x}; \alpha)$, y $\phi_\beta(x)$ es convexa con respecto a \mathbf{x} cuando $f(\mathbf{x}; \mathbf{y})$, es convexa con respecto a \mathbf{x} , y si además, las restricciones son tales que X es un conjunto convexo, la minimización conjunta es un caso de programación convexa.

Los interesados en la demostración del teorema 2, la pueden encontrar en R. Tyrrell Rockafellar y Stanislav Uryasev [70].

De acuerdo al Teorema 2, para determinar \mathbf{x}^* que minimiza β ES, no es necesario trabajar directamente con la función $\phi_{\beta}(x)$, lo cual es difícil, debido a la naturaleza de su definición en términos del valor $\alpha_{\beta}(x)$ de β VaR y de las propiedades matemáticas de ese valor. En cambio, se puede trabajar con la expresión $F_{\beta}(\mathbf{x};\alpha)$ la cual es convexa con respecto a α , e incluso, regularmente, con respecto a $(\mathbf{x};\alpha)$.

La optimización apoyada por el Teorema 2, puede combinarse con la aproximación de la función $F_{\beta}(\mathbf{x};\alpha)$. La convexidad de $f(\mathbf{x};\mathbf{y})$ con respecto a \mathbf{x} , genera la convexidad de la aproximación de $F_{\beta}(\mathbf{x};\alpha)$ en (51).

El problema de minimizar F_{β} en $X \times \mathfrak{R}$ entra en la categoría de la llamada optimización estocástica, o más específicamente, de programación estocástica, debido a la presencia de una "esperanza" en la definición de $F_{\beta}(\mathbf{x};\alpha)$. Existe literatura, por lo menos para los casos que involucran la convexidad, para resolver tales problemas (Birge y Louveaux [10], Ermoliev y West [31], Kall y Wallace [46], Kan y Kibzun [47], Pflug [67], Prekopa [68]).

Hasta ahora la solución al problema se ha dirigido a minimizar β ES, o en otros términos, al problema

$$(A) \text{ Min } \phi_{\beta}(x) \quad \text{sobre } x \in X;$$

en base al Teorema 2, cuando F_{β} se minimiza sobre $X \times \mathfrak{R}$. Sin embargo, el problema relacionado a encontrar un portafolio que minimiza el β VaR (Kast, Luciano y Peccati [48], Mauser y Rosen [55]), es decir que resuelva el problema

