



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

UNA APLICACIÓN DE MEDIDAS DE RIESGO Y ÓRDENES
ESTOCÁSTICOS

T E S I S

QUE PARA OBTENER EL GRADO DE:

ACTUARIA

PRESENTA:

BRENDA LILIANA TRANSITO PIMENTEL

DIRECTOR DE TESIS:

DR. SERGIO IVÁN LÓPEZ ORTEGA



CIUDAD UNIVERSITARIA, CDMX

JUNIO, 2019



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.

Una aplicación de Medidas de Riesgo y Órdenes estocásticos

por

Brenda Liliana Transito Pimentel

Tesis presentada para obtener el grado de

Actuaria

en la

FACULTAD DE CIENCIAS

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad Universitaria, CDMX. Junio, 2019

Índice general

1. Órdenes estocásticos	4
1.1. Primer Orden Estocástico	4
1.2. Uso de Acoplamiento para demostrar Dominio Estocástico	8
1.3. Ordenamiento por Cociente de Verosimilitud	10
1.4. Problema de Inversión de un Período	12
1.5. Segundo Orden Estocástico	15
1.5.1. Variables Aleatorias Normales	17
1.5.2. Otros resultados de Segundo Orden Estocástico	20
2. Medidas de Riesgo	23
2.1. Introducción	23
2.2. Valor en Riesgo (VaR)	24
2.3. CVaR	25
3. Órdenes estocásticos y Medidas de Riesgo	29
3.1. Órdenes estocásticos	29
3.2. Medidas de Riesgo Coherentes	39
4. Optimización de Medidas de Riesgo	51
4.1. Metodología	53
4.2. Rendimientos	53
4.3. Series de Tiempo Multivariadas	54

4.3.1.	Estacionariedad débil y matrices de correlación cruzada	54
4.3.2.	Matrices de correlación cruzada	55
4.3.3.	Estimación de las Matrices de Correlación Cruzada	57
4.3.4.	Prueba de Portmanteau Multivariado	57
4.3.5.	Modelos de Vectores Autorregresivos	58
4.3.6.	Condiciones de estacionariedad y Momentos del Modelo VAR(1)	59
4.3.7.	Vector AR(p)	61
4.3.8.	Construcción del Modelo VAR(p)	63
4.3.9.	Predicción	66
4.3.10.	Comprobación de Supuestos	67
4.4.	Optimización por Simplex	69
4.4.1.	Método Simplex Revisado	71
4.4.2.	Algoritmo del método simplex revisado	72
4.5.	Ejemplo	73
4.6.	Resultados	74
A.	Apéndice	86
A.1.	Tipos de convergencia	86
A.2.	Convergencia	86
A.3.	Integrabilidad uniforme	87
B.	Código	91
B.1.	Minimizar CVaR de un portafolio con distribución Normal Multivariada	91
B.2.	Ajuste de Series de tiempo y simulación de datos	97
B.3.	Optimización del CVaR	103
B.4.	Optimización en t periodos	109

Índice de tablas

4-1. Tabla Simplex	71
4-2. Vector de medias	73
4-3. Matriz de Varianzas y Covarianzas	74
4-4. Tabla de rendimientos	74
4-5. C-VaR mínimo y ponderaciones en cada período para un portafolio de 3 activos .	81
4-6. C-VaR mínimo y ponderaciones en cada período para un portafolio con 9 acciones	83
4-7. C-VaR mínimo y ponderaciones en cada período para un portafolio con 9 acciones	83

Índice de figuras

3-1. Representación típica de segmentos de F_X y F_Y cuando $X \leq_{cx} Y$	35
4-1. Rendimientos del portafolio	75
4-2. Correlaciones del portafolio de 9 activos	76
4-3. Correlaciones del portafolio de 3 activos	77
4-4. Dinámica de las ponderaciones del portafolio de tres activos durante 10 días. . .	82
4-5. Dinámica de las ponderaciones del portafolio de nueve activos durante 10 días. .	84

Introducción

En 1952, Harry Markovitz propuso un modelo basado en encontrar las ponderaciones de un portafolio que minimizara la varianza y por mucho tiempo este modelo fue aceptado universalmente. Sin embargo, se comprobó que el modelo, más que medir el riesgo, medía la dispersión de los activos. Como consecuencia de las grandes crisis financieras, se planteó la necesidad de que las medidas de riesgo tenían que expresar las pérdidas potenciales a cierto nivel de confianza α y en un horizonte de tiempo dado. Por ello, surgieron las medidas de riesgo.

Otra manera de expresar lo anterior es la siguiente: Los inversionistas buscarán perder, en términos monetarios, lo menos posible. Es decir, si tienen una gama de activos disponibles en los que pueden invertir, siempre elegirán aquellos que les generen las mínimas pérdidas posibles. Es decir, desean “ordenar” las pérdidas generadas por los activos financieros con base en la aleatoriedad de sus comportamientos y así elegir el activo que les genere la menor pérdida. En esta interpretación entran naturalmente los órdenes estocásticos.

Si se comparan la pérdida entre un activo 1 y un activo 2, y se sabe que bajo un cierto “orden” (que es necesario definir con cuidado) las pérdidas generadas por el activo 1 son menores a las pérdidas generadas por el activo 2, es natural cuestionarse si al medir el riesgo asociado a estos activos el orden se preserva.

A lo largo de este trabajo, estudiamos los distintos órdenes estocásticos que existen, sus propiedades y las relaciones que existen entre ellos. También, se examinan las medidas de riesgo en general y los casos notables particulares dados por el VaR y el C-VaR. Se muestra la representación que tienen las medidas de riesgo que cumplen ciertas condiciones y se responde la pregunta formulada en el párrafo anterior. Finalmente, se presenta una implementación para

la optimización del C-VaR a un nivel de confianza α dado para portafolios de inversión.

Este trabajo consta de 4 capítulos y dos apéndices. En el primer capítulo se presenta la definición y la caracterización del primer orden estocástico. Posteriormente, se estudia otro criterio para saber cuándo una variable aleatoria domina en primer orden estocástico a otra y después se generaliza el primer orden estocástico para n variables. También se define el orden por cociente de verosimilitud, se estudia la relación de este orden con el primer orden estocástico y se ejemplifica con un problema de inversión de un periodo. Finalmente, se presenta la definición y caracterización del segundo orden estocástico y se estudia el caso particular de variables aleatorias gaussianas. En este primer capítulo, se ejemplifican a los órdenes estocásticos con variables aleatorias que representan rendimientos, sin embargo, posteriormente, se ejemplificarán como pérdidas.

En el segundo capítulo se presenta la definición formal de Medida de Riesgo y sus propiedades; se dan las condiciones para que las medidas de riesgo sean coherentes. Se definen el VaR y el C-VaR, y finalmente se demuestra que el C-VaR es una medida de riesgo coherente.

En el tercer capítulo, se muestra la relación existente entre Órdenes Estocásticos y Medidas de Riesgo para variables aleatorias definidas en L^1 , para ello se estudia la propiedad de Fatou para medidas de riesgo. El objetivo de este capítulo es estudiar cuándo las medidas de riesgo son consistentes con el primer orden estocástico, el orden convexo y el orden convexo creciente. Para demostrar la consistencia de este último orden, se extendió la representación de una medida de riesgo invariante bajo traslaciones y coherente con la propiedad de Fatou en L^∞ presentada en Kusuoka (2001) [7] al espacio L^1 .

En el cuarto capítulo se presenta la metodología para la optimización del C-VaR en t periodos, a un nivel de confianza α , de dos portafolios de rendimientos pertenecientes a acciones que cotizan en Bolsa Mexicana de Valores, el primer portafolio consiste de 3 activos y el segundo consiste de 9 activos. Dado que se necesitan q escenarios para la misma, se realizó un ajuste de series de tiempo multivariadas para posteriormente simular dichos escenarios. La optimización se traduce a un problema de programación lineal, que se resolvió mediante la implementación del método simplex revisado.

El apéndice A presenta conceptos de convergencia de variables aleatorias y algunos resultados de integrabilidad uniforme, que se utilizan en el capítulo 3. El apéndice B tiene los códigos de la implementación de la optimización en t periodos del C-VaR, así como el ajuste de series de tiempo multivariadas de los rendimientos que constituyen los portafolios.

Capítulo 1

Órdenes estocásticos

Los órdenes estocásticos tienen como objetivo dar una pauta para “ordenar” variables aleatorias. Imaginemos que somos inversionistas y que estamos por decidir entre invertir nuestro dinero en un activo que nos genera un rendimiento X o en un activo que genera un rendimiento Y . El primer orden estocástico nos dice que vamos a preferir la inversión Y siempre y cuando, la probabilidad de que Y exceda un nivel t , es al menos tan grande como la probabilidad de que X exceda ese mismo nivel.

1.1. Primer Orden Estocástico

En esta primera sección presentaremos la definición formal del primer orden estocástico, así como algunos resultados elementales del mismo.

Definición 1.1.1. Sean X y Y variables aleatorias, decimos que Y domina estocásticamente a X en primer orden, si para todo t ocurre que,

$$P(X > t) \leq P(Y > t).$$

Es decir, para cada t , es al menos tan probable que Y exceda el valor t como es que X lo haga; se denotará por $X \leq_{st} Y$.

Notemos que en el caso en que X y Y son variables continuas lo anterior es equivalente a que $P(X \geq t) \leq P(Y \leq t)$.

Lemma 1.1.1. *Si X es una variable aleatoria no negativa, entonces*

$$E(X) = \int_0^{\infty} P(X > t) dt.$$

Demostración.

Sea $t > 0$, definamos la siguiente variable aleatoria $I(t)$ como sigue

$$I(t) = \begin{cases} 1 & \text{si } t < X, \\ 0 & \text{si } t \geq X. \end{cases}$$

Entonces,

$$\int_0^{\infty} I(t) dt = \int_0^X I(t) dt + \int_X^{\infty} I(t) dt = X.$$

Por lo tanto,

$$E(X) = E\left(\int_0^{\infty} I(t) dt\right) = \int_0^{\infty} E(I(t)) dt = \int_0^{\infty} P(X > t) dt,$$

donde usamos el Teorema de Fubini en la primera igualdad de la expresión anterior. ■

Lemma 1.1.2. *Si $X \leq_{st} Y$ entonces $E(X) \leq E(Y)$.*

Demostración.

Supongamos que X y Y son variables aleatorias no negativas. Haciendo uso del Lema 1.1.1 y del hecho que X domina estocásticamente a Y ocurre que

$$E(X) = \int_0^{\infty} P(X > t) dt \leq \int_0^{\infty} P(Y > t) dt = E(Y).$$

Ahora, probemos el resultado en general. Notemos que cualquier número real a se puede

expresar como la diferencia de su parte positiva y su parte negativa, es decir,

$$a = a^+ - a^-,$$

donde

$$a^+ = \max(a, 0), \quad a^- = \max(-a, 0).$$

Entonces, podemos expresar a X y Y como diferencia de su parte positiva y negativa:

$$X = X^+ - X^-, \quad Y = Y^+ - Y^-.$$

Sea $t \geq 0$

$$\begin{aligned} P(X^+ > t) &= P(X > t) \\ &\leq P(Y > t) && (\text{pues } X \leq_{st} Y) \\ &= P(Y^+ > t). \end{aligned}$$

También ocurre que

$$\begin{aligned} P(X^- > t) &= P(-X > t) \\ &= P(X < -t) \\ &\geq P(Y < -t) && (\text{pues } X \leq_{st} Y) \\ &= P(-Y > t) \\ &= P(Y^- > t). \end{aligned}$$

Por lo tanto, $X^+ \leq_{st} Y^+$ y $X^- \geq_{st} Y^-$. Notemos que todas estas variables aleatorias son no negativas, haciendo uso del primer resultado tenemos que $E(X^+) \leq E(Y^+)$ y $E(X^-) \geq E(Y^-)$.

Por lo tanto

$$E(X) = E(X^+) - E(X^-) \leq E(Y^+) - E(Y^-) = E(Y).$$

■

Es natural pensar que si se prefiere la inversión en el instrumento Y sobre la inversión en el

instrumento X , el rendimiento promedio que nos da invertir en Y debe ser al menos tan grande como el rendimiento promedio que nos da invertir en X .

Con ayuda de los Lemas anteriores probaremos el siguiente resultado.

Teorema 1.1.1. $X \leq_{st} Y$ si, y sólo si, $E(h(X)) \leq E(h(Y))$ para toda función h no decreciente.

Demostración.

\Rightarrow) Sea h una función creciente ¹ y supongamos que $X \leq_{st} Y$, primero probaremos que $h(X) \leq_{st} h(Y)$.

Como h es una función creciente, para cada t , existe un valor al que llamaremos

$$h^{-1}(t) = \inf\{x \in \mathbb{R} : h(x) \leq t\},$$

de modo que el evento $h(X) > t$ es equivalente al evento $X > h^{-1}(t)$. Así,

$$\begin{aligned} P(h(X) > t) &= P(X > h^{-1}(t)) \\ &\leq P(Y > h^{-1}(t)) \quad (\text{pues } X \leq_{st} Y) \\ &= P(h(Y) > t) \end{aligned}$$

Con lo que probamos que $h(X) \leq_{st} h(Y)$. Haciendo uso del Lema 1.1.2, tenemos que $E(h(X)) \leq E(h(Y))$.

\Leftarrow) Supongamos que $E(h(X)) \leq E(h(Y))$ para toda función creciente h . Para un t fijo, definamos la función h_t como sigue;

$$h_t(x) = \begin{cases} 0 & \text{si } x \leq t, \\ 1 & \text{si } x > t. \end{cases}$$

Notemos que h_t es una función creciente así que

$$E(h_t(X)) \leq E(h_t(Y)).$$

¹Durante toda la tesis utilizaremos crecientes para referirnos a funciones no decrecientes. En el caso en que la función sea estrictamente creciente lo diremos explícitamente.

Por otro lado, tenemos que

$$P(X > t) = E(h_t(X)) \leq E(h_t(Y)) = P(Y > t),$$

por lo que $X \leq_{st} Y$. ■

1.2. Uso de Acoplamiento para demostrar Dominio Estocástico

Un método para mostrar que $X \leq_{st} Y$ es encontrando dos variables aleatorias X' y Y' que tengan la misma distribución que X y Y respectivamente y que sean tales que $X' < Y'$ (es decir, que $X(\omega) \leq Y(\omega)$ para todo ω salvo en un conjunto de probabilidad cero). Para tales variables ocurre que si $Y' > t$ entonces $X' > t$, por lo que

$$P(X' > t) \leq P(Y' > t).$$

Por construcción tenemos que $P(X' > t) = P(X > t)$ y $P(Y' > t) = P(Y > t)$, probando así que $X \leq_{st} Y$.

A este proceso se le conoce como acoplamiento.

Lemma 1.2.1. *Si F es una función de distribución continua y U es una variable aleatoria uniforme en $(0, 1)$ entonces, la variable aleatoria $F^{-1}(U)$ tiene función de distribución F , donde F^{-1} es definida por*

$$F^{-1}(u) = \inf\{t \in \mathbb{R} : F(t) \geq u\},$$

y cumple que $F(F^{-1}(u)) = u$.

Demostración.

F es creciente pues es función de distribución, lo que implica que $a \leq x$ y $F(a) \leq F(x)$ sean

equivalentes. Por lo tanto,

$$\begin{aligned} P(F^{-1}(U) \leq x) &= P(F(F^{-1}(U)) \leq F(x)) \\ &= P(U \leq F(x)) \\ &= F(x). \end{aligned}$$

■

Proposición 1.2.1. *Dos variables aleatorias X y Y satisfacen $X \leq_{st} Y$ si y sólo si, existen variables aleatorias X' y Y' con la misma distribución que X y Y respectivamente, tales que $X' \leq Y'$.*

Demostración.

\Rightarrow) Supongamos que $X \leq_{st} Y$. Sean X y Y variables aleatorias, con funciones de distribución F_X y F_Y respectivamente. Dado que $X \leq_{st} Y$ ocurre que $F_X(x) \geq F_Y(x)$ para toda x , entonces por lo que

$$F_Y(F_X^{-1}(u)) \leq F_X(F_X^{-1}(u)) = u = F_Y(F_Y^{-1}(u)),$$

de lo que se sigue que

$$F_X^{-1}(x) \leq F_Y^{-1}(x) \quad (\text{pues } F_Y \text{ es creciente}) \quad (1-1)$$

Sea U una variable aleatoria con distribución Uniforme en $(0, 1)$ y sea $X' = F_X^{-1}(U)$ así como $Y' = F_Y^{-1}(U)$. Por el resultado de la ecuación (1-1) tenemos que $X' \leq Y'$ y haciendo uso del Lema 1.2.1 tenemos que X' y Y' tienen la misma distribución de X y Y respectivamente.

\Leftarrow) Supongamos que $X' \leq Y'$. Si $Y' > t$ entonces $X' > t$, por lo tanto

$$P(X' > t) \leq P(Y' > t).$$

Por hipótesis tenemos que $P(X' > t) = P(X > t)$ y $P(Y' > t) = P(Y > t)$, probando así que $X \leq_{st} Y$.

■

Teorema 1.2.1. Sean (X_1, \dots, X_n) y (Y_1, \dots, Y_n) vectores de variables aleatorias independientes e independientes entre vectores. Suponga que $X_i \leq_{st} Y_i$ para cada $i = 1, \dots, n$. Entonces ocurre que $g(X_1, \dots, X_n) \leq_{st} g(Y_1, \dots, Y_n)$ para toda $g(x_1, \dots, x_n)$ creciente en cada componente.

Demostración.

Sea $g(x_1, \dots, x_n)$ una función creciente en cada componente. Sea F_{X_i} la función de distribución de X_i y F_{Y_i} la función de distribución de Y_i para cada $i = 1, \dots, n$. Sean U_1, \dots, U_n variables aleatorias independientes con distribución Uniforme en $(0, 1)$ y definamos

$$X'_i = F_{X_i}^{-1}(U_i), \quad Y'_i = F_{Y_i}^{-1}(U_i), \quad i = 1, \dots, n.$$

Haciendo uso de la Proposición 1.2.1 tenemos que $X'_i \leq Y'_i$ para cada i , por ser g creciente en cada componente, se sigue que $g(X'_1, \dots, X'_n) \leq g(Y'_1, \dots, Y'_n)$. Finalmente, notemos que por el Lema 1.2.1 tenemos que $g(X'_1, \dots, X'_n)$ tiene la misma distribución que $g(X_1, \dots, X_n)$ y $g(Y'_1, \dots, Y'_n)$ tiene la misma distribución que $g(Y_1, \dots, Y_n)$ por lo que concluimos $g(X_1, \dots, X_n) \leq_{st} g(Y_1, \dots, Y_n)$. ■

1.3. Ordenamiento por Cociente de Verosimilitud

El orden por cociente de verosimilitud mide de manera puntual, si es más probable que Y tome el valor $t > 0$ a que X lo haga. Si volvemos al ejemplo del inversionista que está interesado en obtener rendimientos a un cierto nivel t , tal inversionista prefiere el activo financiero que con mayor probabilidad alcanza el nivel deseado.

Definición 1.3.1. Sean X y Y variables aleatorias continuas con función de densidad f_X y f_Y respectivamente. Decimos que Y es en cociente de verosimilitud más grande que X si $\frac{f_Y(x)}{f_X(x)}$ es creciente en x en la región en la que $f_X(x) > 0$ o $f_Y(x) > 0$.

Análogamente, si X y Y son variables aleatorias discretas, decimos que Y es en cociente de verosimilitud más grande que X si $\frac{P(Y=x)}{P(X=x)}$ es creciente en x en la región en la que $P(X = x) > 0$

o $P(Y = x) > 0$.

Proposición 1.3.1. *Si Y es en cociente de verosimilitud más grande que X , entonces Y es estocásticamente mayor que X .*

Demostración.