$$(B) \text{ Min } \alpha_{\beta}(x) \quad \text{sobre } x \in X;$$

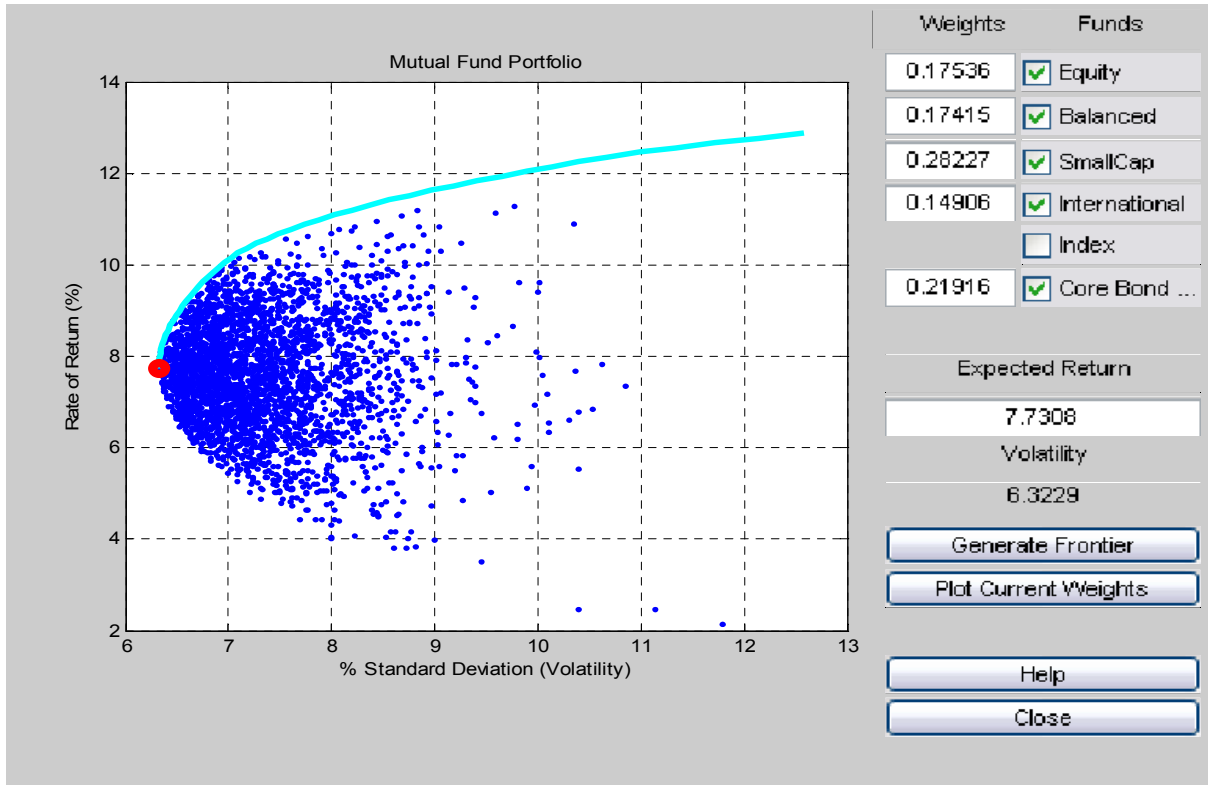
no se cubre directamente, porque $\phi_{\beta}(x) \geq \alpha_{\beta}(x)$, sin embargo, de acuerdo al Teorema 2, minimizar $F_{\beta}(\mathbf{x};\alpha)$ sobre $X \times \mathfrak{R}$ para resolver (A) también determina el β VaR de un portafolio \mathbf{x}^* que minimiza β ES y aunque sería mejor resolver (B), en términos de administración de riesgos, parece ser mejor resolver el problema (A) que el problema (B).

En este último sentido es útil comparar (A) y (B) con el problema de minimizar la varianza (vea Markowitz [53]):

$$(C) \text{ Min } \sigma^2(x) \quad \text{sobre } x \in X;$$

Un rasgo matemático del problema (C) es que se reduce a un problema de programación cuadrática, pero al igual que (B), la solución a (C) ha sido cuestionada por su conveniencia.

Jonas Palmquist, Stanislav Uryasev, and Pavlo Krokmal [44], han demostrado que los tres problemas (A), (B) y (C) son equivalentes, en el sentido de que producen la misma frontera eficiente, la cual consiste de los puntos, tales que dado el rendimiento y el riesgo de un punto en la frontera, este no puede tener igual o mayor rendimiento con un riesgo menor, ni igual o menor riesgo con un rendimiento mayor, como se muestra en la siguiente gráfica donde los puntos frontera corresponden a la línea color fucsia.



Teorema 3. Considere la función $\phi_\beta(x)$ y la función de rendimientos $R(x)$, ambas dependientes del vector de decisión \mathbf{x} , y considere los tres problemas siguientes:

$$(P1) \min_x \phi_\beta(x) - \mu_1 R(x) \quad x \in X, \quad \mu_1 \geq 0$$

$$(P2) \min_x \phi_\beta(x), \quad R(x) \geq \rho \quad x \in X,$$

$$(P3) \min_x -R(x), \quad \phi_\beta(x) \leq w, \quad x \in X,$$

las fronteras eficientes riesgo-rendimiento para tales problemas se pueden identificar, variando los parámetros μ_1 , ρ y w . Si $\phi_\beta(x)$ es convexa, $R(x)$ cóncava y el conjunto X es convexo, entonces los tres problemas generan la misma frontera eficiente.

El teorema 3 permite formular el problema de cualquier de las tres formas para generar fronteras eficientes, y dado que el teorema 2 permite que la función $F_\beta(\mathbf{x};\alpha)$ sea usada en lugar de $\phi_\beta(x)$ para resolver el problema (P2), también se puede usar dicha función en lugar de $\phi_\beta(x)$ en los problemas (P1) y (P3)

Teorema 4. Los siguientes problemas de minimización

$$(P3) \min_x -R(x), \quad \phi_\beta(x) \leq w, \quad x \in X,$$

y

$$(P3') \min_{(x,\alpha) \in X \times \mathfrak{R}} -R(x), \quad F_\beta(x,\alpha) \leq w, \quad x \in X,$$

son equivalentes en el sentido de que producen el mismo valor mínimo. Además, si la restricción ES en (P3) esta activa, el par (\mathbf{x}^*, α^*) , contiene el mínimo de (P3') si y solo si \mathbf{x}^* contiene el mínimo de (P3) y $\alpha^* \in A_\beta(\mathbf{x}^*)$. En particular, cuando el conjunto $A_\beta(\mathbf{x}^*)$ se reduce a un solo punto, minimizar $-R(\mathbf{x})$ sobre (\mathbf{x}, α) produce un par (\mathbf{x}^*, α^*) , tal que \mathbf{x}^* minimiza los rendimientos y α^* da el correspondiente β VaR

Teorema 5. Los dos siguientes problemas de minimización

$$(P1) \min_x \phi_\beta(x) - \mu_1 R(x) \quad x \in X, \quad \mu_1 \geq 0$$

y

$$(P1') \min_{(x,\alpha) \in X \times \mathfrak{R}} F_\beta(x,\alpha) - \mu_1 R(x) \quad x \in X, \quad \mu_1 \geq 0$$

son equivalentes en el sentido de que producen el mismo valor mínimo. Además, el par (\mathbf{x}^*, α^*) contiene el mínimo de (P1') si y solo si \mathbf{x}^* contiene el mínimo de (P1) y $\alpha^* \in A_\beta(\mathbf{x}^*)$. En particular, cuando el conjunto $A_\beta(\mathbf{x}^*)$ se reduce a un punto, minimizar $F_\beta(\mathbf{x};\alpha)$ sobre (\mathbf{x}, α) en $X \times \mathfrak{R}$ produce un par (\mathbf{x}^*, α^*) tal que \mathbf{x}^* minimiza $\phi_\beta(x) - \mu_1 R(x)$ y α^* da el correspondiente β VaR.

La formulación de problemas equivalentes presentados en los teoremas 2, 4 y 5, pueden combinarse con la aproximación de la integral $F_\beta(\mathbf{x};\alpha)$, a fin de obtener restricciones lineales que son más fáciles de manejar.

De acuerdo al teorema 4 podemos reemplazar la restricción

$$\phi_\beta(x) \leq w$$

en el problema de optimización (P3) por la restricción

$$F_{\beta}(\mathbf{x};\alpha) \leq w:$$

restricción que puede ser aproximada por

$$\tilde{F}_{\beta} \leq w$$

y reducida, de acuerdo a la ecuación (51), al siguiente sistema de restricciones lineales

$$\alpha + \frac{1}{k(1-\beta)} \sum_{i=1}^k z_i \leq w \quad (53)$$

$$z_j \geq f(\mathbf{x}, \mathbf{y}_j) - \alpha \quad z_j \geq 0 \quad j = 1, \dots, k \quad (54)$$

A fin de definir con mayor claridad el problema de programación lineal, que nos permite minimizar el ES, considere un portafolio de n instrumentos financieros diferentes y sea $\mathbf{x}^{\circ} = (x^{\circ}_1, x^{\circ}_2, \dots, x^{\circ}_n)$ el vector con los porcentajes iniciales del portafolio y sea $\mathbf{x} = (x_1, x_2, \dots, x_n)$ el vector con los porcentajes óptimos de cada instrumento en el portafolio y que es lo que queremos encontrar.

Los precios iniciales para cada instrumento están dados por $\mathbf{q} = (q_1, q_2, \dots, q_n)$, por lo que el producto de $\mathbf{q}'\mathbf{x}^{\circ}$ es el valor inicial del portafolio, los precios para cada instrumento al final del periodo están dados por $\mathbf{y} = (y_1, y_2, \dots, y_n)$, con estos supuestos tenemos que la función de perdida en el periodo es

$$f(\mathbf{x}, \mathbf{y}; \mathbf{x}^{\circ}, \mathbf{q}) = \mathbf{q}'\mathbf{x}^{\circ} - \mathbf{y}'\mathbf{x}$$

y tomemos el problema de optimización (P3'), es decir minimizaremos el negativo de los rendimientos, que es igual a maximizar los mismos, sujeto a que el ES sea menor a un cierto umbral (w), con la función $F_{\beta}(\mathbf{x};\alpha)$ de acuerdo a la ecuación (51) y linealizada como el conjunto de restricciones (53) y (54).