Suponga que X y Y tienen función de densidad (o de masa) f_X y f_Y respectivamente, suponga que $\frac{f_Y(x)}{f_X(x)}$ es creciente en x . Para cada a , queremos probar que

$$\int_{x>a} f_X(x)dx \leq \int_{x>a} f_Y(x)dx.$$

O equivalentemente, si X y Y son variables aleatorias discretas:

$$\sum_{x>a} P(X = x) \leq \sum_{x>a} P(Y = x).$$

Caso 1: $f_X(a) \leq f_Y(a)$.

Si $x > a$, entonces $\frac{f_Y(x)}{f_X(x)} \geq \frac{f_Y(a)}{f_X(a)} \geq 1$. Por lo tanto, $f_Y(x) \geq f_X(x)$ cuando $x \geq a$. Y así,

$$\int_a^\infty f_X(x)dx \leq \int_a^\infty f_Y(x)dx.$$

Caso 2: $f_Y(a) < f_X(a)$.

Si $x \leq a$, entonces $\frac{f_Y(x)}{f_X(x)} \leq \frac{f_Y(a)}{f_X(a)} < 1$, por lo que

$$\int_{x\leq a} f_Y(x)dx \leq \int_{x\leq a} f_X(x)dx,$$

obteniendo el complemento tenemos que

$$\int_{x>a} f_Y(x)dx \geq \int_{x>a} f_X(x)dx.$$



1.4. Problema de Inversión de un Período

Supongamos que tenemos una cantidad inicial w de dinero y decidimos invertir una parte, a dicha cantidad la denotaremos por y y es tal que $0 \leq y \leq w$. Supongamos también que al invertir y , dicha inversión nos genera la cantidad $yX + (1 + r)(w - y)$ al final de un período, donde X es una variable aleatoria no negativa con una distribución conocida y r es una tasa de interés libre de riesgo a la que invertimos el resto del dinero. Si $u(\cdot)$ es una función de utilidad cóncava y creciente, el objetivo es maximizar el valor esperado de la utilidad al final de un período, es decir, si $\beta = 1 + r$, el objetivo es encontrar y tal que se alcance el valor máximo

$$M = \max_{0 \leq y \leq w} E(u(yX + \beta(w - y))).$$

Ahora, supongamos que X es una variable aleatoria continua con función de densidad f_X . Entonces,

$$M = \max_{0 \leq y \leq w} \int_0^{\infty} u(yX + \beta(w - y)) f_X(x) dx.$$

Diferenciando para encontrar los puntos críticos de la expresión anterior

$$\begin{aligned} \frac{d}{dy} \int_0^{\infty} u((x - \beta)y + \beta w) f_X(x) dx &= \int_0^{\infty} u'((x - \beta)y + \beta w) (x - \beta) f_X(x) dx \\ &= \int_0^{\infty} h(y, x) f_X(x) dx, \end{aligned} \tag{1-2}$$

donde

$$h(y, x) = u'((x - \beta)y + \beta w) (x - \beta).$$

Igualando a 0 obtenemos el valor máximo y a invertir, al cual llamaremos y_{f_X} que cumple

$$\int_0^{\infty} h(y_{f_X}, x) f_X(x) dx = 0. \quad (1-3)$$

Para estudiar a profundidad el problema de inversión, analizaremos el comportamiento de la función $h(y, x)$ variando el parámetro β correspondiente a la tasa libre de riesgo.

Lemma 1.4.1. *Sea x fija y $h(y, x)$ decreciente en y . Ocurre,*

$$h(y, x) \leq 0 \text{ si } x \leq \beta,$$

$$h(y, x) \geq 0 \text{ si } x \geq \beta.$$

Demostración.

Utilizaremos en esta demostración la notación $f(y) \uparrow y$ que denota que la función $f(y)$ es creciente en y , y también $f(y) \downarrow y$ que denotará que la función $f(y)$ es decreciente en y .

Caso 1: $x \leq \beta$.

Por hipótesis $h(y, x)$ es decreciente en y , lo cual implica lo siguiente:

$$\begin{aligned} x \leq \beta &\Rightarrow (x - \beta)y + \beta w \downarrow y \\ &\Rightarrow u'((x - \beta)y + \beta w) \uparrow y \text{ (} u \text{ es cóncava entonces } u'(v) \downarrow v) \\ &\Rightarrow h(y, x) = (x - \beta)u'((x - \beta)y + \beta w) \downarrow y. \end{aligned}$$

Además, $h(y, x) \leq 0$ ya que $x - \beta \leq 0$ y $u' \geq 0$ (pues u es creciente).

Caso 2: $x \geq \beta$

$$\begin{aligned} x \geq \beta &\Rightarrow (x - \beta)y + \beta w \uparrow y \\ &\Rightarrow u'((x - \beta)y + \beta w) \downarrow y \text{ (} u \text{ es cóncava entonces } u'(v) \downarrow v) \\ &\Rightarrow h(y, x) = (x - \beta)u'((x - \beta)y + \beta w) \downarrow y. \end{aligned}$$

Finalmente, $h(y, x) \geq 0$ ya que $x - \beta \geq 0$ y $u' \geq 0$. ■

Continuando con el problema de inversión, consideremos los siguientes dos escenarios de inversión con la misma cantidad inicial de dinero w : el primero tiene una variable aleatoria multiplicativa X_1 , el segundo tiene una variable aleatoria multiplicativa X_2 , donde X_1 tiene función de densidad f_{X_1} y X_2 tiene función de densidad f_{X_2} . ¿Bajo qué condiciones de f_{X_1}

y f_{X_2} , la cantidad óptima invertida en el segundo escenario será al menos tan grande como la cantidad óptima invertida en el primer escenario? Es decir, ¿Cuándo $y_{f_{X_1}} \leq y_{f_{X_2}}$? Para contestar dichas preguntas, presentamos los siguientes teoremas.

Teorema 1.4.1. *Si f_{X_1} y f_{X_2} son funciones de densidad de variables aleatorias no negativas y $\frac{f_{X_2}}{f_{X_1}}$ es creciente en x , entonces $y_{f_{X_1}} \leq y_{f_{X_2}}$. Es decir, cuando X_2 es en orden de cociente de verosimilitud mayor que X_1 , entonces la cantidad óptima invertida en el segundo escenario es al menos tan grande como la cantidad óptima invertida en el primer escenario.*

Demostración.

De la ecuación (1-3) sabemos que la cantidad óptima de inversión de X_1 , la cual denotamos por $y_{f_{X_1}}$, satisface:

$$\int_0^{\infty} h(y_{f_{X_1}}, x) f_{X_1}(x) dx = 0.$$

Queremos probar que si $\frac{f_{X_2}(x)}{f_{X_1}(x)}$ es creciente en x , entonces $y_{f_{X_1}} \leq y_{f_{X_2}}$, donde $y_{f_{X_2}}$ es tal que

$$\int_0^{\infty} h(y_{f_{X_2}}, x) f_{X_2}(x) dx = 0.$$

Sabemos que $h(y, x)$ es decreciente en y y por el Lema 1.4.1 se sigue que la desigualdad $y_{f_{X_1}} \leq y_{f_{X_2}}$ es equivalente a la desigualdad $\int_0^{\infty} h(y_{f_{X_2}}, x) f_{X_2}(x) dx \leq \int_0^{\infty} h(y_{f_{X_1}}, x) f_{X_2}(x) dx$. Por lo tanto, es suficiente probar que

$$\int_0^{\infty} h(y_{f_{X_1}}, x) f_{X_2}(x) dx \geq 0.$$

Ahora,

$$\int_0^{\infty} h(y_{f_{X_1}}, x) f_{X_2}(x) dx = \int_0^{\beta} h(y_{f_{X_1}}, x) f_{X_2}(x) dx + \int_{\beta}^{\infty} h(y_{f_{X_1}}, x) f_{X_2}(x) dx.$$

Si $x \leq \beta$, entonces $\frac{f_{X_2}(x)}{f_{X_1}(x)} \leq \frac{f_{X_2}(\beta)}{f_{X_1}(\beta)}$ y así $f_{X_2}(x) \leq \frac{f_{X_2}(\beta)}{f_{X_1}(\beta)} f_{X_1}(x)$. También, si $x \leq \beta$, por el

Lema 1.4.1, $h(y_{f_{X_1}}, x) \leq 0$. Por lo tanto,

$$\int_0^\beta h(y_{f_{X_1}}, x) f_{X_2}(x) dx \geq \frac{f_{X_2}(\beta)}{f_{X_1}(\beta)} \int_0^\beta h(y_{f_{X_1}}, x) f_{X_1}(x) dx. \quad (1-4)$$

Si $x \geq \beta$, entonces $\frac{f_{X_2}(x)}{f_{X_1}(x)} \geq \frac{f_{X_2}(\beta)}{f_{X_1}(\beta)}$ y por el Lema 1.4.1, $h(y_{f_{X_1}}, x) \geq 0$. Por lo tanto,

$$\int_\beta^\infty h(y_{f_{X_1}}, x) f_{X_2}(x) dx \geq \frac{f_{X_2}(\beta)}{f_{X_1}(\beta)} \int_\beta^\infty h(y_{f_{X_1}}, x) f_{X_1}(x) dx. \quad (1-5)$$

Así, por las ecuaciones (1-4) y (1-5) tenemos que

$$\int_0^\infty h(y_{f_{X_1}}, x) f_{X_2}(x) dx \geq \frac{f_{X_2}(\beta)}{f_{X_1}(\beta)} \int_0^\infty h(y_{f_{X_1}}, x) f_{X_1}(x) dx = 0.$$

■

1.5. Segundo Orden Estocástico

Para que Y domine estocásticamente a X necesitamos que $E(h(X)) \leq E(h(Y))$ para toda función creciente h . Estamos interesados en caracterizar el orden cuando cambiamos la condición de funciones h crecientes, por funciones h crecientes y cóncavas.

Siguiendo con el ejemplo del inversionista, supongamos que los rendimientos de las acciones X y Y están sujetos a una función de utilidad cóncava y creciente. Diremos que preferiremos la inversión Y si, en promedio, nos genera una utilidad al menos tan grande como la utilidad promedio que nos genera X . A este orden se le conoce como segundo orden estocástico.

Definición 1.5.1. Decimos que Y domina en segundo orden estocástico a X y se denota como $X \leq_{icv} Y$ si

$$E(h(X)) \leq E(h(Y)),$$

para toda función h cóncava y creciente.

Para un valor a fijo, definamos la siguiente función h_a como sigue

$$h_a(x) = \begin{cases} x & \text{si } x \leq a, \\ a & \text{si } x > a. \end{cases}$$

Notemos que $h_a(x)$ es una línea recta que cuando llega al valor a se vuelve constante, por lo que es una función creciente y cóncava. Podemos expresarla como

$$h_a(X) = a - (a - h_a(X)),$$

aplicando el Lema 1.1.1 para la variable aleatoria definida como $a - h_a(X)$, tenemos;

$$\begin{aligned} E(h_a(x)) &= a - E(a - h_a(X)) \\ &= a - \int_0^\infty P(a - h_a(X) > t) dt \\ &= a - \int_0^\infty P(h_a(X) < t - a) dt \\ &= a - \int_0^\infty P(X < t - a) dt \\ &= a - \int_{-\infty}^a P(X < y) dy. \end{aligned}$$

De lo anterior, se tiene que si Y domina en segundo orden estocástico a X , entonces

$$\int_{-\infty}^a P(X < y) dy \leq \int_{-\infty}^a P(Y < y) dy \quad \forall a \in \mathbb{R}. \quad (1-6)$$

Dicha condición es suficiente para que $X \leq_{icv} Y$, de lo que se tiene el siguiente Teorema.

Teorema 1.5.1. *Y domina estocásticamente a X en segundo orden si y sólo si se cumple la condición de la ecuación (1-6).*

1.5.1. Variables Aleatorias Normales

Para ejemplificar el segundo orden estocástico presentaremos algunos resultados para variables aleatorias con distribución Normal. Estudiaremos este caso en particular pues es común modelar los rendimientos de acciones con esta distribución por su simplicidad analítica. Antes de ello, presentaremos algunos resultados generales.

Proposición 1.5.1. *Si $f(x)$ y $g(x)$ son funciones crecientes en x , entonces, para cualquier variable aleatoria X se cumple*

$$E[f(X)g(X)] \geq E[f(X)]E[g(X)].$$

Si f o g es una función creciente y la otra es una función decreciente, entonces,

$$E[f(X)g(X)] \leq E[f(X)]E[g(X)].$$

Demostración.

Sean X y Y variables aleatorias independientes e idénticamente distribuidas, suponga que $f(x)$ y $g(x)$ son funciones crecientes en x . Entonces $f(X) - f(Y)$ y $g(X) - g(Y)$ tienen el mismo signo. Por lo que

$$(f(X) - f(Y))(g(X) - g(Y)) \geq 0,$$

o equivalentemente

$$f(X)g(X) + f(Y)g(Y) \geq f(X)g(Y) + f(Y)g(X).$$

Tomando esperanzas en ambos lados

$$E[f(X)g(X)] + E[f(Y)g(Y)] \geq E[f(X)g(Y)] + E[f(Y)g(X)].$$

Dado que X y Y son independientes, se sigue que

$$E[f(X)g(X)] + E[f(Y)g(Y)] \geq E[f(X)]E[g(Y)] + E[f(Y)]E[g(X)], \quad (1-7)$$

dado que X y Y tienen la misma distribución, entonces $E[f(X)g(X)] = E[f(Y)g(Y)]$, $E[f(X)] = E[f(Y)]$ y $E[g(X)] = E[g(Y)]$. Por lo tanto

$$2E[f(X)g(X)] \geq 2E[f(X)]E[g(X)],$$

y así

$$E[f(X)g(X)] \geq E[f(X)]E[g(X)].$$

Ahora, si f es decreciente y g es creciente, entonces $-f$ es creciente por lo que

$$E[-f(X)g(X)] \geq E[-f(X)]E[g(X)]$$

Por linealidad de la esperanza y multiplicando ambos lados por -1 tenemos que

$$E[f(X)g(X)] \leq E[f(X)]E[g(X)].$$

■

Lemma 1.5.1. Si $E[X] = 0$ y $c \geq 1$, entonces $cX \leq_{icv} X$.

Demostración.

Sea h una función cóncava y creciente, y sea $c \geq 1$. La expansión en serie de Taylor (con residuo) de $h(cx)$ en x , para algún w entre x y cx , nos dice que

$$\begin{aligned} h(cx) &= h(x) + h'(x)(cx - x) + h''(w)(cx - x)^2/2! \\ &\leq h(x) + h'(x)(cx - x). \end{aligned}$$

La desigualdad se da ya que h es una función cóncava lo que implica que $h''(w) \leq 0$. Dado que dicha desigualdad se da para cualquier valor x , se sigue

$$h(cX) \leq h(X) + (c-1)Xh'(X).$$

Tomando esperanzas en ambos lados

$$E[h(cX)] \leq E[h(X)] + (c-1)E[Xh'(X)].$$

Definamos a $f(x) = x$, entonces

$$E[h(cX)] \leq E[h(X)] + (c-1)E[X]E[h'(X)]$$

La desigualdad anterior se da ya que f es una función creciente y h es cóncava por lo que $h'(x)$ es una función decreciente en x y finalmente por hipótesis $E[X] = 0$, entonces $E[h(cX)] \leq E[h(X)]$. ■

Teorema 1.5.2. Si X_i , $i = 1, 2$ son variables aleatorias normales con medias μ_i y varianzas σ_i respectivas, entonces $\mu_1 \leq \mu_2$, $\sigma_1 \geq \sigma_2$ implica que $X_1 \leq_{icv} X_2$

Demostración.

Suponga que $\mu_1 \leq \mu_2$ y $\sigma_2 \leq \sigma_1$. Sea Z una variable aleatoria normal estándar. Sea $c = \sigma_1/\sigma_2 \geq 1$, por el Lema 1.5.1 se sigue que, $\sigma_2 Z \geq_{icx} c\sigma_2 Z = \sigma_1 Z$. Ahora, sea $h(X)$ una función cóncava y creciente en x . Entonces

$$\begin{aligned} E[h(\mu_1 + \sigma_1 Z)] &\leq E[h(\mu_2 + \sigma_1 Z)] \\ &\leq E[h(\mu_2 + \sigma_2 Z)]. \end{aligned}$$

La primera desigualdad se da ya que $\mu_1 \leq \mu_2$ y h es una función creciente. La segunda desigualdad se da pues $g(x) = h(\mu_2 + x)$, es una función cóncava y creciente en x y $\sigma_1 Z \leq_{icx} \sigma_2 Z$.

El resultado se da, pues $\mu_i + \sigma_i Z$ con $i = 1, 2$, son variables aleatorias con media μ_i y varianza σ_i^2 . ■

1.5.2. Otros resultados de Segundo Orden Estocástico

Supongamos que queremos invertir nuestro dinero en más de un activo financiero y que tenemos $X_1, \dots, X_n, Y_1, \dots, Y_n$ posibles activos a elegir. Decidimos crear dos portafolios, uno con los activos X_1, \dots, X_n y otro con los activos Y_1, \dots, Y_n , es decir, $X = \sum_{i=1}^n X_i$ y $Y = \sum_{i=1}^n Y_i$. Tenemos como antes, una función de utilidad $g(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ cóncava y creciente. Para tomar una decisión, queremos saber qué condiciones deben de cumplir estos activos para que, en promedio, una inversión nos genere mayor utilidad que la otra.

Teorema 1.5.3. *Sea X_1, \dots, X_n y Y_1, \dots, Y_n vectores de variables aleatorias independientes. Si $X_i \leq_{icv} Y_i$ para cada $i = 1, \dots, n$ entonces $\sum_{i=1}^n X_i \leq_{icv} \sum_{i=1}^n Y_i$.*

Demostración.

Sea h una función cóncava y creciente. Por demostrar que $E[h(\sum_{i=1}^n X_i)] \leq_{icv} E[h(\sum_{i=1}^n Y_i)]$. La demostración la haremos por inducción sobre n . Sabemos que cuando $n = 1$ se cumple por hipótesis. Supongamos que se cumple para vectores de tamaño $n - 1$. Ahora consideremos dos vectores de variables aleatorias independientes de tamaño n , X_1, \dots, X_n y Y_1, \dots, Y_n , también supongamos que son independientes entre ellos. Primero, mostraremos que $\sum_{i=1}^n X_i \leq_{icv} \sum_{i=1}^{n-1} Y_i + X_n$. Para verificarlo definiremos, para cada x , la función $h_x(a)$ como $h_x(a) = h(x+a)$. Notemos que h_x es una función cóncava y creciente.