La función objetivo definida como los rendimientos sobre el periodo de optimización, la podemos considerar como el valor esperado al final del periodo dividido por el valor inicial, que podemos expresar como.

$$R(\mathbf{x}) = \frac{1}{\mathbf{q}'\mathbf{x}^{\circ}} \sum_{i=1}^n E(\mathbf{y}_i) x_i$$

Uno de los n instrumentos, puede ser considerado como dinero efectivo, con lo que el modelo, puede ser usado sin modificaciones cuando el inversionista tiene, solo efectivo en el periodo inicial, además de permitir conservar efectivo para los siguientes periodos. Dicho efectivo, también genera un cierto rendimiento.

En el planteamiento del problema, se supondrá un costo de transacción lineal y proporcional al valor de los instrumentos comercializados, es decir, con cada instrumento tenemos asociado un costo de transacción c_i , que se paga cuando se compra o se vende

el instrumento i , considerando que para el instrumento al que asignamos el efectivo el costo es igual a cero,

Dados esos costos, tenemos una restricción de balance como sigue

$$\sum_{i=1}^n q_i x_i^0 = \sum_{i=1}^n c_i q_i |x_i^0 - x_i| + \sum_{i=1}^n q_i x_i$$

Dado que supusimos una cuota comercial proporcional, el problema puede ser replanteado usando el siguiente conjunto de restricciones lineales

$$\sum_{i=1}^n q_i x_i^0 = \sum_{i=1}^n c_i q_i (\underline{\delta}_i + \overline{\delta}_i) + \sum_{i=1}^n q_i x_i$$

$$x_i^0 - \underline{\delta}_i + \overline{\delta}_i = x_i \quad i=1, \dots, n$$

$$\underline{\delta}_i \geq 0, \quad \overline{\delta}_i \geq 0, \quad i=1, \dots, n$$

La restricción no lineal $\underline{\delta}_i \overline{\delta}_i = 0$ puede omitirse, toda vez que la compra y venta del mismo instrumento i , no puede ser óptima.

También se puede considerar que los cambios de posición pueden ser acotados, donde la cota, puede ser por ejemplo un número fijo o ser proporcional a la posición inicial del instrumento.

$$0 \leq \underline{\delta}_i \leq \underline{\delta}_i^{\max}, \quad 0 \leq \overline{\delta}_i \leq \overline{\delta}_i^{\max}, \quad i=1, \dots, n$$

Asimismo, se puede considerar, que las posiciones algunas veces pueden ser acotadas

$$\underline{x}_i \leq x_i \leq \overline{x}_i \quad i=1, \dots, n$$

Además no se permitirá que un instrumento i tenga más que un porcentaje v_i dado del total del valor del portafolio

$$q_i x_i \leq v_i \sum_{k=1}^n x_k q_k \quad i=1, \dots, n$$

Esta restricción tiene sentido solo cuando posiciones cortas no son permitidas.

Por otra parte, las regulaciones existentes imponen requerimientos de capital, proporcionales al VaR de un portafolio. Estos requisitos pueden ser forzados por

restricciones sobre el ES, dado que $ES \geq VaR$, la cota superior de ES (w) puede seleccionarse como el máximo VaR,

$$\alpha + \frac{1}{k(1-\beta)} \sum_{i=1}^k z_i \leq w$$

$$z_j \geq \sum_{i=1}^n (-y_{ij} x_i + q_i x_i^0) - \alpha \quad z_j \geq 0, \quad j=1, \dots, k$$

De acuerdo con lo anterior, tenemos que el problema de minimizar el negativo de los rendimientos, con un ES menor a un umbral (w) se plantea como un problema de Programación lineal como sigue:

$$\min_{x, \alpha} \frac{1}{q'x^0} \sum_{i=1}^n E(y_i) x_i$$

sujeto a las restricciones

$$\alpha + \frac{1}{k(1-\beta)} \sum_{i=1}^k z_i \leq w$$

$$z_j \geq \sum_{i=1}^n (-y_{ij} x_i + q_i x_i^0) - \alpha \quad z_j \geq 0, \quad j=1, \dots, k$$

$$\sum_{i=1}^n q_i x_i^0 = \sum_{i=1}^n c_i q_i (\underline{\delta}_i + \overline{\delta}_i) + \sum_{i=1}^n q_i x_i$$

$$x_i^0 - \underline{\delta}_i + \overline{\delta}_i = x_i \quad i=1, \dots, n$$

$$0 \leq \underline{\delta}_i \leq \underline{\delta}_i^{\max}, \quad 0 \leq \overline{\delta}_i \leq \overline{\delta}_i^{\max}, \quad i=1, \dots, n$$

$$\underline{x}_i \leq x_i \leq \overline{x}_i \quad i=1, \dots, n$$

$$x_i \leq v_i \sum_{k=1}^n x_k q_k \quad i=1, \dots, n$$

Resolviendo este problema, se obtiene el vector óptimo x^* , así como el correspondiente VaR (α^*), y el máximo rendimiento esperado, así mismo resolviendo este mismo problema para diferentes ESs del portafolio, es decir para diferentes valores de w , se puede obtener la frontera eficiente para el portafolio.

Para n instrumentos y k escenarios, la formulación del problema de programación lineal presentado tiene $3n + k + 1$ variables y $2(n + 1) + k$ restricciones, sin incluir las cotas sobre $\underline{\delta}_i$, $\bar{\delta}_i$ y \mathbf{x}_i . El número de coeficientes distintos de cero en las restricciones es de $6n + nk + 3k + 1$.

Anteriormente se señaló que la función definida en (50) es aproximada por la suma sobre todos los escenarios, los cuales pueden ser generados como sigue:

- Si se supone una distribución conjunta para los rendimientos de todos los instrumentos, se pueden generar los escenarios con simulaciones de Monte Carlo
- También se pueden usar los datos históricos sin asumir distribución alguna, dividiendo las series en J periodos (escenarios), calculando el rendimiento sobre cada uno de esos periodos. La longitud de el periodo Δt , es igual a la longitud de el periodo sobre el cual se esta optimizando el portafolio.

Por ejemplo, con datos históricos, para optimizar sobre periodos de un día, si se tienen series de precios diarios, se toman los precios al cierre del día, de dos periodos consecutivos, p_t y p_{t+1} . De manera similar, para periodos de dos semanas, se toman los precios de p_t y de p_{t+10} , con estos pares se generan los escenarios para el instrumento i como sigue:

$$r_{i,j} = \frac{p_{i,t+\Delta t}}{p_{i,t}}$$

para cada periodo (escenario) se puede calcular el rendimiento esperado como:

$$E[r_i] = \frac{1}{k} \sum_{j=1}^k r_{i,j}$$

Estos rendimientos pueden ser usados para calcular los precios en cada escenario y el precio esperado para cada instrumento al final del periodo sobre el cual se esta optimizando como sigue:

$$y_{i,j} = q_i r_{i,j}$$

$$E[y_i] = \frac{1}{k} \sum_{j=1}^k y_{i,j}$$

7. CONCLUSIONES

En este trabajo se presentaron conceptos y metodologías básicas para calcular el VaR y sus posibles alternativas, utilizando metodologías paramétricas y no-paramétricas, asimismo, se incluyó un capítulo para la optimización del portafolio relacionado con el valor en riesgo. Todo ello, con el propósito de contribuir a la divulgación de estas técnicas, así como de sus objetivos y los supuestos en los que se sustentan.

En general se puede concluir que las metodologías para calcular el VaR utilizan herramientas bastante sofisticadas y complejas que al ser usadas requieren no sólo habilidades matemáticas, sino también mucho cuidado en la interpretación de los resultados y transparencia al declarar sus alcances. Como cualquier herramienta de análisis sofisticado, el cálculo de VaR puede proporcionar información precisa cuando se usa correctamente, pero puede ser muy peligrosa cuando no se atienden los supuestos subyacentes o se realizan aproximaciones no validadas tanto metodológicamente como por la experiencia y la evidencia empírica.

También es claro que, una buena administración de riesgos, requiere no solo la aplicación aislada de una sola metodología al riesgo de mercado, sino la aplicación de varias de ellas, no solo para confirmar o reconfirmar los resultados, sino para redondear o complementar las deficiencias de una u otra metodología.

Así mismo, la instrumentación de cualquiera de estas metodologías requiere de la existencia de una base de datos de todos los posibles instrumentos, toda vez que todas las metodologías, requieren de series de al menos dos años de los instrumentos que se están considerando, así como de aquellos que en un momento dado, se piensa sean parte del portafolio, para obtener resultados confiables.