Entonces

$$\begin{aligned}
E \left[h \left(\sum_{i=1}^n X_i \right) \middle| X_n = x \right] &= E \left[h \left(x + \sum_{i=1}^{n-1} X_i \right) \middle| X_n = x \right] \\
&= E \left[h \left(x + \sum_{i=1}^{n-1} X_i \right) \right] \text{ por independencia} \\
&= E \left[h_x \left(\sum_{i=1}^{n-1} X_i \right) \right] \\
&\leq E \left[h_x \left(\sum_{i=1}^{n-1} Y_i \right) \right] \text{ por hipótesis de inducción} \quad (1-8) \\
&= E \left[h \left(x + \sum_{i=1}^{n-1} Y_i \right) \right] \\
&= E \left[h \left(x + \sum_{i=1}^{n-1} Y_i \right) \middle| X_n = x \right] \text{ por independencia} \\
&= E \left[h \left(X_n + \sum_{i=1}^{n-1} Y_i \right) \middle| X_n = x \right].
\end{aligned}$$

Por lo tanto,

$$E \left[h \left(\sum_{i=1}^n X_i \right) \middle| X_n = x \right] \leq E \left[h \left(X_n + \sum_{i=1}^{n-1} Y_i \right) \middle| X_n = x \right],$$

tomando esperanza en ambos lados

$$E \left[h \left(\sum_{i=1}^n X_i \right) \right] \leq E \left[h \left(X_n + \sum_{i=1}^{n-1} Y_i \right) \right].$$

Por lo que $\sum_{i=1}^n X_i \leq_{icv} \sum_{i=1}^{n-1} Y_i + X_n$. Ahora, para completar la prueba, mostraremos que $\sum_{i=1}^{n-1} Y_i + X_n \leq_{icv} \sum_{i=1}^n Y_i$. Para ello, notemos que

$$\begin{aligned}
E \left[h \left(\sum_{i=1}^{n-1} Y_i + X_n \right) \middle| \sum_{i=1}^{n-1} Y_i = y \right] &= E[h_y(X_n)] \\
&\leq E[h_y(Y_n)] = E \left[h \left(\sum_{i=1}^n Y_i \right) \middle| \sum_{i=1}^{n-1} Y_i = y \right]. \quad (1-9)
\end{aligned}$$

La desigualdad se da ya que h_y es una función cóncava, creciente y por la independencia

entre variables aleatorias. Concluimos que

$$E \left[h \left(\sum_{i=1}^{n-1} Y_i + X_n \right) \mid \sum_{i=1}^{n-1} Y_i = y \right] \leq E \left[h \left(\sum_{i=1}^n Y_i \right) \mid \sum_{i=1}^{n-1} Y_i = y \right],$$

y al tomar esperanza

$$E \left[h \left(\sum_{i=1}^{n-1} Y_i + X_n \right) \right] \leq E \left[h \left(\sum_{i=1}^n Y_i \right) \right].$$

Por lo tanto $\sum_{i=1}^{n-1} Y_i + X_n \leq_{icv} \sum_{i=1}^n Y_i$. Y así $\sum_{i=1}^n X_i \leq_{icv} \sum_{i=1}^n Y_i$.

■

Capítulo 2

Medidas de Riesgo

2.1. Introducción

El riesgo es la posibilidad de que ocurra un evento que se traduzca en pérdidas. En el caso financiero estas pérdidas serían sufridas por alguno de los participantes (inversionistas, deudores o entidades financieras) de los mercados financieros. El riesgo está asociado al comportamiento aleatorio de los activos financieros y a los factores que los afectan; a mayor aleatoriedad mayor riesgo.

Las medidas de riesgo sirven para cuantificar las pérdidas generadas ante cambios inusuales en el comportamiento de los activos.

En términos matemáticos, las medidas de riesgo asignan un número real al riesgo dado por una variable aleatoria $X : \Omega \rightarrow \mathbb{R}$. Supondremos desde ahora que la variable aleatoria X describe la pérdida potencial asociada a un activo, sin embargo, permitiremos que esta variable aleatoria tome valores negativos, lo cual representará una ganancia.

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad, sea χ el espacio de todas las variables aleatorias definidas en tal espacio. Una medida de riesgo es entonces, una función de valor real, $\rho : \chi \rightarrow \mathbb{R}$. Las medidas de riesgo pueden cumplir ciertas propiedades:

1. **Invarianza bajo traslación:** para toda $X \in \chi$ y cualquier constante $m \in \mathbb{R}$ ocurre que,
$$\rho(X + m) = \rho(X) + m.$$

Si suponemos que m es el riesgo asociado a un portafolio determinista, entonces, el riesgo de la suma del activo X y m , será el riesgo que produce el activo riesgoso más el riesgo cierto.

2. **Subaditividad:** si $X, Y \in \chi$ entonces $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

Esta propiedad está asociada a la idea de diversificar una cartera. Nos dice que el riesgo asociado a una inversión que incluye a los activos X, Y es a lo más tan riesgosa como invertir por separado dichos activos.

3. **Homogeneidad positiva** si $X \in \chi$ y $\lambda > 0$, entonces $\rho(\lambda X) = \lambda\rho(X)$.

Esta propiedad implica que el riesgo es proporcional al tamaño de la inversión (λ) sobre el activo X .

4. **Monotonidad** si $X \leq Y$ entonces $\rho(X) \leq \rho(Y)$.

Si el activo Y genera mayores rendimientos que el activo X , en consecuencia el riesgo asociado a Y será mayor que el riesgo asociado a X .

5. **Convexidad** $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$ para toda $\lambda \in [0, 1]$.

Esta propiedad se deriva de la propiedad subaditividad y homogeneidad positiva.

6. **Ley de invarianza;** si $P_X = P_Y$, entonces $\rho(X) = \rho(Y)$.

Si dos activos X y Y se rigen por la misma medida de probabilidad (tienen la misma distribución de probabilidad), entonces el riesgo asociado a dichos activos será el mismo para ambos.

Definición 2.1.1. *Decimos que ρ es una medida de riesgo coherente si es monótona, invariante bajo traslaciones, subaditiva y positivamente homogénea.*

2.2. Valor en Riesgo (VaR)

Por sus siglas en inglés *Value-at-Risk* (Valor en Riesgo), es una medida de riesgo que cuantifica la pérdida potencial sobre un activo financiero a un nivel de confianza α en un horizonte

de tiempo dado. El Valor en Riesgo no es una medida de riesgo coherente pues no cumple la propiedad de subaditividad.

Definición 2.2.1. Sea $\alpha \in (0, 1)$ un nivel de confianza fijo. Entonces el VaR de la pérdida de una cartera, X , a un nivel de confianza, α , se define como;

$$VaR_X(\alpha) = q_X(\alpha) = \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\}$$

donde $F_X(\cdot)$ es la función de distribución de la variable aleatoria X .

2.3. CVaR

Por sus siglas en inglés *Conditional Value at Risk* (Valor condicional en Riesgo), también conocida como el déficit esperado, es una medida de riesgo que cuantifica el riesgo que tiene un activo financiero en las colas de la distribución de la variable aleatoria de pérdida.

Definición 2.3.1. Sea X la variable aleatoria de pérdida, el déficit esperado a un nivel de confianza $\alpha \in (0, 1)$ se define como:

$$\rho_\alpha(X) = \frac{1}{\alpha} \int_{1-\alpha}^1 q_X(u) du = \frac{1}{\alpha} \int_{1-\alpha}^1 VaR_X(u) du. \quad (2-1)$$

Notemos que $\rho_\alpha : L^1 \rightarrow \mathbb{R}$ está bien definida, $\rho_1(X) = E[X]$ y $\rho_\alpha(X)$ es decreciente en α .

Más adelante demostraremos que el CVaR es una medida de riesgo coherente. Por ahora mostraremos un resultados que expresa al CVaR por medio del VaR.

Lemma 2.3.1. Para $\alpha \in (0, 1)$ y $X \in L^1$ ocurre

$$\rho_\alpha(X) = VaR_X(1 - \alpha) + \frac{1}{\alpha} E[(X - VaR_X(1 - \alpha))_+].$$

Demostración.

$$\begin{aligned}
\rho_\alpha(X) &= \frac{1}{\alpha} \int_{1-\alpha}^1 F_X^{-1}(u) du = \frac{1}{\alpha} \int_{1-\alpha}^1 \left(F_X^{-1}(1-\alpha) + F_X^{-1}(u) - F_X^{-1}(1-\alpha) \right) du \\
&= F_X^{-1}(1-\alpha) + \frac{1}{\alpha} \int_{1-\alpha}^1 (F_X^{-1}(u) - F_X^{-1}(1-\alpha)) du \\
&= VaR_X(1-\alpha) + \frac{1}{\alpha} E[(F_X^{-1}(U_X) - VaR_X(1-\alpha))_+] \\
&= VaR_X(1-\alpha) + \frac{1}{\alpha} E[(X - VaR_X(1-\alpha))_+].
\end{aligned} \tag{2-2}$$

■

Lemma 2.3.2. Para $\alpha \in (0, 1)$

$$VaR_X(1-\alpha) \in \operatorname{argmin}_{t \in \mathbb{R}} \left\{ t + \frac{1}{\alpha} E[(X-t)_+] \right\}. \tag{2-3}$$

Demostración.

Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que $t \mapsto t + \frac{1}{\alpha} E[(X-t)_+]$, o equivalentemente,

$$f(t) = t + \frac{1}{\alpha} \int_t^\infty (x-t) f_X(x) dx = t + \frac{1}{\alpha} \int_t^\infty (1 - F_X(x)) dx, \quad t \in \mathbb{R},$$

donde la última igualdad se da por integración por partes. Sea $t_0 = VaR_X(1-\alpha)$. Para $t_1 > t_0$ tenemos

$$f(t_1) - f(t_0) = (t_1 - t_0) - \frac{1}{\alpha} \int_{t_0}^{t_1} (1 - F_X(x)) dx.$$

Por la definición de $VaR_X(1-\alpha)$, tenemos que para $x \in (t_0, t_1)$, $F_X(x) \geq 1-\alpha$ y así $1 - F_X(x) \leq \alpha$. Por lo tanto,

$$f(t_1) - f(t_0) \geq (t_1 - t_0) - \frac{1}{\alpha} (\alpha)(t_1 - t_0) = 0$$

Para $t_2 < t_0$, tenemos

$$f(t_2) - f(t_0) = \frac{1}{\alpha} \int_{t_2}^{t_0} (1 - F_X(x)) dx - (t_0 - t_2).$$

De la definición de $VaR_X(1 - \alpha)$, tenemos que para $x \in (t_2, t_0)$, $F_X(x) < 1 - \alpha$ y así $1 - F_X(x) > \alpha$. Por lo tanto,

$$f(t_2) - f(t_0) > (t_2 - t_0) - \frac{1}{\alpha}(\alpha)(t_2 - t_0) = 0$$

Con lo que concluimos que, $t_0 \in \operatorname{argmin}_{t \in \mathbb{R}} f(t)$, es decir, se cumple (2-3). ■

Lemma 2.3.3. Para $\alpha \in (0, 1)$

$$\rho_\alpha(X) = \min_{t \in \mathbb{R}} \left\{ t + \frac{1}{\alpha} E[(X - t)_+] \right\}. \quad (2-4)$$

Demostración.

Por la representación del Lema 2.3.1 tenemos que $\rho_\alpha(X) = VaR_X(1 - \alpha) + \frac{1}{\alpha} E[(X - VaR_X(1 - \alpha))_+]$ y el Lema 2.3.2 se cumple (2-4). ■

Con ayuda de los resultados anteriores procedemos a mostrar que ρ_α es una medida de riesgo consistente. Para ello, debemos mostrar que CVaR es monótona, invariante bajo traslación, subaditiva y homogénea positiva.

1. Monotonía.

Sea $X \leq Y$, por el Teorema 3.1.1 presentado en el Capítulo 3, en la parte de la demostración donde se prueba que 2. implica 3., tenemos que $q_X(\alpha) \leq q_Y(\alpha)$ para toda $\alpha \in (0, 1)$, por lo tanto

$$\int_{1-\alpha}^1 q_X(u) du \leq \int_{1-\alpha}^1 q_Y(u) du,$$

y así $\rho_\alpha(X) \leq \rho_\alpha(Y)$.

2. Invariancia bajo traslaciones.

Sea X y $m \in \mathbb{R}$,

$$q_{X+m}(\alpha) = \inf\{x \in \mathbb{R} | P(X + m \leq x) \geq \alpha\} = \inf\{x \in \mathbb{R} | P(X \leq x - m) \geq \alpha\}.$$

Por otra parte, sabemos que $q_X(\alpha)$ es tal que

$$q_X(\alpha) = \inf\{x \in \mathbb{R} | P(X \leq x) \geq \alpha\},$$

por lo que $q_{X+m}(\alpha) = q_X(\alpha) + m$, y así $\rho_\alpha(X + m) = \rho_\alpha(X) + m$.

3. Homogeneidad positiva.

Sea X y $\lambda > 0$,

$$q_{\lambda X}(\alpha) = \inf\{x \in \mathbb{R} | P(\lambda X \leq x) \geq \alpha\} = \inf\{x \in \mathbb{R} | P(X \leq x/\lambda) \geq \alpha\}$$

Por otra parte, sabemos que $q_X(\alpha)$ es tal que

$$q_X(\alpha) = \inf\{x \in \mathbb{R} | P(X \leq x) \geq \alpha\},$$

por lo que $q_{\lambda X}(\alpha) = \lambda q_X(\alpha)$, y así $\rho_\alpha(\lambda X) = \lambda \rho_\alpha(X)$.

4. Subaditividad.

Sean X, Y variables aleatorias tales que $X \leq Y$. Sea $t_1 = VaR_X(1 - \alpha)$ y $t_2 = VaR_Y(1 - \alpha)$. Por el Lema 2.3.1, por un lado, tenemos que

$$\rho_\alpha(X) + \rho_\alpha(Y) = t_1 + t_2 + \frac{1}{\alpha} E[(X - t_1)_+] + \frac{1}{\alpha} E[(Y - t_2)_+].$$

Notemos que $(x + y)_+ \leq (x)_+ + y_+$ para toda $x, y \in \mathbb{R}$. Por lo tanto, si $t_0 = t_1 + t_2$,

$$\begin{aligned} \rho_\alpha(X) + \rho_\alpha(Y) &= t_0 + \frac{1}{\alpha} E[(X - t_1)_+ + (Y - t_2)_+] \\ &\geq t_0 + \frac{1}{\alpha} E[(X + Y - t_0)_+] \\ &\geq \min_{t \in \mathbb{R}} \left\{ t + \frac{1}{\alpha} E[(X + Y - t)_+] \right\} \text{ por el Lema 2.3.3} \\ &= \rho_\alpha(X + Y). \end{aligned} \tag{2-5}$$

Con lo que demostramos que CVaR es una medida de riesgo consistente.

Capítulo 3

Órdenes estocásticos y Medidas de Riesgo

En este capítulo estamos interesados en la consistencia entre órdenes estocásticos y medidas de riesgo, es decir, saber cuándo ocurre que si $X \preceq Y$, donde \preceq es un orden estocástico, implica $\rho(X) \leq \rho(Y)$.

Buscaremos establecer las condiciones que deben de cumplir las medidas de riesgo para que sean consistentes con los órdenes estocásticos. Trabajaremos con el primero orden estocástico, el orden convexo y el orden convexo creciente.

Para ello, dividiremos este capítulo en tres secciones, en la primera mostraremos algunas implicaciones entre órdenes estocásticos, en la segunda sección mostraremos algunos resultados de medidas de riesgo y finalmente, en la tercera sección, mostraremos algunos resultados de consistencia entre órdenes estocásticos y medidas de riesgo con ayuda de las secciones anteriores.

3.1. Órdenes estocásticos

En el primer capítulo ejemplificamos los órdenes estocásticos con variables aleatorias X y Y que medían los rendimientos de activos, sin embargo, en este capítulo estaremos trabajando con medidas de riesgo por lo que ahora, estas variables aleatorias medirán las pérdidas potenciales.

Por ejemplo, si $X \leq_{st} Y$, preferiremos invertir en el activo asociados a X pues la probabilidad de que la pérdida asociada al activo X sea mayor a un cierto valor t , es menor a la probabilidad de que la pérdida asociada al activo Y sobrepase el mismo nivel.

Definición 3.1.1. Sean X, Y variables aleatorias, definimos las siguientes relaciones de orden

$X \leq_{st} Y$, si $E[f(X)] \leq E[f(Y)]$ para toda función f creciente.

$X \leq_{cx} Y$, si $E[f(X)] \leq E[f(Y)]$ para toda función f convexa.

$X \leq_{icx} Y$, si $E[f(X)] \leq E[f(Y)]$ para toda función f creciente y convexa.

$X \leq_{icv} Y$, si $E[f(X)] \leq E[f(Y)]$ para toda función f creciente y cóncava.

Para toda variable aleatoria X denotamos a $F_X(t) = P(X \leq t), t \in \mathbb{R}$ a la función de distribución y

$$q_X(\alpha) = \inf\{x \in \mathbb{R} \mid P(X \leq x) \geq \alpha\}, \quad 0 < \alpha < 1,$$

a la función cuantil, que es la inversa generalizada de la función de distribución.

Teorema 3.1.1. Para X y Y variables aleatorias con funciones de densidad F_X y F_Y respectivamente, los siguientes enunciados son equivalentes:

1. $X \leq_{st} Y$.
2. Existe un espacio de probabilidad $(\Omega', \mathcal{A}', P')$ y variables aleatorias X' y Y' con funciones de distribución F_X y F_Y en este espacio, tales que $X'(\omega') \leq Y'(\omega')$ para todo $\omega' \in \Omega'$.
3. $F_X(t) \geq F_Y(t)$ para todo t .
4. $q_X(\alpha) \leq q_Y(\alpha)$ para toda $\alpha \in (0, 1)$.

Demostración.

(1) \Rightarrow (2) Se demuestra en la Proposición 1.2.1.

(2) \Rightarrow (3) Sea U una variable aleatoria con distribución uniforme en $(0, 1)$ Definimos a $X' = F_X^{-1}(U)$ y a $Y' = F_Y^{-1}(U)$. Por hipótesis sabemos $X' \leq Y'$, es decir, $F_X^{-1}(u) \leq F_Y^{-1}(u)$ por lo

que

$$F_X(u) \geq F_Y(u).$$

(1) \Leftrightarrow (3) Ocurre que

$$\begin{aligned} P(X > t) &\leq P(Y > t) \quad (X \leq_{st} Y) \\ \Leftrightarrow 1 - F_X(t) &\leq 1 - F_Y(t) \\ \Leftrightarrow -F_X(t) &\leq -F_Y(t) \\ \Leftrightarrow F_X(t) &\geq F_Y(t). \end{aligned} \tag{3-1}$$

(1) \Leftrightarrow (4) Se prueba en la Proposición 1.2.1 considerando $X' = F_X^{-1}(\alpha)$ y $Y' = F_Y^{-1}(\alpha)$, con $\alpha \in (0, 1)$. ■

El orden \leq_{icx} en Actuaría es mejor conocido como orden *stop-loss*. La razón es que el orden \leq_{icx} se mantiene si y sólo si las correspondientes transformaciones *stop-loss* están ordenadas. La transformación *stop-loss* π_X de una variable aleatoria X se define como

$$\pi_X(t) = E[(X - t)_+] = E[\text{máx}\{X - t, 0\}].$$

Integrando por partes;

$$\begin{aligned} E[\text{máx}\{X - t, 0\}] &= \int_t^\infty (x - t)f_X(x)dx \\ &= \int_t^\infty xf_X(x)dx - \int_t^\infty tf_X(x)dx \\ &= - \int_t^\infty x(-f_X(x))dx - t\bar{F}_X(t) \end{aligned}$$

$$\text{sea } u = x \quad dv = -f_X(x)dx$$

$$du = dx \quad v = \bar{F}_X(x) \tag{3-2}$$

$$\begin{aligned} &= -x\bar{F}_X(x)|_t^\infty + \int_t^\infty \bar{F}_X(x)dx - t\bar{F}_X(t) \\ &= x\bar{F}_X(x)|_\infty^t + \int_t^\infty \bar{F}_X(x)dx - t\bar{F}_X(t) \\ &= t\bar{F}_X(t) + \int_t^\infty \bar{F}_X(x)dx - t\bar{F}_X(t) \\ &= \int_t^\infty \bar{F}_X(x)dx. \end{aligned}$$

En teoría de decisiones $X \leq_{icx} Y$ significa que cualquier tomador de decisiones averso al riesgo prefiere el riesgo de X al riesgo de Y .

El orden cóncavo creciente \leq_{icv} corresponde al orden de los rendimientos en lugar de las pérdidas. Si $X \leq_{icv} Y$, donde X y Y son rendimientos, cualquier tomador de decisiones averso al riesgo preferirá Y en lugar de X .