Un requisito extra para la aplicación de cualquiera de estas metodologías, es el conocimiento sobre los instrumentos, toda vez que la valuación del portafolio, requiere del cálculo de precios de los instrumentos, sobre todo de aquellos como las opciones, bonos y derivados, que en general, no es simple, aún y cuando existen formulas al respecto, pero que regularmente son aproximadas.

Finalmente, es necesario estar claro de que los resultados que se obtienen con cualquier metodología, son solo estimaciones de lo que puede ocurrir y el nivel o pérdida en que se puede incurrir, lo cual no necesariamente coincide con la realidad.

REFERENCIAS

- [1] Acerbi C., Nardio C., Sirtori C. (2001) Expected Shortfall as a Tool for Financial Risk Management. Working paper <http://www.gloriamundi.org/var/wps.html>
- [2] Andersen T.G. and T. Bollerslev (1997), "Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *Journal of Finance*, 52, 975-1005.
- [3] Andersen T.G., Bollerslev T., Christoffersen P.F., and Diebold F.X. (2005), "Practical Volatility and Correlation Modeling for Financial Market Risk Management," in M. Carey and R. Stulz (eds.), *Risks of Financial Institutions*. Chicago: University of Chicago Press for National Bureau of Economic Research.
- [4] Andrews D.W.K. (1988), Laws of large numbers for dependent non-identically distributed random variables, *Econometric Theory*, 4: 458-467.
- [5] Artzner P., Delbaen F., Eber J. and Heath D. (1997), Thinking Coherently, *RISK* 10.
- [6] Artzner P., Delbaen F., Eber J. and Heath D. (1999) Coherent measures of risk. *Mathematical Finance* 9(3), 203-228.
- [7] Baillie R.T. Bollerslev T. and Mikkelsen H.O. (1996), "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 74, 3-30.
- [8] Basle Committee on Banking Supervision (1996), Amendment to the Capital Accord to Incorporate Market Risks, Basle, January.
- [9] Beder T.S. (1995), VaR: Seductive but Dangerous, *Financial Analyst Journal*, Sep-Oct,
- [10] Birge J.R. and Louveaux F. (1997). *Introduction to Stochastic Programming*, Springer Verlag, pp. 448.
- [11] Bollerslev T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31
- [12] Bollerslev T, Engle R.F. and Nelson D.B. 1994, ARCH models, in Engle R.F. and McFadden eds., *Handbook of Econometrics*, Vol. 4, Elsevier Science B.V. Amsterdam
- [13] Bollerslev T. and Wooldridge J.M. (1992), Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances, *Econometric Reviews* 11
- [14] Bollerslev T., Engle R.F. and Wooldridge J.M. (1988), "A Capital Asset Pricing Model with Time Varying Covariances," *Journal of Political Economy*, 96, 116-131.
- [15] Boudoukh J., Richardson M. and Whitelaw R.F. (1998). The best of both worlds, *Risk*, 11: 64-67.
- [16] Buchinsky M. (1998), Recent Advances in Quantile Regression Models: a Practical Guideline for Empirical Research. *Journal of Human Resources* 33.
- [17] Cardenas J, Fruchard E., Picron J.F., Reyes C., Walters K., and Yang W. (1999) Monte Carlo within a day *Risk* 12:2
- [18] Coles Stuart (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London.
- [19] Chernozhukov V. and L. Umantsev (2000), Conditional Value-at-Risk: Aspects of Modelling and Estimation, Stanford University
- [20] Danielsson, J. and C. de Vries (1998), Beyond the Sample: Extreme Quantile and Probability Estimation, London School of Economics, Discussion Paper 298.
- [21] Danielsson Jon y Casper G. de Vries. (1997). "Tail Index and Quantile Estimation with Very High Frequency Data". *Journal of Empirical Finance*. Vol. 4, pp. 241-257.
- [22] Danielsson Jon, L. de Haan, L. Peng y C. de Vries (2001). "Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation." *Journal of Multivariate Analysis*. No. 76, pp. 226-248.
- [23] Ding Z., Granger C.W.J. and Engle R.F. (1993), "A Long Memory Property of Stock Market Returns and a New Model," *Journal of Empirical Finance*, 1, 83-106.
- [24] Diebold F., Schuermann T. and Stroughair J. (1999), Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management, in *Advances in Computational Finance*, Kluwer Academic Publishers, Amsterdam.
- [25] Elton E.J. Gruber M.J. y Michaely R. (1990): "The Structure of Spot Rates and Immunization", *Journal of Finance*, Vol. XLV, 2, Junio. 629-642.
- [26] Embrechts P., Klüppelberg C. y Mikosch T. (1997) *Modelling Extremal Events for Insurance and Finance*. Berlin. Springer-Verlag
- [27] Engle R.F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50.
- [28] Engle R.F. and G. Gonzalez-Rivera (1991), Semiparametric ARCH Models, *Journal of Business and Economic Statistics* 9
- [29] Engle R.F. and S. Manganelli (1999), CAViaR: Conditional Value at Risk by Quantile Regression, NBER, Working Paper 7341.
- [30] Engle R. F. and V. Ng (1993) Measuring and Testing the Impact of News On Volatility, *Journal of Finance*, 48: 1749-1778.

- [31] Ermoliev Yu. and Wets, R. J-B. (Eds.) (1988). Numerical Techniques for Stochastic Optimization, Springer Series in Computational Mathematics, 10.
- [32] Fama E. (1965), The Behavior of Stock Market Prices, *Journal of Business* 38.
- [33] Glasserman P., Heidelberger P., and Shahabuddin P. (1999a) Importance sampling and stratification for value-at-risk, in Computational Finance 1999 (Proceedings of the Sixth International Conference on Computational Finance), Y.S. Abu-Mostafa, B. LeBaron, A.W. Lo, and A.S.Weigend, eds., MIT Press, Cambridge, Mass.
- [34] Glasserman P., Heidelberger P., and Shahabuddin P. (1999b) Stratification issues in estimating value-at-risk, in Proceedings of the 1999 Winter Simulation Conference, pp351-358, IEEE Computer Society Press, Piscataway, New Jersey.
- [35] Glasserman P., Heidelberger P., and Shahabuddin P. (1999c) Variance reduction techniques for estimating value-at-risk," IBM Research Report RC 21577, Yorktown Heights, New York. Also available at <http://www.ssrn.com> To appear in Management Science.
- [36] Glosten L.R., Jagannathan R. and Runkle D. (1993), "On the Relation Between the expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779-1801.
- [37] Gouriéroux, C. and J. Jasak (1998), Truncated Maximum Likelihood, Goodness of Fit Tests and Tail Analysis. Unpublished Manuscript.
- [38] Granger C.W.J. White H. and Kamstra M. (1989), Interval forecasting. An analysis based upon ARCH-quantile estimators, *Journal of Econometrics*, 40: 87-96.
- [39] Hendricks D. (1996): Evaluation of Value-at-Risk Models Using Historical Data. *Federal Reserve Bank of New York Economic Policy Review*, Vol.2.
- [40] Hendricks D. and Hirtle B. (1997). Bank Capital Requirements for Market Risk: The Internal Models Approach. *Federal Reserve Bank of New York. Economic Policy Review* 3
- [41] Hogg R., and Klugman S. (1984), *Loss Distributions*, Wiley, New York.
- [42] Huber P. (1981), *Robust Statistics*, John Wiley and Sons, New York.
- [43] Hull J. and White A. (1998), Incorporating Volatility Updating into the Historical Simulation Method for VaR, *The Journal of Risk* 1.
- [44] Jonas Palmquist, Stanislav Uryasev, and Pavlo Krokmal Portfolio Optimization With Conditional Value-At-Risk Objective And Constraints Research Report #99-14 Center for Applied Optimization Dept. of Industrial and Systems Engineering University of Florida, Gainesville, FL 32611 Date: November 20, 1999
- [45] Jorion P. In Defense of VaR, <http://www.gsm.uci.edu/jorion/oc/ntalib2.html> 1997.
- [46] Kall P. and Wallace S.W. (1995). Stochastic Programming, John Wiley & Sons, pp. 320.
- [47] Kan Y.S., and Kibzun A.I. (1996). Stochastic Programming Problems with Probability and Quantile Functions, John Wiley & Sons, pp. 316.
- [48] Kast R., Luciano E., and Peccati L. (1998). VaR and Optimization. 2nd International Workshop on Preferences and Decisions, Trento, July 13 1998.
- [49] Konno H. and Yamazaki H. (1991). Mean Absolute Deviation Portfolio Optimization Model and Its Application to Tokyo Stock Market. *Management Science*. 37, 519-531.
- [50] Leadbetter M.R., Lindgren G. and Rootzen H. (1983), *Extremes and Related Properties of Random Sequences and Processes*, New York : Springer-Verlag.
- [51] Mandelbrot B. (1963). The Variations of Certain Speculative Prices, *Journal of Business*:
- [52] Manganelli S. Y. and Engle R. (2001). "Value at Risk Models in Finance." Working Paper N°75, European Central Bank.
- [53] Markowitz H.M. (1952). Portfolio Selection. *Journal of finance*. Vol.7, 1, 77-91.
- [54] Martin R.D. y Zamar R. (1993): Efficiency-constrained bias-robust estimation of location. *Annals of Statistics*, Vol.21, 1.
- [55] Mausser H. and Rosen D. (1999). Beyond VaR: From Measuring Risk to Managing Risk. *ALGO Research Quarterly*. Vol.1, 2, 5{20.
- [56] McNeil A.J. and Frey R. (2000), Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach, *Journal of Empirical Finance*,
- [57] Morgan J.P. y Reuters (1996): RiskMetrics. *Technical Document*.
- [58] Navarro E. y Nave J.M. (1997): A Two Factor Duration Model for Interest Rate Risk Management. *Investigaciones Económicas*. Vol. XXI. Enero. 55-74.
- [59] Navarro E. y Nave J.M. (1998): Un modelo para la evaluación de la gestión de carteras de renta fija. Documentos de Trabajo. Facultad de Ciencias Económicas y Empresariales de Albacete. Universidad de Castilla-La Mancha. Doc. 1998.
- [60] Nelson D.B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347-370.
- [61] Newey W. and Powell J. (1991), Efficient Estimation of Linear and Type I Censored Regression Models Under Conditional Quantile Restrictions, *Econometric Theory*