Por otro lado, si definimos a

$$\phi_X(t) = E[(t - X)_+] = E[\text{máx}\{0, t - X\}].$$

Similarmente, podemos integrar por partes y obtener

$$\begin{aligned}
 E[\text{máx}\{0, t - X\}] &= \int_{-\infty}^t (t - x)f_X(x)dx \\
 &= \int_{-\infty}^t tf_X(x)dx - \int_{-\infty}^t xf_X(x)dx \\
 &= tF_X(t) - \int_{-\infty}^t xf_X(x)dx \\
 \text{sea } u &= x \quad dv = f_X(x)dx \\
 du &= dx \quad v = F_X(x) \\
 &= tF_X(t) - \left(xF_X(x) \Big|_{-\infty}^t - \int_{-\infty}^t F_X(x)dx \right) \\
 &= tF_X(t) - tF_X(t) + \int_{-\infty}^t F_X(x)dx \\
 &= \int_{-\infty}^t F_X(x)dx.
 \end{aligned} \tag{3-3}$$

Teorema 3.1.2. *Para variables aleatorias X y Y con función de distribución F_X y F_Y respectivamente las siguientes proposiciones son equivalentes:*

1. $X \leq_{cx} Y$.
2. Existe un espacio de probabilidad $(\Omega', \mathcal{A}', P')$ y variables aleatorias X' y Y' en él con funciones de distribución F_X y F_Y tales que $E[Y'|X'] = X'$.
3. $\pi_X(t) \leq \pi_Y(t)$, $\phi_X(t) \leq \phi_Y(t)$ para toda t , y $E[X] = E[Y]$.

4.

$$\int_{\alpha}^1 q_X(s)ds \leq \int_{\alpha}^1 q_Y(s)ds, \tag{3-4}$$

que es equivalente a la desigualdad

$$\int_0^{\alpha} q_X(s)ds \geq \int_0^{\alpha} q_Y(s)ds, \tag{3-5}$$

para toda $\alpha \in (0, 1)$, y $E[X] = E[Y]$.

Demostración.

(1) \Rightarrow (2) La prueba es compleja y queda fuera de los alcances de esta tesis por lo que no se realizará, se puede consultar en Strassen (1965) [10].

(1) \Rightarrow (3) Como $f(x) = (x - t)_+$ es una función convexa y $X \leq_{cx} Y$ entonces $\pi_X(t) \leq \pi_Y(t)$. Por la igualdad (3-2) tenemos que:

$$\int_a^\infty \bar{F}_X(s) ds \leq \int_a^\infty \bar{F}_Y(s) ds. \quad (3-6)$$

También $f(x) = (t - x)_+$ es una función convexa y $X \leq_{cx} Y$ entonces $\phi_X(t) \leq \phi_Y(t)$. Por la igualdad (3-3) tenemos que:

$$\int_{-\infty}^t F_X(x) dx \leq \int_{-\infty}^t F_Y(x) dx. \quad (3-7)$$

Por otro lado, $f(x) = x$ y $-f(x)$ son funciones convexas por lo que $E[X] \leq E[Y]$ y $-E[X] \leq -E[Y]$, es decir, $E[X] \geq E[Y]$ y así $E[X] = E[Y]$.

(2) \Rightarrow (1) Sea ϕ una función convexa, por la desigualdad de Jensen tenemos que:

$$E[\phi(X)] = E[\phi(\hat{X})] = E[\phi(E[\hat{Y}|\hat{X}])] \leq E[E[\phi(\hat{Y})|\hat{X}]] = E[\phi(\hat{Y})] = E[\phi(Y)],$$

por lo tanto $X \leq_{cx} Y$.

(1) \Leftrightarrow (4)

(1) \Rightarrow (4) La prueba se realizará solo para funciones de distribución continuas. Supongamos que F_X y F_Y no son idénticas, como $E[X] = E[Y]$, se sigue que F_X y F_Y deben cruzarse entre sí al menos una vez. Si se cumple la ecuación (3-6), entonces, si hay una primera vez que F_X cruza F_Y , deberá de hacerlo por debajo. De igual forma, si hay una última vez que F_X cruza a F_Y también lo hará por debajo.

Sean (y_0, α_0) , (y_1, α_1) y (y_2, α_2) los puntos consecutivos de corte, como se muestra en la figura 3.1. Notemos que (y_0, α_0) puede ser $(-\infty, 0)$, y que (y_2, α_2) puede ser $(\infty, 1)$. Notemos

que bajo el supuesto de continuidad tenemos que

$$\alpha_i = F_X(y_i) = F_Y(y_i), \quad i = 0, 1, 2.$$

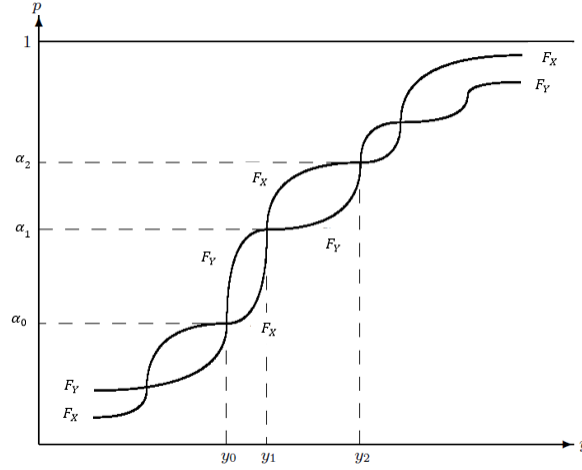


Figura 3-1: Representación típica de segmentos de F_X y F_Y cuando $X \leq_{cx} Y$.

Por hipótesis $X \leq_{cx} Y$. Entonces

$$\int_{y_2}^{\infty} \bar{F}_X(x) dx \leq \int_{y_2}^{\infty} \bar{F}_Y(x) dx,$$

y así

$$\begin{aligned} \int_{\alpha_2}^1 F_X^{-1}(u) du &= y_2(1 - \alpha_2) + \int_{y_2}^{\infty} \bar{F}_X(x) dx \\ &\leq y_2(1 - \alpha_2) + \int_{y_2}^{\infty} \bar{F}_Y(x) dx \\ &= \int_{\alpha_2}^1 F_Y^{-1}(u) du. \end{aligned} \tag{3-8}$$

Ahora, para $u \in [\alpha_1, \alpha_2]$ tenemos que $F_X^{-1}(u) - F_Y^{-1}(u) \leq 0$. Así, $\int_{\alpha}^1 (F_X^{-1}(u) - F_Y^{-1}(u)) du$ es decreciente en $\alpha \in [\alpha_1, \alpha_2]$. Por lo tanto, por la ecuación (3-8) tenemos que:

$$\int_{\alpha}^1 F_X^{-1}(u)du \leq \int_{\alpha}^1 F_Y^{-1}(u)du \text{ para } \alpha \in [\alpha_1, \alpha_2]. \quad (3-9)$$

Dado que $X \leq_{cx} Y$ tenemos también que

$$\int_{-\infty}^{y_0} F_X(x)dx \leq \int_{-\infty}^{y_0} F_Y(x)dx,$$

y así

$$\begin{aligned} \int_0^{\alpha_0} F_X^{-1}(u)du &= y_0\alpha_0 - \int_{-\infty}^{y_0} F_X(x)dx \\ &\geq y_0\alpha_0 - \int_{-\infty}^{y_0} F_Y(x)dx \\ &= \int_0^{\alpha_0} F_Y^{-1}(u)du. \end{aligned} \quad (3-10)$$

Ahora, para $u \in [\alpha_0, \alpha_1]$ tenemos que $F_X^{-1}(u) - F_Y^{-1}(u) \geq 0$. Entonces $\int_0^{\alpha} (F_X^{-1}(u) - F_Y^{-1}(u))du$ es creciente en $\alpha \in [\alpha_0, \alpha_1]$. Entonces, por (3-10) tenemos que

$$\int_0^{\alpha_0} F_X^{-1}(u)du \geq \int_0^{\alpha_0} F_Y^{-1}(u)du \text{ para } \alpha \in [\alpha_0, \alpha_1]. \quad (3-11)$$

(4) \Rightarrow (1) Ahora, supongamos que se cumple la ecuación (3-4), entonces

$$\int_{\alpha_2}^1 F_X^{-1}(u)du \leq \int_{\alpha_2}^1 F_Y^{-1}(u)du,$$

y así

$$\begin{aligned}
\int_{y_2}^{\infty} \bar{F}_X(x) dx &= \int_{\alpha_2}^1 F_X^{-1}(u) du - y_2(1 - \alpha_2) \\
&\leq \int_{\alpha_2}^1 F_Y^{-1}(u) du - y_2(1 - \alpha_2) \\
&= \int_{y_2}^{\infty} \bar{F}_Y(x) dx.
\end{aligned} \tag{3-12}$$

Ahora, para $x \in [y_1, y_2]$ tenemos que $\bar{F}_X(x) - \bar{F}_Y(x) \leq 0$, entonces $\int_y^{\infty} (\bar{F}_X(x) - \bar{F}_Y(x)) dx$ es creciente en $y \in [y_1, y_2]$. Por lo tanto, por la ecuación (3-12) tenemos que

$$\int_y^{\infty} \bar{F}_X(x) dx \leq \int_y^{\infty} \bar{F}_Y(x) dx \text{ para } y \in [y_1, y_2]. \tag{3-13}$$

Ahora, supongamos que se cumple la ecuación (3-5), tenemos que

$$\int_0^{\alpha_0} F_X^{-1}(u) du \geq \int_0^{\alpha_0} F_Y^{-1}(u) du. \tag{3-14}$$

Así

$$\begin{aligned}
\int_{-\infty}^{y_0} F_X(x) dx &= y_0 \alpha_0 - \int_0^{\alpha_0} F_X^{-1}(u) du \\
&\leq y_0 \alpha_0 - \int_0^{\alpha_0} F_Y^{-1}(u) du \\
&= \int_{-\infty}^{y_0} F_Y(x) dx.
\end{aligned} \tag{3-15}$$

Ahora, para $x \in [y_0, y_1]$ tenemos que $F_X(x) - F_Y(x) \leq 0$. Así, $\int_{-\infty}^y (F_X(x) - F_Y(x)) dx$ es decreciente en $y \in [y_0, y_1]$. Por lo tanto, por la ecuación (3-15) tenemos que

$$\int_{-\infty}^y F_X(x) dx \leq \int_{-\infty}^y F_Y(x) dx \text{ para } y \in [y_0, y_1]. \tag{3-16}$$

Por las ecuaciones (3-13) y (3-16) tenemos que $X \leq_{cx} Y$. ■

Teorema 3.1.3. Sean X y Y variables aleatorias en el espacio de probabilidad (Ω, \mathcal{A}, P) tales

que $X \leq_{icx} Y$. Entonces, existe una variable aleatoria Z en el mismo espacio de probabilidad que cumple:

$$X \leq_{st} Z \leq_{cx} Y.$$

Demostración.

Sean variables aleatorias X', Y' un espacio de probabilidad $(\Omega', \mathcal{A}', P')$ tales que

$$X =_{st} X',$$

$$Y =_{st} Y',$$

$E[Y'|X' = x]$ es creciente en x y $E[Y'|X'] = X'$. Definamos a $Z' = E[Y'|X']$, lo cual implica $E[Y'|Z'] = E[Y'|E[Y'|X']] = E[Y'|X'] = Z'$, por el Teorema 3.1.3 tenemos que $Z' \leq_{cx} Y'$.

Por construcción tenemos que $X' = E[X'|Y'] = Z'$, en particular $X' \leq Z'$, por el Teorema 1.1.1 tenemos que $X' \leq_{st} Z'$. Tomando Z con la misma distribución que Z' , obtenemos el resultado. ■

Posteriormente, necesitaremos la relación entre el orden convexo y el concepto de fusión. Utilizaremos la notación μ_P para la medida de probabilidad P y δ_c para la medida de un punto en c .

Definición 3.1.2. Una medida de probabilidad P es una fusión de una medida de probabilidad Q , si existen medidas de probabilidad P_1, P_2 y una constante $0 \leq \alpha_1$ tal que

$$P = \alpha P_1 + (1 - \alpha) \delta_{\mu_{P_2}} \quad \text{y} \quad Q = \alpha P_1 + (1 - \alpha) P_2 \quad (3-17)$$

En palabras, una medida P es una fusión de Q , si P puede verse como la medida Q tras haber colapsado una parte de la masa de probabilidad $(1 - \alpha)$.

Dada nuestra definición, es claro que $P \leq_{cx} Q$, si P es una fusión de Q .

Teorema 3.1.4. Para medidas de probabilidad P y Q con la misma esperanza finita, son equivalentes

1. $P \leq_{cx} Q$;

2. Existe una sucesión de medidas de probabilidad (Q_n) tales que $Q_0 = Q$, Q_{n+1} es una fusión de Q_n , y Q_n converge débilmente a P ;

Si se cumple alguna de las equivalencias, podemos construir variables aleatorias X, X_0, X_1, X_2, \dots con $X \sim P$ y $X_n \sim Q_n$ tales que $E|X - X_n| \rightarrow 0$.

Omitiremos la prueba de este Teorema, pero se puede consultar en Elton y Hill (1992) [3] y en Strassen (1965) [10].

3.2. Medidas de Riesgo Coherentes

En esta sección estudiaremos las condiciones para que las medidas de riesgo sean consistentes con los órdenes estocásticos. Trabajaremos con los órdenes; \leq_{st} , \leq_{icx} , \leq_{cx} pues el orden \leq_{icv} corresponde al orden de rendimientos en lugar de pérdidas.

Continuando con la definición 2.1.1, una medida de riesgo ρ es coherente si es monótona, invariante bajo traslación, subaditiva y es homogénea positiva.

En Artzner (1999) [1] se muestra la siguiente representación para medidas de riesgo coherentes.

Teorema 3.2.1. *Sea $\rho : \chi \rightarrow \mathbb{R}$ una medida de riesgo coherente, dicha medida de riesgo tiene representación*

$$\rho(X) = \sup_{Q \in \mathcal{Q}} E_Q(X), \quad (3-18)$$

donde \mathcal{Q} es un espacio de medidas de probabilidad.

El teorema anterior nos dice que $\rho(X)$ es la peor pérdida esperada bajo Q , donde Q varía en algún espacio de medidas de probabilidad.

De ahora en adelante supondremos que (Ω, \mathcal{A}, P) es un espacio de probabilidad con P no atómica. Esto implica que para cualquier distribución Q en \mathbb{R} existe una variable aleatoria $X : \Omega \rightarrow \mathbb{R}$ con $Q = P_X$.

Usualmente, en la literatura de medidas de riesgo se asume que el espacio en el que se trabaja es $L^\infty(\Omega, \mathcal{A}, P)$ el espacio de las variables aleatorias acotadas. Sin embargo, trabajaremos en el espacio $\chi = L^1(\Omega, \mathcal{A}, P)$, es decir, toda variable tiene media finita bajo P , permitiendo que las medidas de riesgo tomen el valor $+\infty$. Así, las medidas de riesgo son un mapeo $\rho : \chi \rightarrow \mathbb{R} \cup \{\infty\}$.

Denotaremos por $X_n \rightarrow_d X$ la convergencia débil, $X_n \rightarrow_p X$ la convergencia en probabilidad y $X_n \rightarrow_{L^1} X$ la convergencia en L^1 , es decir, $E|X - X_n| \rightarrow 0$.

En la literatura de medidas de riesgo, donde solo se consideran variables aleatorias en L^∞ , usualmente se asume el supuesto de semi-continuidad, conocido como la *propiedad de Fatou*. Si (X_n) es una sucesión de variables aleatorias uniformemente acotadas con $X_n \rightarrow_p X$, entonces

$$\rho(X) \leq \liminf_{n \rightarrow \infty} \rho(X_n). \quad (3-19)$$

Como trabajaremos en el espacio L^1 , esta condición se modifica de la siguiente manera.

Si X, X_1, X_2, \dots son variables aleatorias integrables con $X_n \rightarrow_{L^1} X$, entonces

$$\rho(X) \leq \liminf_{n \rightarrow \infty} \rho(X_n). \quad (3-20)$$

Recordemos que $X_n \rightarrow_{L^1} X$ se cumple si y sólo si, se cumplen las siguientes dos condiciones $X_n \rightarrow_p X$ y (X_n) es uniformemente integrable.

Ahora, probaremos bajo qué supuestos las medidas de riesgo en espacios de probabilidad no atómicos son consistentes con el orden estocástico usual y el orden convexo.

Teorema 3.2.2. *Suponga que (Ω, \mathcal{A}, P) es un espacio de probabilidad estándar con P no atómico, sea ρ una medida de riesgo monótona e invariante bajo ley. Entonces $X \leq_{st} Y$ implica que $\rho(X) \leq \rho(Y)$.*

Demostración.

Como todos los espacios de probabilidad no atómicos son isomorfos de Borel, podemos elegir sin pérdida de generalidad, $\Omega = (0, 1)$ y P la medida de Lebesgue. Si X y Y son variables aleatorias con $X \leq_{st} Y$, por el Teorema 3.1.1, existen variables aleatorias X' y Y' en (Ω, \mathcal{A}, P)

con la misma distribución de X y Y respectivamente, que satisfacen $X' \leq Y'$. Por lo tanto, la monotonicidad e invarianza de ley de la medida de riesgo implica

$$\rho(X) = \rho(X') \leq \rho(Y') = \rho(Y). \quad (3-21)$$

■

Teorema 3.2.3. *Suponga que (Ω, \mathcal{A}, P) es un espacio de probabilidad estándar con P no atómico, sea ρ una medida de riesgo convexa, invariante bajo ley y con la propiedad de Fatou. Entonces $X \leq_{cx} Y$ implica $\rho(X) \leq \rho(Y)$.*

Demostración.

Primero, definiremos $X_1 := X$, $X_2 := Y$ y asumimos que P_{X_1} es una fusión de P_{X_2} . Esto significa que existen medidas de probabilidad P_1 y P_2 tal que

$$P_{X_1} = \alpha P_1 + (1 - \alpha)\delta_{\mu_{P_2}} \text{ y } P_{X_2} = \alpha P_1 + (1 - \alpha)P_2. \quad (3-22)$$

Sean W, I, Y_1, Y_2, \dots variables aleatorias con $P(I = 1) = \alpha = 1 - P(I = 0)$, W tiene distribución P_1 ($W \sim P_1$) y Y_1, Y_2, \dots es una sucesión de variables aleatorias independientes con distribución P_2 . $Y_i \sim P_2, \forall i = 1, 2, \dots$. Definimos $V_n = IW + (1 - I)Y_n$, $n = 1, 2, \dots$. Notemos que V_n tiene la misma distribución que X_2 y que $V_0 = IW + (1 - I)E[Y_1]$ tiene la misma distribución que X_1 . Entonces

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n V_i &= \frac{1}{n} \sum_{i=1}^n (IW + (1 - I)Y_i) \\ &= IW + (1 - I) \frac{1}{n} \sum_{i=1}^n Y_i. \end{aligned} \quad (3-23)$$

Ahora,

$$\begin{aligned}
E \left| V_0 - \frac{1}{n} \sum_{i=1}^n V_i \right| &= E \left| IW + (1-I)E[Y_1] - \left(IW + (1-I)\frac{1}{n} \sum_{i=1}^n Y_i \right) \right| \\
&= E \left| (1-I) \left(E[Y_1] - \frac{1}{n} \sum_{i=1}^n Y_i \right) \right| \\
&\leq E \left| E[Y_1] - \frac{1}{n} \sum_{i=1}^n Y_i \right| \text{ pues } I \in \{0, 1\} \\
&= E \left| \frac{1}{n} \sum_{i=1}^n (Y_i - E[Y_i]) \right| \xrightarrow{L.G.N} 0.
\end{aligned} \tag{3-24}$$

Así, por la invarianza en ley, convexidad y la propiedad de Fatou de la medida de riesgo tenemos

$$\rho(X_1) = \rho(V_0) \leq \liminf_{n \rightarrow \infty} \rho \left(\frac{1}{n} \sum_{i=1}^n V_i \right) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \rho(V_i) = \rho(V_1) = \rho(X_2). \tag{3-25}$$

Ahora, consideremos variables aleatorias arbitrarias X_1 y X_2 con $X_1 \leq_{cx} X_2$. De acuerdo al Teorema 3.1.4 existe una sucesión de medidas de probabilidad (P_n) tal que $P_0 = P_{X_2}$, P_{n+1} es una fusión de P_n , y $P_n \rightarrow_d P_{X_1}$, y una sucesión $\{Z_n\}$ de variables aleatorias con distribución $\{P_n\}$ tales que $Z_n \rightarrow_{L^1} Z$ para alguna variable aleatoria Z con la misma distribución que X_1 . De la primera parte de la prueba, se sigue que $\rho(Z_{n+1}) \leq \rho(Z_n)$, y aplicando la propiedad de Fatou

$$\rho(X_1) = \rho(Z) \leq \liminf_{n \rightarrow \infty} \rho(Z_n) \leq \rho(Z_0) = \rho(X_2). \tag{3-26}$$

■

De manera análoga a como obtuvimos el Teorema 3.1.3, también podemos combinar el Teorema 3.2.2 y 3.2.3 para obtener el siguiente resultado.