- [62] Pickands J. (1975), Statistical Inference Using Extreme Order Statistics, *The Annals of Statistics* 3:
- [63] Powell J. (1984), Least absolute deviations estimation for the censored regression model, *Journal of Econometrics*, 25: 303-325.
- [64] Powell J. (1986), Censored regression quantiles, *Journal of Econometrics*, 32: 143-155.
- [65] Price K. and Storn R. (1997), Differential Evolution, *Dr. Dobb's Journal*, April, 18-24.
- [66] Pflug G.Ch. (1996). Optimization of Stochastic Models: The Interface Between Simulation and Optimization. Kluwer Academic Publishers, Dordrecht, Boston.
- [67] Pflug G.Ch. (2000). Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk. In."Probabilistic Constrained Optimization: Methodology and Applications", Ed. S. Uryasev, Kluwer Academic Publishers, 2000
- [68] Prekopa A. (1995). Stochastic Programming. Kluwer Academic Publishers, Dordrecht, Boston.
- [69] Rockafellar R.T. (1970). Convex Analysis. Princeton Mathematics, Vol. 28, Princeton Univ.Press.
- [70] Rockafellar R. Tyrrell and Stanislav Uryasev (1999). Optimization of Conditional Value-at-Risk
- [71] Rousseeuv P.J. and Leroy A.M. (1987): Robust Regression and Outlier Detection. Ed. John Wiley and Sons.
- [72] Shor N.Z. (1985). Minimization Methods for Non-Differentiable Functions. Springer-Verlag.
- [73] Taleb N. Against Value at Risk: Nassim Taleb Replies to Philippe Jorion, <http://pw1.netcom.com/ntaleb/jorion.html> 1997.
- [74] Taylor S.J. (1986) *Modeling Financial Time Series*. Chichester, UK: John Wiley and Sons.
- [75] Uryasev S. (1995). Derivatives of Probability Functions and Some Applications. *Annals of Operations Research*, V56, 287-311.
- [76] White H. (1994), *Estimation, Inference and Specification Analysis*, Cambridge University Press
- [77] White H and Domowitz I. (1984) Nonlinear regression with dependent observations, *Econometrica* 52:143-161
- [78] Young M.R. (1998). A Minimax Portfolio Selection Rule with Linear Programming Solution. *Management Science*. Vol.44, No. 5, 673-683.
- [79] Zakořian, J.M. (1994), "Threshold Heteroskedastic Models," *Journal of Economic Dynamics and Control*, 18, 931-955.