Teorema 3.2.4. *Suponga que (Ω, \mathcal{A}, P) es un espacio de probabilidad estándar con P no atómico, sea ρ una medida de riesgo monótona, convexa e invariante bajo ley con la propiedad*

de Fatou. Entonces $X \leq_{icx} Y$ implica $\rho(X) \leq \rho(Y)$.

Demostración.

Por el resultado del Teorema 3.1.3, existe Z en (Ω, \mathcal{A}, P) tal que $X \leq_{st} Z \leq_{cx} Y$. Como ρ es una medida de riesgo monótona e invariante bajo ley, por el Teorema ??

$$\rho(X) \leq \rho(Z).$$

También, ρ es una medida de riesgo convexa, invariante bajo ley y con la propiedad de Fatou, por el Teorema 3.2.2

$$\rho(Z) \leq \rho(Y).$$

Por lo que se sigue que $\rho(X) \leq \rho(Y)$. ■

Recordemos la definición de la medida de riesgo CVaR (*Conditional Value at Risk*), que es definida para un valor fijo $\alpha \in (0, 1]$ como

$$\rho_\alpha(X) = \frac{1}{\alpha} \int_{1-\alpha}^1 q_X(u) du.$$

Y para el caso $\alpha = 0$ por

$$\rho_0(x) = \text{ess.sup}(X) = \inf\{a \geq 0 | P(\{\omega : |X(\omega)| > a\}) = 0\}.$$

Notemos que $\rho_\alpha : L^1 \rightarrow \mathbb{R} \cup \{\infty\}$, $\alpha \in [0, 1]$ está bien definida, $\rho_1(X) = E[X]$ y $\rho_\alpha(X)$ es decreciente en α .

Teorema 3.2.5. *Sea (Ω, \mathcal{A}, P) un espacio de probabilidad estándar con P no atómica. Una medida de riesgo $\rho : L^1(\Omega, \mathcal{A}, P) \rightarrow \mathbb{R}$ es invariante bajo traslaciones y coherente con la propiedad de Fatou si y sólo si la siguiente representación se mantiene:*

$$\rho(X) = \sup \left\{ \int_0^1 \rho_\alpha(X) \mu(d\alpha); \mu \in P_0 \right\}, \quad (3-27)$$

donde P_0 es un conjunto de medidas de probabilidad compacto y convexo en $[0, 1]$.

Demostración.

\Leftrightarrow) En la sección de Medidas de Riesgo, mostramos que el CVaR es una medida de riesgo coherente. Ahora probaremos que $\rho(X)$ es una medida de riesgo coherente.

1. Monotonía

Sean X, Y variables aleatorias, dado que el CVaR cumple la monotonía, tenemos que $\rho_\alpha(X) \leq \rho_\alpha(Y)$. Sea $\mu \in P_0$,

$$\int_0^1 \rho_\alpha(X) \mu(d\alpha) \leq \int_0^1 \rho_\alpha(Y) \mu(d\alpha),$$

tomando supremo de ambos lados, $\rho(X) \leq \rho(Y)$.

2. Invarianza bajo traslación

Sea X y $m \in \mathbb{R}$, dado que el CVaR es una medida de riesgo coherente tenemos que; $\rho_\alpha(X + m) = \rho_\alpha(X) + m$. Sea $\mu \in P_0$,

$$\int_0^1 \rho_\alpha(X + m) \mu(d\alpha) = \int_0^1 \rho_\alpha(X) \mu(d\alpha) + m,$$

tomando supremo de ambos lados, $\rho(X + m) = \rho(X) + m$.

3. Homogeneidad positiva

Sea X y $\lambda > 0$, dado que el CVaR es una medida de riesgo coherente, se tiene que $\rho_\alpha(\lambda X) = \lambda \rho_\alpha(X)$. Sea $\mu \in P_0$,

$$\int_0^1 \rho_\alpha(\lambda X) \mu(d\alpha) = \lambda \int_0^1 \rho_\alpha(X) \mu(d\alpha),$$

tomando supremo de ambos lados, $\rho(\lambda X) = \lambda \rho(X)$.

4. Subaditividad

Sean X, Y variables aleatorias, dado que el CVaR es una medida de riesgo coherente,

$\rho_\alpha(X + Y) \leq \rho_\alpha(X) + \rho_\alpha(Y)$. Sea $\mu \in P_0$,

$$\int_0^1 \rho_\alpha(X + Y)\mu(d\alpha) \leq \int_0^1 \rho_\alpha(X)\mu(d\alpha) + \int_0^1 \rho_\alpha(Y)\mu(d\alpha),$$

tomando supremo de ambos lados

$$\begin{aligned} \rho(X + Y) &\leq \sup \left\{ \int_0^1 \rho_\alpha(X)\mu(d\alpha) + \int_0^1 \rho_\alpha(Y)\mu(d\alpha); \mu \in P_0 \right\} \\ &\leq \sup \left\{ \int_0^1 \rho_\alpha(X)\mu(d\alpha); \mu \in P_0 \right\} + \sup \left\{ \int_0^1 \rho_\alpha(Y)\mu(d\alpha); \mu \in P_0 \right\} \quad (3-28) \\ &= \rho(X) + \rho(Y). \end{aligned}$$

5. Invarianza en ley

Sean X y Y variables aleatorias, tal que $P_X = P_Y$, entonces $q_X(\alpha) = q_Y(\alpha)$ por lo que $\rho_\alpha(X) = \rho_\alpha(Y)$ y así $\rho(X) = \rho(Y)$.

Por lo que se sigue que ρ una medida de riesgo invariante en ley y coherente, basta demostrar que ρ satisface la propiedad de Fatou, es decir, para $X, X_1, X_2, \dots \in L^1$ con $X_n \rightarrow_P X$ y $\{X_n\}$ uniformemente integrable, se sigue que

$$\rho(X) \leq \liminf_{n \rightarrow \infty} \rho(X_n).$$

Primero, notemos que $X_n \rightarrow_P X$ implica que $q_{X_n}(\alpha) \rightarrow q_X(\alpha)$ para casi toda $\alpha \in (0, 1)$ y $q_X(U)$ tiene la misma función de distribución F_X , donde U es una variable aleatoria con distribución uniforme en $(0, 1)$, es decir, $q_{X_n}(U) \stackrel{d}{=} X_n$.

Se sigue, por la proposición A.3.1, que son equivalentes

1. $E[|q_{X_n}|] \leq C \forall n$,
2. $\forall \epsilon > 0$ existe $\delta > 0$ tal que $\forall \mathcal{A} \in \mathcal{F}$

$$P(\mathcal{A}) \leq \delta \rightarrow \sup_n E[|q_{X_n}(U)|1_{\mathcal{A}}] \leq \epsilon. \quad (3-29)$$

De esta última expresión y de que $q_{X_n}(U)$ tiene la misma ley que X_n , se sigue que $\{q_{X_n}(U)\}$ es uniformemente integrable.

Entonces,

$$\begin{aligned}\rho_\alpha(X) &= \frac{1}{\alpha} \int_{1-\alpha}^1 q_X(u) du = \frac{1}{\alpha} \int_{1-\alpha}^1 \lim_{n \rightarrow \infty} q_{X_n}(u) du \\ &= \lim_{n \rightarrow \infty} \rho_\alpha(X_n).\end{aligned}\tag{3-30}$$

Notemos que $X_n \rightarrow_P X$ implica $\liminf_{n \rightarrow \infty} \rho_0(X_n) \geq \rho_0(X)$.

$$\begin{aligned}P(|X| > a + \epsilon) &= P(|X| > a + \epsilon, |X_n| > a) + P(|X| > a + \epsilon, |X_n| \leq a) \\ &\leq P(|X_n| > a) + P(|X - X_n| \geq \epsilon).\end{aligned}\tag{3-31}$$

Dado que $X_n \rightarrow_P X$, existe $N = N(\epsilon)$ tal que $\forall \epsilon > 0$

$$P(|X - X_k| > \epsilon) < \epsilon \quad \forall k > N.$$

Definimos a los siguientes conjuntos $A(a + \epsilon) = \{a > 0 : P(|X| > a + \epsilon) = 0\}$, $B_k(a) = \{a > 0 : P(|X_k| > a) = 0\}$ y demostraremos que $B_k(a) \subseteq A(a + \epsilon)$, $\forall k > N$.

Sea $a \in B_k$, es decir, $P(|X_k| > a) = 0$. Por la ecuación (3-31) tenemos que

$$P(|X| > a + \epsilon) \leq P(|X - X_n| > \epsilon),$$

dado que $k \geq N$,

$$P(|X| > a + \epsilon) < \epsilon \quad \forall \epsilon > 0,$$

así $P(|X| > a + \epsilon) = 0$, con lo que concluimos que $B_k(a) \subseteq A(a + \epsilon)$ y en consecuencia $\inf A(a + \epsilon) \leq \inf B_n(a)$.

Sea $A(a) = \{a > 0 : P(|X| > a) = 0\}$ y $\{A(a + \epsilon)\}$ para $\epsilon > 0$, una sucesión. Notemos que $\{A(a + \epsilon)\} \downarrow A(a)$. Por teorema de continuidad de la probabilidad, se sigue que $A(a) = \bigcap_{\epsilon > 0} A(a + \epsilon)$.

ϵ), por lo que $\inf A(a) \leq \inf B_k(a)$, es decir, $\rho_0(X) \leq \rho_0(X_k)$ y así; $\rho_0(X) \leq \liminf_{k \rightarrow \infty} \rho_0(X_k)$.

Sea $\mu \in P_0$ arbitraria. Recordemos que $\alpha \mapsto \rho_\alpha(X)$ es decreciente y $\rho_1 = E[X]$. Suponga que $N(\epsilon)$ es tal que $|E[X_n] - E[X]| \leq \epsilon$ para toda $n \geq N(\epsilon)$. Así, para toda $n \geq N(\epsilon)$ y para toda $\alpha \in [0, 1]$ tenemos que

$$E[X_n] \leq \rho_\alpha(X_n).$$

Por otro lado,

$$-\epsilon \leq E[X_n] - E[X] \leq \epsilon$$

$$E[X] - \epsilon \leq E[X_n] \leq \rho_\alpha(X_n),$$

y así $\rho_\alpha(X_n) \geq E[X] - \epsilon$. Utilizando el Lema de Fatou, tenemos que

$$\liminf_{n \rightarrow \infty} \int_0^1 \rho_\alpha(X_n) \mu(d\alpha) \geq \int_0^1 \rho_\alpha(X) \mu(d\alpha). \quad (3-32)$$

Tomando límite inferior en ambos lados de la desigualdad tenemos

$$\liminf_{n \rightarrow \infty} \rho(X_n) \geq \liminf_{n \rightarrow \infty} \int_0^1 \rho_\alpha(X_n) \mu'(d\alpha) \geq \int_0^1 \rho_\alpha(X) \mu'(d\alpha). \quad (3-33)$$

Dado que la desigualdad se cumple para toda $\mu' \in P_0$, tomando supremo sobre toda $\mu' \in P_0$ obtenemos:

$$\liminf_{n \rightarrow \infty} \rho(X_n) \geq \rho(X).$$

\Rightarrow) Ahora, debemos mostrar que para cualquier medida de riesgo ρ' coherente e invariante en ley con la propiedad de Fatou tiene la forma (3-27). Consideremos la restricción de ρ' en L^∞ . Por el Teorema 4 en Kusuoka (2001) [7], hay una representación como en (3-27), la cual denotaremos por ρ . Tal representación es profunda. Nos contentaremos con mostrar que podemos extenderla a L^1 . Para ello mostraremos que $\rho = \rho'$ se mantiene para variables no acotadas en L^1 . Supongamos primero que $X \in L^1$ y X es acotada por debajo. Consideremos la sucesión $X_n := X \wedge n$. Notemos que X_n es acotada y $X_n \leq X_{n+1}$ para $n \geq 1$, por el Teorema A.2.1 de convergencia monótona $E[X_n] \uparrow E[X]$.

Veamos que $\{X_n\}$ es uniformemente integrable,

$$|X_n| \leq |X| \quad \forall n,$$

por lo tanto $E[|X_n|] \leq E[|X|]$ y así $\sup_n E[|X_n|] \leq E[|X|]$, por lo que $\{X_n\}$ es U.I. Por la monotonicidad y por la propiedad de Fatou de ρ y ρ' y del hecho de que $\rho = \rho'$ en L^∞ , tenemos que

$$\begin{aligned} \rho(X) &\leq \liminf_{n \rightarrow \infty} \rho(X_n) = \liminf_{n \rightarrow \infty} \rho'(X_n) \\ &\leq \rho'(X) \leq \liminf_{n \rightarrow \infty} \rho'(X_n) = \liminf_{n \rightarrow \infty} \rho(X_n) \leq \rho(X). \end{aligned} \tag{3-34}$$

Así, $\rho(X) = \rho'(X)$ para toda $X \in L^1$ acotada por debajo. Ahora, sea $X \in L^1$ arbitraria y consideremos $X_n := X \vee (-n)$. Tenemos que $X_n \downarrow X$, X_n es acotada por debajo $X_{n+1} \leq X_n$ para toda $n \geq 1$ por el teorema A.2.1 de convergencia monótona $E[X_n] \downarrow E[X]$.

Veamos que $\{X_n\}$ es uniformemente integrable,

$$|X_n| = \begin{cases} n & \text{si } X \leq -n \\ |X| & \text{si } X > -n. \end{cases}$$

Notemos que $|X_n| \leq |X_{n+1}|$ para $n \geq 1$, más aún $|X_n| \leq |X|$, lo cual implica que $E[|X_n|] \leq E[|X|]$ y así $\{X_n\}$ es uniformemente integrable. Ahora veamos que esta sucesión $\{X_n\}$ cumple $\lim_{n \rightarrow \infty} \rho(X_n) = \rho(X)$. Sea n suficientemente grande y fijo. Tenemos que $q_{X_n}(u) = q_X(u)$ para toda $u \geq \alpha^*$ y $\rho_0(X_n) = \rho_0(X)$. Así, $\rho_\alpha(X_n) = \rho_\alpha(X)$ para toda $\alpha \leq 1 - \alpha^*$. Ahora, para $\alpha > 1 - \alpha^*$ tenemos que

$$|\rho_\alpha(X_n) - \rho_\alpha(X)| = \frac{1}{\alpha} \left| \int_{1-\alpha}^1 q_{X_n}(u) - q_X(u) du \right| \leq \frac{1}{1 - \alpha^*} |E[X] - E[X_n]|.$$

La última expresión se puede hacer lo suficientemente pequeña tomando un n suficientemente grande. Por lo tanto, $\rho_\alpha \rightarrow \rho_\alpha(X)$ uniformemente para $\alpha \in [0, 1]$, lo cual implica que

para $\epsilon > 0$ existe $N(\epsilon)$ tal que para $n \geq N(\epsilon)$ y $\mu \in P_0$:

$$\int_0^1 \rho_\alpha(X_n) \mu(d\alpha) \leq \int_0^1 \rho_\alpha(X) \mu(d\alpha) + \epsilon.$$

Tomando supremos sobre toda $\mu \in P_0$, tomando $\epsilon \downarrow 0$ y con la propiedad de Fatou se sigue que

$$\rho(X) \leq \liminf_{n \rightarrow \infty} \rho(X_n) \leq \limsup_{n \rightarrow \infty} \rho(X_n) \leq \rho(X),$$

y así $\lim_{n \rightarrow \infty} \rho(X_n) = \rho(X)$.

Ahora mostraremos que $\rho(X) = \rho'(X)$. Sea $X_n = X + Y_n$ con $Y_n := (-X - n) \vee 0$. Por subaditividad de ρ' tenemos que

$$\rho(X_n) = \rho'(X_n) \leq \rho'(X) + \rho'(Y_n) = \rho'(X) + \rho(Y_n).$$

Notemos que $Y_n \rightarrow_{L^1} 0$ y $\rho(0) = 0$, por lo que tomando límite cuando $n \rightarrow \infty$ en la última expresión tenemos que $\rho(X) \leq \rho'(X)$. La propiedad de Fatou para ρ' implica

$$\rho'(X) \leq \liminf_{n \rightarrow \infty} \rho'(X_n) = \liminf_{n \rightarrow \infty} \rho(X_n) = \rho(X)$$

Y así $\rho = \rho'$. ■

Con los resultados anteriores podemos obtener fácilmente el siguiente resultado.

Teorema 3.2.6. *Sea (Ω, \mathcal{A}, P) un espacio de probabilidad estándar con P no atómica. Sea $\rho : L^1(\Omega, \mathcal{A}, P) \rightarrow \mathbb{R}$ una medida de riesgo coherente e invariante en ley con la propiedad de Fatou. Entonces $X \leq_{icx} Y$ implica $\rho(X) \leq \rho(Y)$.*

Demostración.

Por el resultado del Teorema 3.1.3, existe Z en (Ω, \mathcal{A}, P) tal que $X \leq_{st} Z \leq_{cx} Y$.

Por un lado $X \leq_{st} Z$, por el Teorema 3.1.1 tenemos que $q_X(\alpha) \leq q_Z(\alpha)$ para toda $\alpha \in (0, 1)$, por lo tanto

$$\rho_\alpha(X) = \frac{1}{\alpha} \int_{1-\alpha}^1 q_X(s) ds \leq \frac{1}{\alpha} \int_{1-\alpha}^1 q_Z(s) ds = \rho_\alpha(Z).$$

Por otro lado $Z \leq_{cx} Y$, por el Teorema 3.1.2 tenemos que

$$\int_{1-\alpha}^1 q_Z(s) ds \leq \int_{1-\alpha}^1 q_Y(s) ds \text{ para } \alpha \in (0, 1),$$

por lo que $\rho_\alpha(Z) \leq \rho_\alpha(Y)$. Por la representación en el Teorema 3.2.5, se sigue que $\rho(X) \leq \rho(Z)$ y $\rho(Z) \leq \rho(Y)$, por lo tanto $\rho(X) \leq \rho(Y)$. ■

Capítulo 4

Optimización de Medidas de Riesgo

Consideremos un problema de optimización de inversión de carteras en un periodo, la regla de decisión se denota por el vector columna π que representa las fracciones del total del portafolio que se le asignará a cada activo. Se asume que $\pi^T \cdot e = 1$, donde e es el vector de unos con la misma longitud que el número de activos del portafolio m . Supongamos que todos los recursos están invertidos en estos m activos.

El rendimiento/ganancia se denota con la variable aleatoria r cuya función de distribución depende de la política π . Sea $R_t = g(r)$ el rendimiento agregado, $Z^\pi = Z(\pi)$ el valor de la medida de riesgo correspondiente a la regla de decisión π y con respecto a g .

Minimización del C-VaR de un Portafolio

Como antes, supondremos que hay m acciones en el mercado y un inversionista está interesado en invertir en ellos. El capital inicial es conocido, suponemos que no es una pérdida y en general asumimos que es de \$1. Sea $Y = (Y_1, Y_2, \dots, Y_m)^T$ el vector aleatorio de porcentajes de rendimiento de m acciones. Por conveniencia supondremos que $Y_k, k = 1, \dots, m$ son variables aleatorias continuas.

Si el capital inicial es \$1 y la política de inversión π es utilizada, el rendimiento aleatorio total será:

$$R(\pi, Y) = 1 + \pi^T Y.$$

Cambiaremos la notación del Valor en Riesgo, α - VaR asociado al rendimiento total $R(\pi, Y)$ se denotará como

$$\zeta_\alpha^\pi = \zeta_\alpha(\pi) = \min\{\zeta | P_Y(R(\pi, Y) \leq \zeta) \geq \alpha\}.$$

Y el valor condicionado al riesgo, denotado por α - $CVaR$ asociado al rendimiento total $R(\pi, Y)$ se denotará como

$$\phi_\alpha^\pi = \phi_\alpha(\pi) = E[R(\pi, Y) | R(\pi, Y) \geq \zeta_\alpha].$$

Para un valor fijo α , si elegimos $Z(\pi) = \zeta_\alpha(\pi)$, entonces utilizamos el VaR como medida de riesgo, de otro modo, si $Z(\pi) = \phi_\alpha(\pi)$ utilizaremos el $CVaR$ como medida de riesgo.

Como $\phi_\alpha(\pi)$ es una función convexa en π podemos calcularla como sigue:

$$\phi_\alpha(\pi) = \min_{\zeta} F_\alpha(\pi, \zeta),$$

donde $F_\alpha(\pi, \zeta) = \zeta + \frac{1}{1-\alpha} E_Y[|R(\pi, Y) - \zeta|^+]$.

Supondremos que la distribución del rendimiento aleatorio Y_t es conocida y esta denotada por $p(y)$. Además, podemos generar muestras y^k , $k = 1, \dots, q$.

Así, podemos obtener una aproximación $\tilde{F}_\alpha(\pi, \zeta)$ de $F_\alpha(\pi, \zeta)$ como sigue:

$$\tilde{F}_\alpha(\pi, \zeta) = \zeta + \frac{1}{1-\alpha} \sum_{k=1}^q [1 + \pi^T y^k - \zeta]^+.$$

Basado en lo anterior, tenemos el siguiente algoritmo de optimización para obtener una política óptima (cartera de inversión).

$$\left\{ \begin{array}{l} \min \zeta + \frac{1}{q(1-\alpha)} \sum_{k=1}^q u_k \\ \text{sujeto a} \\ \pi_j \geq 0, j = 1, \dots, m; \sum_{j=1}^m \pi_j = 1; \\ u_k \geq 0, u_k + (1 + \pi^T y^k) - \zeta, k = 1, \dots, q, \end{array} \right. \quad (4-1)$$

donde u_k representa una variable de holgura.

4.1. Metodología

En esta subsección se presenta la optimización del C-VAR de dos portafolios de rendimientos, en n periodos, uno con 3 acciones y otro con 9 acciones que cotizan en la Bolsa Mexicana de Valores. Dado que la optimización en cada tiempo, requiere q escenarios del comportamiento del portafolio de rendimientos de los activos, se realizó un ajuste de series de tiempo multivariadas, para posteriormente, con el modelo obtenido, realizar la simulación de los escenarios.

La primera optimización el portafolio consta de los rendimientos de las acciones de; BIMBOA (Grupo Bimbo, S.A.B. de C.V.), AMXL (America Movil, S.A.B. de C.V.) y LIVERPOL (El puerto de Liverpool, S.A.B. de C.V.). Este primer portafolio esta constituido de dicha forma ya que buscamos diversificar el riesgo, más adelante veremos que los activos están poco correlacionados.

La segunda optimización el portafolio consta de los rendimientos de las acciones de: BIMBOA (Grupo Bimbo, S.A.B. de C.V.), MFRISCOA (Minera FRISCO, S.A.B. de C.V.) , ELEMENT (Elementia, S.A.B. de C.V.) , AMXL (America Movil, S.A.B. de C.V.), AEROMEX (Grupo Aeroméxico, S.A.B. de C.V.) , TLEVICPO (Grupo Televisa, S.A.B. de C.V.) , LIVERPOLC (El puerto de Liverpool, S.A.B. de C.V.), HCITY (Hoteles City Express, S.A.B. de C.V.) Y AC (Arca Continental, S.A.B. de C.V.). Con este último portafolio se desea diversificar el riesgo, ya que dichas empresas pertenecen a distintos sectores.

4.2. Rendimientos

La mayoría de los estudios financieros incluyen rendimientos, en lugar de precios, de acciones. Una de la razones es que las series de tiempo de rendimientos tienen propiedades estadísticas más atractivas.

Sea P_t el precio de una acción a tiempo t . Supondremos que las acciones con las que trabajaremos no pagan dividendos.

Definición 4.2.1. Rendimiento simple a un periodo

Mantener el activo durante un periodo desde la fecha $t - 1$ hasta la fecha t resultaría en un

simple rendimiento bruto

$$1 + R_t = \frac{P_t}{P_{t-1}}$$

o equivalentemente

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

El correspondiente rendimiento neto simple de un período o el rendimiento simple es

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}.$$

Se tomaron 443 observaciones diarias que consideran los precios de las acciones del 03/01/2017 al 12/09/2018 y se obtuvieron los rendimientos simples diarios del mismo.

4.3. Series de Tiempo Multivariadas

Una serie de tiempo multivariada consiste de varias series de tiempo denominadas componentes. Sea $\mathbf{r}_t = (r_{1t}, r_{2t}, \dots, r_{kt})'$ el vector de k rendimientos a tiempo t , donde a' denota la transpuesta de a . Por ejemplo, un inversionista tiene acciones en AMX, Bimbo, CEMEX y GCARSO consideremos los rendimientos diarios en un período de $t = 5$ de las empresas. Entonces, r_{1t} denota los rendimientos diarios de las acciones de AMX, r_{2t} son los rendimientos diarios de Bimbo y así sucesivamente.

4.3.1. Estacionariedad débil y matrices de correlación cruzada

Una serie de tiempo k -dimensional $\mathbf{r}_t = (r_{1t}, r_{2t}, \dots, r_{kt})'$ es débilmente estacionaria si su primer y segundo momento son invariantes en el tiempo. En particular, el vector de medias y la matriz de covarianzas de una serie de tiempo débilmente estacionaria es constante en el tiempo.

A menos que se indique lo contrario, supondremos que las series de tiempo de los rendimientos accionarios son débilmente estacionarias.

Para una serie de tiempo estacionaria \mathbf{r}_t , definimos el vector de medias y la matriz de covarianzas como

$$\mu = E(\mathbf{r}_t), \quad \Gamma_0 = E[(\mathbf{r}_t - \mu)(\mathbf{r}_t - \mu)'], \quad (4-2)$$

donde la esperanza es tomada elemento por elemento sobre la distribución conjunta de (\mathbf{r}_t) . La media μ es un vector de dimensión k de esperanzas no condicionales de los componentes de (\mathbf{r}_t) . La matriz de covarianzas Γ_0 es una matriz de dimensión $k \times k$. El i -ésimo elemento de la diagonal de Γ_0 es la varianza de (\mathbf{r}_t) , mientras que la entrada (i, j) de Γ_0 es la covarianza entre r_{it} y r_{jt} . Los elementos del vector de medias y la matriz de covarianzas se pueden ver como $\mu = (\mu_1, \dots, \mu_k)'$ y $\Gamma_0 = [\Gamma_{ij}(0)]$, respectivamente.

4.3.2. Matrices de correlación cruzada

Sea \mathbf{D} la matriz diagonal de dimensión $k \times k$ constituida por las desviaciones estándar de r_{it} para $i = 1, \dots, k$. En otras palabras, $\mathbf{D} = \text{diag}\{\sqrt{\Gamma_{11}(0)}, \dots, \sqrt{\Gamma_{kk}(0)}\}$. El retraso cero, es la matriz de correlación cruzada de \mathbf{r}_t definida como

$$\rho_0 = [\rho_{ij}(0)] = \mathbf{D}^{-1}\Gamma_0\mathbf{D}^{-1}.$$

Mas específicamente, el elemento (i, j) de ρ_0 es,

$$\rho_{ij}(0) = \frac{\Gamma_{ij}(0)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}} = \frac{\text{Cov}(r_{it}, r_{jt})}{\text{std}(r_{it})\text{std}(r_{jt})},$$

el coeficiente de correlación entre r_{it} y r_{jt} . En el análisis de series de tiempo, a cada elemento se le conoce como un coeficiente de correlación concurrente o contemporáneo, ya que es la correlación de dos series a tiempo t . Se cumple también que $\rho_{ij}(0) = \rho_{ji}(0)$, $-1 \leq \rho_{ij} \leq 1$, y $\rho_{ii} = 1$ para $1 \leq i, j \leq k$.

Así, $\rho(0)$ es una matriz simétrica con elementos diagonales unitarios.

Un tema importante en el análisis de series de tiempo multivariadas es la relación entre el

tiempo y el retraso de las componentes de las series. Para terminar, las matrices de correlación cruzada se utilizan para medir la fuerza de dependencia entre las series de tiempo. La matriz de covarianza cruzada de \mathbf{r}_t con retraso l se define como

$$\Gamma_l = [\Gamma_{ij}(l)] = E[(\mathbf{r}_t - \mu)(\mathbf{r}_{t-l} - \mu)'], \quad (4-3)$$

donde μ es el vector de medias de \mathbf{r}_t . Por lo tanto, el elemento (i, j) de Γ_l es la covarianza entre r_{it} y $r_{j,t-l}$. Para series débilmente estacionarias, la matriz de covarianza cruzada Γ_l es una función de l , no del índice de tiempo. La matriz de correlación cruzada (MCC) con retraso l de \mathbf{r}_t es definida como

$$\rho_l = [\rho_{ij}(l)] = D^{-1}\Gamma_l D^{-1}, \quad (4-4)$$

donde, como antes, D es la matriz diagonal de desviaciones estándar de las series individuales r_{it} . Por definición,

$$\rho_{ij}(l) = \frac{\Gamma_{ij}(l)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}} = \frac{Cov(r_{it}, r_{j,t-l})}{std(r_{it})std(r_{jt})}, \quad (4-5)$$

es el coeficiente de correlación entre r_{it} y $r_{j,t-l}$. Cuando $l > 0$, este coeficiente mide la dependencia lineal de r_{it} en $r_{j,t-l}$, que ocurrió antes del tiempo t . Entonces, si $\rho_{ij}(l) \neq 0$ y $l > 0$, decimos que la serie r_{jt} lleva la serie r_{it} en el retraso l . De manera similar, $\rho_{ji}(l)$ mide la dependencia lineal de r_{jt} y $r_{i,t-l}$, y decimos que la serie r_{it} lleva la serie r_{jt} en el retraso l si $\rho_{ji}(l) \neq 0$ y $l > 0$. Los elementos de la diagonal de $\rho_{ij}(l)$ son los coeficientes de correlación de retraso l de r_{it} .

De las observaciones anteriores tenemos algunos resultados. En general, $\rho_{ij}(l) \neq \rho_{ji}(l)$ ya que los dos coeficientes de correlación miden diferentes relaciones lineales entre $\{r_{it}\}$ y $\{r_{jt}\}$. Por lo tanto, Γ_l y ρ_l en general, no son simétricas.

Usando $Cov(x, y) = Cov(y, x)$ y bajo el supuesto de estacionariedad débil, tenemos que

$$Cov(r_{it}, r_{j,t-l}) = Cov(r_{j,t-l}, r_{it}) = Cov(r_{it}, r_{j,t+l}) = Cov(r_{it}, r_{j,t-(-l)}),$$

así $\Gamma_{ij}(l) = \Gamma_{ij}(-l)$. Dado que $\Gamma_{ij}(-l)$ es el elemento (i, j) de la matriz de Γ_{-l} y la igualdad se da para $1 \leq i, j \leq k$, tenemos que $\Gamma_l = \Gamma'_{-l}$ y $\rho_l = \rho'_l$.

4.3.3. Estimación de las Matrices de Correlación Cruzada

Dados los datos $\{r_t | t = 1, \dots, T\}$, la matriz de covarianza cruzada Γ_l se puede estimar como

$$\hat{\Gamma}_l = \frac{1}{T} \sum_{t=l+1}^T (r_t - \bar{r})(r_{t-l} - \bar{r})', \quad l \geq 0, \quad (4-6)$$

donde $\bar{r} = (\sum_{t=1}^T r_t) / T$ es el vector de medias muestrales. La matriz de correlaciones cruzadas ρ_l se puede estimar como

$$\hat{\rho}_l = \hat{D}^{-1} \hat{\Gamma}_l \hat{D}^{-1}, \quad l \geq 0, \quad (4-7)$$

donde \hat{D} es la matriz diagonal de dimensión $k \times k$ de la desviación estándar muestral de los componentes de la serie.

4.3.4. Prueba de Portmanteau Multivariado

La prueba univariada Ljung-Box $Q(m)$ se ha generalizado para series de tiempo multivariadas por Hosking (1980), Li y McLeod (1981). Para una serie de tiempo multivariada, la hipótesis nula de la prueba estadística es $H_0 : \rho_1 = \dots = \rho_m = 0$ vs la hipótesis alternativa $H_1 : \rho_i \neq 0$ para algún $i \in \{1, \dots, m\}$. Así, la estadística de prueba se utiliza para probar que no hay autocorrelaciones ni correlaciones cruzadas en el vector de series r_t . El estadístico de prueba toma la siguiente forma

$$Q_k(m) = T^2 \sum_{l=1}^m \frac{1}{T-l} \text{tr}(\hat{\Gamma}'_l \hat{\Gamma}_0^{-1} \hat{\Gamma}_l \hat{\Gamma}_l^{-1}), \quad (4-8)$$

donde T es el tamaño de la muestra, k es la dimensión de r_t , y $\text{tr}(A)$ es la traza de la matriz A , que es la suma de los elementos de la diagonal de A . Bajo la hipótesis nula y bajo ciertas condiciones de regularidad, $Q_k(m)$ sigue una distribución asintótica Ji-Cuadrada con $k^2 m$ gra-

dos de libertad. Si se rechaza la hipótesis nula, entonces se puede ajustar un modelo de series de tiempo multivariado.

4.3.5. Modelos de Vectores Autorregresivos

Un modelo vectorial simple y útil para modelar rendimientos de activos es el modelo de vector de autorregresivos (VAR). Una serie de tiempo multivariada r_t es un proceso VAR de orden 1, o VAR(1) si sigue el modelo:

$$r_t = \phi_0 + \Phi r_{t-1} + a_t, \quad (4-9)$$

donde ϕ_0 es un vector de dimensión k , Φ es una matriz de dimensión $k \times k$, y $\{a_t\}$ es una sucesión de vectores aleatorios no correlacionados con media cero y matriz de varianzas Σ . En la aplicación, la matriz de covarianzas Σ es positiva definida, en otro caso, la dimensión de r_t se reduciría. Se supone que a_t sigue una distribución normal multivariada.

Consideremos el caso bivariado (es decir, $k = 2$, $r_t = (r_{1t}, r_{2t})'$ y $a_t = (a_{1t}, a_{2t})'$). El modelo VAR(1) consiste en las siguientes dos ecuaciones:

$$\begin{aligned} r_{1t} &= \phi_{10} + \Phi_{11}r_{1,t-1} + \Phi_{12}r_{2,t-1} + a_{1t}, \\ r_{2t} &= \phi_{20} + \Phi_{21}r_{1,t-1} + \Phi_{22}r_{2,t-1} + a_{2t}, \end{aligned} \quad (4-10)$$

donde Φ_{ij} es el elemento (i, j) de Φ y ϕ_{i0} es el i -ésimo elemento de ϕ_0 . Basado en la primera ecuación Φ_{12} denota la dependencia lineal de r_{1t} en $r_{2,t-1}$ en presencia de $r_{1,t-1}$. Por lo tanto, Φ_{12} es el efecto condicional de $r_{2,t-1}$ en r_{1t} dado $r_{1,t-1}$. Si $\Phi_{12} = 0$, entonces r_{1t} no depende de $r_{2,t-1}$, y el modelo muestra que r_{1t} solo depende de su propio pasado. De igual manera, si $\Phi_{21} = 0$, entonces la segunda ecuación nos dice que r_{2t} no depende de $r_{1,t-1}$ cuando $r_{2,t-1}$ es dado.

Consideremos las dos ecuaciones juntas. Si $\Phi_{12} = 0$ y $\Phi_{21} \neq 0$, entonces existe una relación no direccional de r_{1t} en r_{2t} . Si $\Phi_{12} = \Phi_{21} = 0$, entonces r_{1t} y r_{2t} no están acoplados. Si $\Phi_{12} \neq 0$

y $\Phi_{21} \neq 0$, entonces no existe una relación entre las dos series.

4.3.6. Condiciones de estacionariedad y Momentos del Modelo VAR(1)

Suponemos que el modelo $VAR(1)$ de la ecuación (4-9) es débilmente estacionario. Tomando esperanzas en el modelo y utilizando $E[a_t] = 0$, tenemos que

$$E[r_t] = \Phi_0 + \Phi E[r_{t-1}]$$

Dado que $E[r_t]$ es invariante en el tiempo, tenemos que

$$\mu = E[r_t] = (I - \Phi)^{-1} \Phi_0$$

siempre que la matriz $I - \Phi$ sea no singular, donde I es la matriz identidad de dimensión $k \times k$.

Utilizando el hecho de que $\Phi_0 = (I - \Phi)\mu$, el modelo $VAR(1)$ en la ecuación (4-9) se puede reescribir como:

$$(r_t - \mu) = \Phi(r_{t-1} - \mu) + a_t$$

Sea $\bar{r}_t = r_t - \mu$ la serie de tiempo de la media. Entonces el modelo $VAR(1)$ se reescribe como:

$$\bar{r}_t = \Phi \bar{r}_{t-1} + a_t. \tag{4-11}$$

Este modelo se puede utilizar para derivar algunas propiedades del modelo $VAR(1)$. Por sustitución repetida, podemos reescribir la ecuación (4-11) como

$$\bar{r}_t = a_t + \Phi a_{t-1} + \Phi^2 a_{t-2} + \Phi^3 a_{t-3} + \dots$$

Esta expresión nos muestra varias características del proceso $VAR(1)$. Dado que a_t es una sucesión no correlacionada, se sigue que $Cov(a_t, r_{t-1}) = 0$. De hecho, a_t no está correlacionada

con r_{t-l} para toda $l > 0$. Por esta razón, a_t se le conoce como el choque o la innovación de la serie a tiempo t .

Si multiplicamos la ecuación (4-9) por a_t' , tomando esperanza, y utilizando el hecho de que a_t es una sucesión no correlacionada al proceso, tenemos que $Cov(r_t, a_t) = \Sigma$.

Para el modelo $VAR(1)$, r_t depende de las pasadas innovaciones a_{t-j} con matriz de coeficientes Φ^j . Para que tal dependencia sea significativa, Φ^j debe converger a cero cuando $j \rightarrow \infty$. Esto se refiere a que los k eigenvalores de Φ deben ser en módulo menores a 1, en otro caso, Φ^j convergerá a una matriz distinta de cero cuando $j \rightarrow \infty$. Más aún, la condición de que los eigenvalores de Φ son menores en módulo a 1 es una condición suficiente y necesaria para que r_t sea débilmente estacionaria siempre que la matriz de covarianzas a_t exista. Dicha condición se reduce del caso univariado del modelo $AR(1)$ en donde se pide que el coeficiente del polinomio de retraso ϕ cumpla que $|\phi| < 1$. Además, ya que

$$|\lambda I - \Phi| = \lambda^k \left| I - \Phi \frac{1}{\lambda} \right|,$$

los eigenvalores de Φ son los inversos de los ceros del determinante $|I - \Phi B|$.

Así, una condición suficiente y necesaria equivalente para la estacionariedad de r_t es que todos las raíces del determinante $|\Phi(B)|$ son en módulo, mayores a uno; es decir, todas las raíces deben estar afuera del círculo unitario en el plano complejo.

Por la ecuación (4-2), la varianza de los rendimientos r_t tienen la forma:

$$Var(r_t) = Cov(r_t, r_t) = \Gamma_0 = \Sigma + \Phi \Sigma \Phi' + \Phi^2 \Sigma (\Phi^2)' + \dots = \sum_{i=0}^{\infty} \Phi^i \Sigma (\Phi^i)',$$

donde se entiende que $\Phi^0 = I$ es la matriz identidad de dimensión $k \times k$. Multiplicando la ecuación (4-11), tomando esperanzas, y utilizando el resultado de que $Cov(a_t, a_{t-j}) = E[a_t \bar{r}'_{t-j}] = 0$ para toda $j > 0$, tenemos que

$$E(\bar{r}_t \bar{r}'_{t-j}) = \Phi E(\bar{r}_t \bar{r}'_{t-j}), \quad l > 0.$$

Por lo tanto,

$$\Gamma_l = \Phi\Gamma_{l-1}, \quad l > 0, \quad (4-12)$$

donde Γ_j es la matriz de covarianza cruzada de retraso j de r_t . Sustituyendo repetidamente la ecuación (4-12) tenemos que

$$\Gamma_l = \Phi^l\Gamma_0 \quad \text{para } l > 0.$$

Y multiplicando la ecuación por $D^{-1/2}$, tenemos que

$$\rho_l = D^{-1/2}\Phi\Gamma_{l-1}D^{-1/2} = D^{-1/2}\Phi D^{1/2}D^{-1/2}\Gamma_{l-1}D^{-1/2} = \Upsilon\rho_{l-1},$$

donde $\Upsilon = D^{-1/2}\Phi D^{1/2}$. Por consiguiente, la matriz de correlaciones cruzadas de un modelo VAR(1) satisface:

$$\rho_l = \Upsilon^l\rho_0 \quad \text{para } l > 0.$$

4.3.7. Vector AR(p)

La generalización del modelo VAR(1) al VAR(p) es sencilla. La serie de tiempo r_t sigue un modelo VAR(p) si satisface

$$r_t = \phi_0 + \Phi_1 r_{t-1} + \dots + \Phi_p r_{t-p} + a_t, \quad p > 0, \quad (4-13)$$

donde ϕ_0 y a_t son como se definió antes, y Φ_j es una matriz de $k \times k$. Utilizando el operador de retraso B , el modelo VAR(p) se puede escribir como

$$(I - \Phi_1 B - \dots - \Phi_p B^p) r_t = \phi_0 + a_t,$$

donde I es la matriz identidad de dimensión $k \times k$. Esta representación se puede escribir de manera compacta como

$$\Phi(B)r_t = \phi_0 + a_t,$$

donde $\Phi(B) = (I - \Phi_1 B - \dots - \Phi_p B^p)$ es una matriz polinomial. Si r_t es débilmente estacionaria, tenemos que

$$\mu = E[r_t] = (I - \Phi_1 B - \dots - \Phi_p B^p)^{-1} \phi_0 = [\Phi(1)]^{-1} \phi_0$$

siempre que exista la inversa. Sea $\bar{r}_t = r_t - \mu$. El modelo VAR(p) es

$$\bar{r}_t = \Phi_1 \bar{r}_{t-1} + \dots + \Phi_p \bar{r}_{t-p} + a_t. \quad (4-14)$$

Utilizando las ecuaciones anteriores y siguiendo los mismos pasos que en el modelo VAR(1), tenemos que

- $Cov(r_t, a_t) = \Sigma$, la matriz de covarianzas de a_t ;
- $Cov(r_{t-l}, a_t) = 0$ para $l > 0$;
- $\Gamma_l = \Phi_1 \Gamma_{l-1} + \dots + \Phi_p \Gamma_{l-p}$ para $l > 0$.

A la última propiedad se le conoce como la ecuación de momento del modelo VAR(p). En términos de la matriz de correlaciones cruzadas, las ecuaciones de momento son

$$\rho_l = \Upsilon_1 \rho_{l-1} + \dots + \Upsilon_p \rho_{l-p} \text{ para } l > 0,$$

donde $\Upsilon_i = D^{-1/2} \Phi_i D^{1/2}$. Para poder entender mejor las propiedades del modelo VAR(p) de la ecuación (4-13) haremos uso de algunos resultados del modelo VAR(1) de la ecuación (4-9). Esto es posible transformando el modelo VAR(p) de r_t en un modelo VAR(1) kp -dimensional. Específicamente, sean $x_t = (\bar{r}'_{t-p+1}, \bar{r}'_{t-p+2}, \dots, \bar{r}'_t)'$ y $b_t = (0, \dots, 0, a'_t)'$ dos procesos kp -dimensionales. La media de b_t es cero, la matriz de covarianzas de b_t es una matriz de dimensión $kp \times kp$ con entradas ceros excepto en la esquina inferior derecha, que es Σ . El modelo VAR(p) de r_t se puede escribir de la forma

$$x_t = \Phi^* x_{t-1} + b_t, \quad (4-15)$$

donde Φ^* es una matriz de dimensión $kp \times kp$ dada por

$$\Phi^* = \begin{bmatrix} 0 & I & 0 & 0 & \dots & 0 \\ 0 & 0 & I & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & I \\ \Phi_p & \Phi_{p-1} & \Phi_{p-2} & \Phi_{p-3} & \dots & \Phi_1 \end{bmatrix},$$

donde 0 e I son la matriz de ceros y la matriz identidad, respectivamente, de dimensión $k \times k$. A Φ^* se le conoce como la matriz *compañera* de la matriz polinomial $\Phi(B)$.

La ecuación (4-15) representa al modelo VAR(1) de x_t , el cual contiene a r_t en sus últimos k componentes. Los resultados del modelo VAR(1) vistos en la subsección anterior se pueden utilizar para derivar las propiedades del modelo VAR(p) vía la ecuación (4-15). Por ejemplo, de la definición, x_t es débilmente estacionaria si y sólo si r_t es débilmente estacionaria. Por lo tanto, una condición suficiente y necesaria para que el modelo VAR(p) sea débilmente estacionario en la ecuación (4-13) es que todos los eigenvalores (4-15) sean menor, en módulo, que uno. De manera similar que en el caso VAR(1), se puede mostrar que es una condición equivalente a todos los ceros del determinante $|\Phi(B)|$ queden fuera del círculo unitario.

En el análisis de series de tiempo financieras, es reelevante, el estudio de las matrices de coeficientes Φ_l del modelo VAR(p). Por ejemplo, si el elemento (i, j) , $\Phi_{ij}(l)$ de Φ_l es cero para toda l , entonces r_{it} no depende de los valores pasados de r_{jt} . La estructura de la matriz de coeficientes Φ_l prevee información sobre la relación de retraso entre los componentes de r_t .

4.3.8. Construcción del Modelo VAR(p)

El concepto de la función de autocorrelación parcial para las series de tiempo univariadas se puede generalizar para especificar el orden p de una serie de vectores. Consideremos los consecutivos modelos VAR:

$$\begin{aligned}
r_t &= \phi_0 + \Phi_1 r_{t-1} + a_t \\
r_t &= \phi_0 + \Phi_1 r_{t-1} + \Phi_2 r_{t-2} + a_t \\
&\vdots \\
r_t &= \phi_0 + \Phi_1 r_{t-1} + \dots + \Phi_i r_{t-i} + a_t \\
&\vdots
\end{aligned} \tag{4-16}$$

Los parámetros de este modelo se pueden estimar por el método de mínimos cuadrados ordinarios.

Sea

$$R = \begin{bmatrix} 1 & r_{t-1} & 0 & 0 & \dots & 0 \\ 1 & r_{t-1} & r_{t-2} & 0 & \dots & 0 \\ 1 & \dots & \dots & \dots & \dots & \dots \\ 1 & r_{t-1} & r_{t-2} & r_{t-3} & \dots & r_{t-p} \end{bmatrix},$$

$$\Phi = \begin{bmatrix} \phi_0 \\ \Phi_1 \\ \vdots \\ \Phi_p \end{bmatrix}, \quad r = \begin{bmatrix} r_t \\ r_t \\ \vdots \\ r_t \end{bmatrix},$$

El objetivo es encontrar los coeficientes de Φ que se ajustan mejor a la rectas, en el sentido de resolver el problema de minimización de los errores cuadrados

$$\hat{\Phi} = \operatorname{argmin}_{\Phi} S(\Phi),$$

donde la función objetivo S está dada por

$$S(\Phi) = \sum_{i=1}^p \left| r_t - \phi_0 - \sum_{j=1}^p r_{t-j} \Phi_j \right|^2 = \|r - R\Phi\|^2.$$

El problema de minimización tiene una única solución, ya que las p columnas de la matriz

R son linealmente independientes, resolviendo las ecuaciones normales

$$(R' R)\hat{\Phi} = R' r.$$

La matriz $R' R$ es conocida como la matriz Gramian de R , es una matriz semi-definida positiva, y la matriz $R' r$ es conocida como la matriz de momentos de los regresores. Finalmente, $\hat{\Phi}$ es el vector de estimadores y está dado por:

$$\hat{\Phi} = (R' R)^{-1} R' r.$$

Para la ecuación i -ésima en (4-16), sea $\hat{\Phi}_j^{(i)}$ es el estimador por mínimos cuadrados de Φ_j y sea $\hat{\Phi}_0^{(i)}$ el estimador de ϕ_0 , donde el superíndice (i) denota que los estimadores corresponden al modelo VAR(i). Por lo tanto, el residuo es

$$\hat{a}_t^{(i)} = r_t - \hat{\phi}_0^{(i)} - \hat{\Phi}_1^{(i)} r_{t-1} - \dots - \hat{\Phi}_i^{(i)} r_{t-i}.$$

Para $i = 0$, el residual se define como $\hat{r}_t^0 = r_t - \bar{r}$, donde \bar{r} es la media muestral de r_t . El residuo de la matriz de covarianza se define como

$$\hat{\Sigma}_i = \frac{1}{T - 2i - 1} \sum_{t=i+1}^T \hat{a}_t^{(i)} (\hat{a}_t^{(i)})', \quad i \geq 0. \quad (4-17)$$

Para especificar el orden p , se prueba la hipótesis $H_0 : \Phi_l = 0$ vs la hipótesis alternativa $H_1 : \Phi_l \neq 0$ para $l = 1, 2, \dots$. Por ejemplo, usando la primera ecuación de (4-16), podemos probar la hipótesis $H_0 : \Phi_1 = 0$ vs la hipótesis alternativa $H_1 : \Phi_1 \neq 0$. La prueba estadística es

$$M(1) = - \left(T - k - \frac{5}{2} \right) Ln \left(\frac{|\hat{\Sigma}_1|}{|\hat{\Sigma}_0|} \right),$$

donde $\hat{\Sigma}_i$ es definida en (4-17) y $|A|$ denota el determinante de la matriz A . En condiciones de regularidad, la prueba estadística $M(1)$ tiene distribución asintótica Ji-cuadrada con k^2 grados de libertad. En general, utilizaremos la ecuación i e $(i - 1)$ de la ecuación (4-16) para la prueba

$H_0 : \Phi_i = 0$ vs la hipótesis alternativa $H_1 : \Phi_i \neq 0$, es decir, probando un modelo $VAR(i)$ vs un modelo $VAR(i - 1)$. El estadístico de prueba es

$$M(i) = - \left(T - k - i - \frac{3}{2} \right) \log \left(\frac{|\hat{\Sigma}_i|}{|\hat{\Sigma}_{i-1}|} \right).$$

Asintóticamente, $M(i)$ distribuye Ji-cuadrada con k^2 grados de libertad. Alternativamente, podemos utilizar el criterio de información de Akaike (AIC) para seleccionar el orden p . Suponemos que a_t distribuye normal multivariado y consideremos la ecuación i -ésima de (4-16). Podemos estimar el modelo por máxima verosimilitud. Para el modelo AR, la estimación por mínimos cuadrados de $\hat{\phi}_0$ y $\hat{\Phi}_j$ es equivalente a la estimación por máxima verosimilitud. Sin embargo, existen diferencias entre las estimaciones de Σ . Bajo máxima verosimilitud Σ es

$$\tilde{\Sigma}_i = \frac{1}{T} \sum_{t=i+1}^T \hat{a}_t^{(i)} \left[\hat{a}_t^{(i)} \right]'$$

El AIC de un modelo VAR(i) bajo el supuesto de normalidad se define como

$$AIC(i) = \log(|\tilde{\Sigma}_i|) + \frac{2k^2 i}{T}.$$

Para un vector de series de tiempo dado, se puede seleccionar el orden del modelo VAR, p como $AIC(p) = \min_{0 \leq i \leq p_0} AIC(i)$, donde p_0 es un entero positivo preespecificado.

Otros criterios de información del modelo $VAR(i)$ son

$$BIC(i) = \log(|\tilde{\Sigma}_i|) + \frac{k^2 i \log(T)}{T},$$

$$HQ(i) = \log(|\tilde{\Sigma}_i|) + \frac{k^2 i \log(\log(T))}{T}.$$

4.3.9. Predicción

Para un modelo VAR(p), el pronóstico para un paso adelante del origen h es $r_h(1) = \phi_0 + \sum_{i=1}^p \Phi_i r_{h+1-i}$, y el error asociado a la predicción es $e_h(1) = a_{h+1}$. La matriz de covarianzas

del error de la predicción es Σ . Para la predicción en dos pasos, tenemos que sustituir r_{h+1} para obtener;

$$r_h(2) = \phi_0 + \Phi_1 r_h(1) + \sum_{i=1}^p \Phi_i r_{h+2-i},$$

y el error asociado a la predicción es

$$e_h(2) = a_{h+2} + \Phi_1[r_t - r_h(1)] = a_{h+2} + \Phi_1 a_{h+1}.$$

La matriz de covarianzas del error de la predicción en dos pasos es $\Sigma + \Phi_1 \Sigma \Phi_1'$. Si r_t es débilmente estacionaria, entonces la predicción en l pasos $r_h(l)$ converge en media al vector μ a medida que el horizonte de predicción l crece. La matriz de covarianzas de los errores de predicción converge a la matriz de covarianzas de r_t .

4.3.10. Comprobación de Supuestos

Para saber qué tan bien ajustado está nuestro modelo;

$$\hat{r}_t = \phi_0 + \hat{\Phi}_1 r_{t-1} + \hat{\Phi}_2 r_{t-2} + \dots + \hat{\Phi}_p r_{t-p} + \hat{a}_t,$$

debemos de comprobar los supuestos sobre los residuos; no correlación, varianza constante y normalidad. Para ello existen las siguientes pruebas.

1. Multivariado de Portmanteau y Breush-Godfrey de errores correlacionados en serie.

La prueba de Portmanteu verifica la ausencia de correlación en serie de los residuos hasta un orden h en un modelo VAR(p), la estadística de prueba se define como

$$Q_h = T \sum_{j=1}^h \text{tr}(\hat{C}_j' \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}),$$

donde $\hat{C}_i = \sum_{t=i+1}^T \hat{a}_t \hat{a}_{t-i}'$. La estadística de prueba tiene una distribución $\chi^2_{(k^2(h-p))}$. Si el valor de h es pequeño o el tamaño de muestra no es tan grande, se recomienda utilizar la

estadística de prueba

$$Q_h^* = T^2 \Sigma_{j=1}^h \frac{1}{T-j} \text{tr}(\hat{C}'_j \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}).$$

La prueba de Breusch-Godfrey LM se basa en una regresión auxiliar

$$\hat{a}_t = \phi_0 + \Phi_1 r_{t-1} + \dots + \Phi_p r_{t-p} + CD_t + \beta_1 \hat{a}_{t-1} + \dots + \beta_h \hat{a}_{t-h} + \varepsilon_t.$$

La prueba de hipótesis en la que se basa es $H_0 : \beta_1 = \dots = \beta_h = 0$ vs $H_1 : \exists \beta_i \neq 0$ para $i = 1, 2, \dots, h$. La estadística de prueba se define como

$$LM_h = T(k - \text{tr}(\bar{\Sigma}_R^{-1} \bar{\Sigma}_e)),$$

donde $\bar{\Sigma}_R$ y $\bar{\Sigma}_e$ asignan la matriz de covarianza residual del modelo restringido y no restringido, respectivamente. La estadística de prueba LM_h sigue una distribución $\chi^2_{(hk^2)}$.

2. Prueba Heteroscedasticidad de los residuos, ARCH-LM.

La prueba multivariada ARCH-LM se basa en la siguiente regresión:

$$\text{vech}(\hat{a}_t, \hat{a}'_t) = \beta_0 + B_1 \text{vech}(\hat{a}_{t-1}, \hat{a}'_{t-1}) + \dots + B_q \text{vech}(\hat{a}_{t-p}, \hat{a}'_{t-p}) + \alpha_t,$$

donde α_t asigna un proceso de error esférico y vech es el operador de apilamiento de columnas para matrices simétricas. La dimensión de β_0 es $\frac{1}{2}k(k+1)$, y las dimensiones de las matrices de coeficientes B_i para $i = 1, \dots, q$, es $(\frac{1}{2}k(+1) \times \frac{1}{2}k(k+1))$. La hipótesis nula es $H_0 : B_1 = \dots = B_q = 0$ vs la hipótesis alternativa $H_1 : B_1 \neq 0, \dots, B_q \neq 0$. La estadística de prueba se define como

$$VARCH_{LM}(q) = \frac{1}{2} T k(k+1) R_m^2,$$

con

$$R_m^2 = 1 - \frac{2}{k(k+1)} \text{tr}(\hat{\Omega}\hat{\Omega}^{-1}),$$

y $\hat{\Omega}$ representa la matriz de covarianza del modelo de regresión definido anteriormente.

La estadística de prueba sigue una distribución $\chi_{(qk^2(k+1)^2/4)}^2$.

3. Prueba de Normalidad para los residuos.

Para comprobar normalidad, utilizaremos la prueba de Jarque-Bera multivariada aplicada a los residuos, donde H_0 : *los residuos siguen una distribución normal multivariada* vs H_1 : *los residuos no siguen una distribución normal multivariada*. El estadístico de prueba para el caso multivariado se define como

$$JB_{mv} = s_3^2 + s_4^2,$$

donde

$$\begin{aligned} s_3^2 &= T b_1' b_1 / 6, \\ s_4^2 &= T (b_2 - 3k)' (b_2 - 3k) / 24, \end{aligned} \tag{4-18}$$

b_1 y b_2 son el tercer y el cuarto momento de los residuos estandarizados $\hat{a}_t^s = \bar{P} - (\hat{a}_t - \bar{\hat{a}}_t)$ y \bar{P} es una matriz triangular con diagonal positiva tal que $\bar{P}\bar{P}' = \bar{\Sigma}_a$, es decir, es la descomposición de Choleski de la matriz de covarianzas de los residuos. La estadística de prueba JB_{mv} sigue una distribución $\chi_{(2k)}^2$ y el sesgo multivariado, s_3^2 , y la prueba de curtosis, s_4^2 , se distribuyen $\chi_{(k)}^2$.

4.4. Optimización por Simplex

Considere el problema de programación lineal en forma matricial:

$$\text{Maximizar } z = CX \text{ sujeto a } AX = b, X \geq 0,$$

donde X es el vector de tamaño n que representa las variables, A es una matriz de dimensión

$(m \times n)$ que representa los coeficientes de restricción, b es el vector columna que representa el lado derecho, y C como un vector n que representa los coeficientes de la función objetivo.

De forma equivalente el problema puede escribirse como

$$\begin{pmatrix} 1 & -C \\ 0 & A \end{pmatrix} \begin{pmatrix} z \\ X \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}.$$

Suponga que B es una base factible del sistema $AX = b$, $X \geq 0$, decimos que un conjunto de m vectores forman una base B si y sólo si los m vectores seleccionados son linealmente independientes. Sea X_B el correspondiente vector de variables no básicas (es decir, cumple $BX_B = b$) y C_B su vector objetivo asociado. Dado que todas las variables no básicas son cero, la solución se calcula entonces como

$$\begin{pmatrix} z \\ X_B \end{pmatrix} = \begin{pmatrix} 1 & -C_B \\ 0 & B \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ b \end{pmatrix} = \begin{pmatrix} 1 & C_B B^{-1} \\ 0 & B^{-1} \end{pmatrix} \begin{pmatrix} 0 \\ b \end{pmatrix} = \begin{pmatrix} C_B B^{-1} b \\ B^{-1} b \end{pmatrix}.$$

La tabla simplex completa en forma matricial puede derivarse a partir de la ecuación matricial original como

$$\begin{pmatrix} 1 & C_B B^{-1} \\ 0 & B^{-1} \end{pmatrix} \begin{pmatrix} 1 & -C \\ 0 & A \end{pmatrix} \begin{pmatrix} z \\ X \end{pmatrix} = \begin{pmatrix} 1 & C_B B^{-1} \\ 0 & B^{-1} \end{pmatrix} \begin{pmatrix} 0 \\ b \end{pmatrix}.$$

Realizando operaciones entre matrices, tenemos que

$$\begin{pmatrix} 1 & C_B B^{-1} A - C \\ 0 & B^{-1} A \end{pmatrix} \begin{pmatrix} z \\ X \end{pmatrix} = \begin{pmatrix} C_B B^{-1} b \\ B^{-1} b \end{pmatrix}.$$

Dado el vector j -ésimo P_j de A , la columna de la tabla simplex asociada a la variable x_j puede escribirse como en la tabla 4-1.

Una propiedad importante de esta tabla es que la inversa B^{-1} es el único elemento que cambia con una nueva iteración.

Básica	X_j	Solución
z	$C_B B^{-1} P_j - c_j$	$C_B B^{-1} b$
X_B	$B^{-1} P_j$	$B^{-1} b$

Tabla 4-1: Tabla Simplex

4.4.1. Método Simplex Revisado

El método simplex consiste en buscar una base factible B posteriormente pasar a una base B^* hasta que se alcance la base óptima.

El problema de programación lineal puede escribirse como

$$\begin{aligned} &\text{Máximizarse o mínimizarse } z = \sum_{j=1}^n c_j x_j. \\ &\text{sujeto a } \sum_{j=1}^n P_j x_j = b, \quad x_j \geq 0, \quad j = 1, \dots, n. \end{aligned}$$

Dados los vectores básicos X_B , su base B , y su vector objetivo C_B , la tabla simplex general muestra que en cualquier iteración del método simplex puede representarse mediante las siguientes ecuaciones:

$$\begin{aligned} z + \sum_{j=1}^n (z_j - c_j) x_j &= C_B B^{-1} b, \\ (X_B)_i + \sum_{j=1}^n (B^{-1} P_j)_i x_j &= (B^{-1} b)_i, \end{aligned}$$

donde la notación $(V)_i$ representa el elemento i del vector V . El costo reducido de x_j se calcula como

$$z_j - c_j = C_B B^{-1} P_j - c_j.$$

Condición de optimalidad

La ecuación z muestra que, en el caso de maximización, un incremento de la variable x_j no básica por encima de su valor actual cero puede mejorar el valor de z (con respecto a su valor actual, $C_B B^{-1} b$) solo si $z_j - c_j < 0$. Para minimización, la condición es $z_j - c_j > 0$. Por lo tanto, el vector de entrada se selecciona como un vector no básico con la condición $z_j - c_j$ más negativa (más positiva) según sea el caso.

Condición de factibilidad

Dado el vector de entrada P_j determinado por la condición de optimalidad, las ecuaciones de restricción se reducen a

$$(X_B)_i = (B^{-1}b)_i - (B^{-1}P_j)x_j.$$

El objetivo es aumentar x_j por encima de cero, reemplazando una de las variables básicas actuales. Siempre y cuando todas las entradas de X_B sean no negativas, dicha condición depende del grado en que se incrementa x_j , es decir,

$$(X_B)_i = (B^{-1}b)_i - (B^{-1}P_j)x_j \geq 0.$$

Si $(B^{-1}P_j) > 0$ para al menos una i , la condición de no negatividad $(X_B) \geq 0$ sobre todas las i , establece el límite en el incremento máximo del valor de la variable de entrada x_j ;

$$x_j = \min_i \left\{ \frac{(B^{-1}b)_i}{(B^{-1}P_j)_i} \mid (B^{-1}P_j)_i > 0 \right\}.$$

Supongamos que $(X_B)_k$ es la variable no básica que corresponde a la relación mínima. Entonces P_k debe ser el vector de salida, y la variable básica asociada debe volverse no básica en la siguiente iteración simplex.

4.4.2. Algoritmo del método simplex revisado

1. Construir una solución factible básica de inicio, con B la base asociada y C_B el vector de coeficientes objetivo.
2. Calcular B^{-1} de la base B .
3. Para cada vector no básico P_j calcular

$$z_j - c_j = C_B B^{-1} P_j - c_j,$$

si $z_j - c_j \geq 0$ en maximización (≤ 0 en minimización) para todos los vectores no básicos,

terminar; la solución óptima es $X_B = B^{-1}b$ y $z = C_B X_B$. En caso contrario, seleccionar el vector de entrada P_j que tiene el valor $z_j - c_j$ más negativo (más positivo), entre todos los vectores no básicos, en caso de que el problema de programación lineal sea de maximización (minimización).

4. Calcular $B^{-1}P_j$, si todos los elementos de $B^{-1}P_j$ son negativos o cero, terminar; la solución es no acotada. En caso contrario, utilizar la prueba de relación para determinar el valor de salida P_j .
5. Reemplazar el vector de salida P_i con el vector de entrada P_j en la matriz de la base, B . Repetir el paso 2 para una nueva iteración.

Una vez implementado el algoritmo simplex revisado para la resolución del problema de programación lineal (4-1), presentamos el siguiente ejemplo.

4.5. Ejemplo

En esta sección ejemplificaremos el problema de optimización planteado en la ecuación (4-1), donde los rendimientos siguen una distribución normal multivariada.

Supongamos que queremos invertir en un portafolio con tres instrumentos financieros: S&P 500, una cartera de bonos del gobierno de Estados Unidos y una cartera de acciones de pequeña capitalización, los rendimientos de estos tres instrumentos están modelados por una distribución normal multivariada.

La media μ de los rendimientos está dada por la Tabla 4-2.

Instrumento	Rendimiento medio
S&P	0.0101110
Bonos Gov.	0.0043532
Small cap	0.0137058

Tabla 4-2: Vector de medias

La matriz de varianzas y covarianzas está dada por la Tabla 4-3.

	S&P	Bonos Gov.	Small cap
S&P	0.00324625	0.00022983	0.00420395
Bonos Gov.	0.00022983	0.00049937	0.00019247
Small cap	0.00420395	0.00019247	0.00764097

Tabla 4-3: Matriz de Varianzas y Covarianzas

Para obtener el C-VaR óptimo, realizamos la implementación en R de un Solver que nos devuelve el C-VaR mínimo y la ponderación de cada instrumento en el portafolio a un nivel β dado por el usuario. Para ello primero genera los q escenarios pertenecientes a una distribución normal multivariada dado el vector de medias μ y la matriz de varianzas y covarianzas Σ . (Ver B.1). Los resultados obtenidos se muestran en la tabla 4-4.

q	beta	S&P	Bonos Gov.	Small cap	C-VaR
10	0.90	0.42184265	0.57815735	0	0.01961583
100	0.90	0	0.91661992	0.08338008	0.01876966
500	0.90	0.03717975	0.92841637	0.03440388	0.02339622
10	0.95	0	0.94807094	0.05192906	0.03894116
100	0.95	0.08542543	0.87515289	0.03942168	0.02476857
500	0.95	0.13659520	0.86340480	0	0.03030310

Tabla 4-4: Tabla de rendimientos

4.6. Resultados

Como se mencionó al principio del capítulo, llevamos a cabo las optimizaciones del CVaR de dos portafolios de rendimientos.

En la Figura 4-1 se observa el comportamiento del portafolio constituido por los rendimientos de las acciones de; BIMBOA, MFRISCOA, ELEMENT, AMXL, AEROMEX, TLEVICPO, LIVERPOLC, HCITY y AC.

Veamos cómo se comporta la correlación entre los rendimientos para así crear un portafolio más pequeño que el original y que sea más sencillo de manejar, también deseamos que este portafolio este diversificado. Aunque también realizaremos la optimización del portafolio que consta de 9 activos.

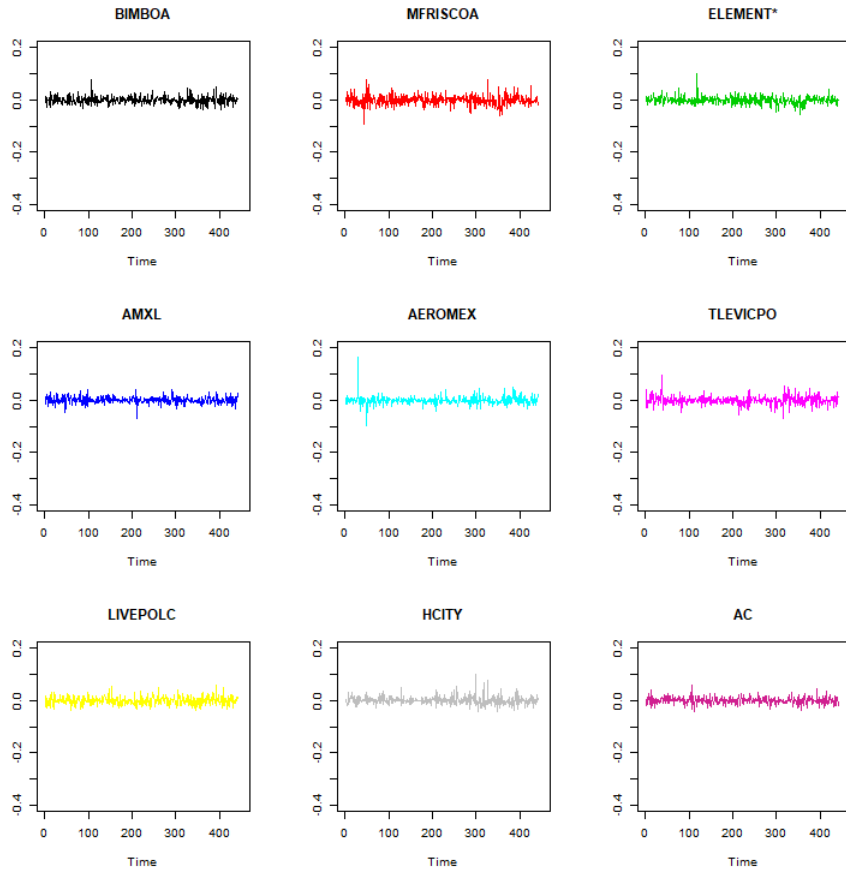


Figura 4-1: Rendimientos del portafolio

En la Figura 4-2 se muestran las correlaciones entre activos, con ayuda del mismo, contruimos el segundo portafolio con los rendimientos de BIMBOA, AMXL y LIVERPOL, ya que los sectores a los que se enfocan son distintos y aparte, las correlaciones entre ellos son pequeñas.

Para cada uno de los portafolios, ajustamos un modelo sencillo de series de tiempo, en este caso un $VAR(1)$, para determinar el grado del modelo, utilizamos el comando *VARselect*, (ver B.2).

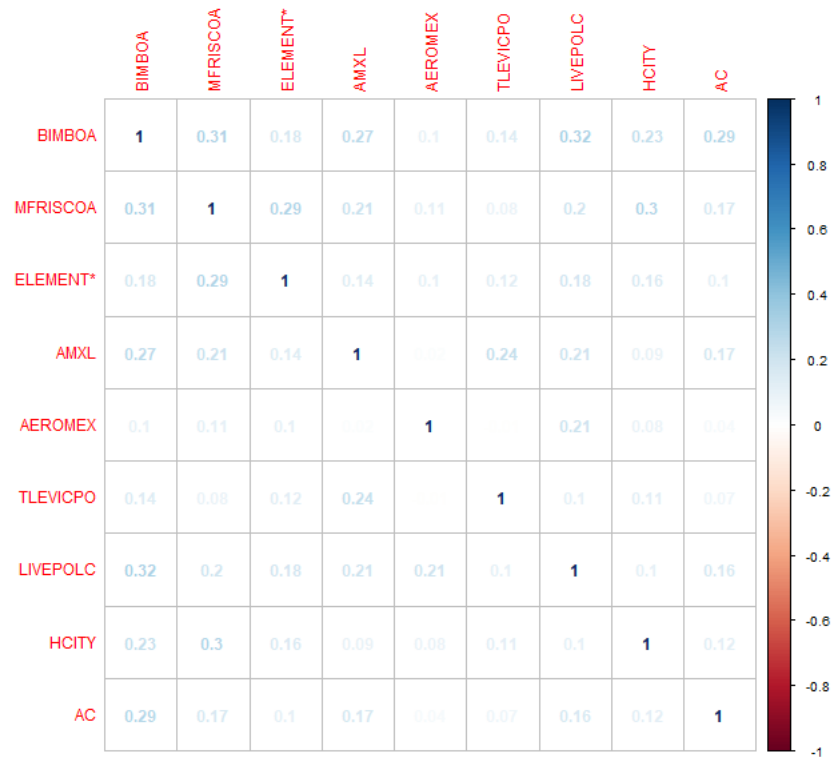


Figura 4-2: Correlaciones del portafolio de 9 activos

Para el modelo de 3 activos, el modelo ajustado fue

$$r_t^1 = \begin{bmatrix} -0.077595421 & -0.026290986 & 0.03629792 \\ 0.070737540 & 0.048403933 & 0.03657111 \\ -0.006063867 & -0.008123815 & -0.08554642 \end{bmatrix} r_{t-1}^1 + a_t^1.$$

con matriz de correlación

$$\Sigma^1 = \begin{bmatrix} 1.978743e-04 & 4.617948e-05 & 4.959373e-05 \\ 4.617948e-05 & 1.598873e-04 & 1.784990e-05 \\ 4.959373e-05 & 1.784990e-05 & 2.339092e-04 \end{bmatrix}.$$

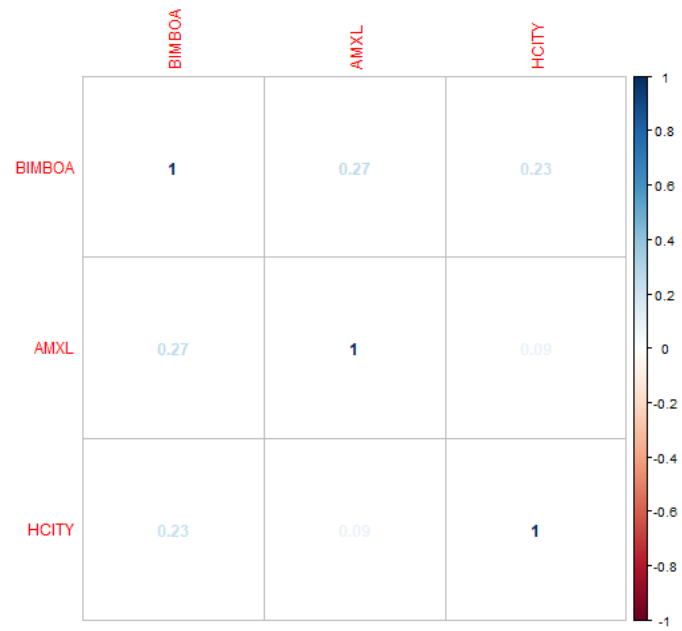


Figura 4-3: Correlaciones del portafolio de 3 activos

Para el modelo de 9 activos, el modelo ajustado fue;

$$r_t^2 = \begin{bmatrix} -0.074884823 & 0.004880266 & -0.010023383 \\ -0.026349423 & -0.017008417 & 0.001745434 \\ -0.015629219 & -0.070542706 & -0.042948108 \\ 0.080317759 & -0.005575245 & -0.056596840 \\ 0.007080103 & -0.035174576 & -0.001297897 \\ -0.025573028 & -0.008375724 & 0.006708607 \\ -0.009353740 & 0.072590185 & 0.075802988 \\ 0.002592313 & -0.055961361 & -0.035110329 \\ 0.051681504 & 0.044017397 & 0.050035746 \end{bmatrix}$$

$$\begin{matrix} -0.027598269 & 0.083105121 & -0.004171595 \\ 0.004273727 & -0.054734760 & 0.003010478 \\ 0.052076735 & -0.074263481 & 0.062821101 \\ 0.049563794 & 0.163365459 & 0.055382911 \\ -0.029525373 & 0.068940222 & -0.094928304 \\ -0.027542148 & 0.030764948 & -0.017859126 \\ 0.005518958 & -0.034411588 & 0.041708282 \\ -0.010438754 & -0.011892762 & -0.005762490 \\ -0.013795975 & -0.004627609 & 0.043919005 \end{matrix}$$

$$\begin{bmatrix}
 0.107904295 & 0.0411929591 & -0.01683901 \\
 0.018782271 & -0.0410469886 & -0.02007165 \\
 -0.085544968 & 0.0742892457 & 0.02905774 \\
 -0.010783541 & 0.0498001731 & 0.02550598 \\
 0.095302866 & -0.0002297404 & 0.02276766 \\
 0.028037556 & -0.0494961914 & -0.05558376 \\
 -0.086767345 & -0.0041460059 & 0.03126217 \\
 -0.008752918 & -0.0784332577 & -0.04944450 \\
 0.026490162 & 0.0144110131 & -0.15911175
 \end{bmatrix} r_{t-1}^2 + a_t^2.$$

Con matriz de covarianzas

$$\Sigma^2 = \begin{bmatrix}
 1.970128e - 04 & 8.244535e - 05 & 3.556036e - 05 \\
 8.244535e - 05 & 3.632135e - 04 & 7.864472e - 05 \\
 3.556036e - 05 & 7.864472e - 05 & 2.014073e - 04 \\
 4.640547e - 05 & 5.096683e - 05 & 2.438885e - 05 \\
 2.096497e - 05 & 3.223571e - 05 & 2.429835e - 05 \\
 3.208008e - 05 & 2.513231e - 05 & 2.802399e - 05 \\
 6.981985e - 05 & 5.845336e - 05 & 4.009170e - 05 \\
 4.915356e - 05 & 8.472513e - 05 & 3.321024e - 05 \\
 5.132991e - 05 & 3.981592e - 05 & 1.834625e - 05
 \end{bmatrix}$$

$4.640547e - 05$	$2.096497e - 05$	$3.208008e - 05$
$5.096683e - 05$	$3.223571e - 05$	$2.513231e - 05$
$2.438885e - 05$	$2.429835e - 05$	$2.802399e - 05$
$1.590317e - 04$	$3.427752e - 06$	$4.642388e - 05$
$3.427752e - 06$	$2.319905e - 04$	$-3.983425e - 08$
$4.642388e - 05$	$-3.983425e - 08$	$2.622341e - 04$
$4.270672e - 05$	$4.526641e - 05$	$3.008375e - 05$
$1.703974e - 05$	$1.850768e - 05$	$2.557080e - 05$
$2.727327e - 05$	$7.928682e - 06$	$1.473063e - 05$

$6.981985e - 05$	$4.915356e - 05$	$5.132991e - 05$
$5.845336e - 05$	$8.472513e - 05$	$3.981592e - 05$
$4.009170e - 05$	$3.321024e - 05$	$1.834625e - 05$
$4.270672e - 05$	$1.703974e - 05$	$2.727327e - 05$
$4.526641e - 05$	$1.850768e - 05$	$7.928682e - 06$
$3.008375e - 05$	$2.557080e - 05$	$1.473063e - 05$
$2.233403e - 04$	$2.537630e - 05$	$3.370475e - 05$
$2.537630e - 05$	$2.321524e - 04$	$2.169039e - 05$
$3.370475e - 05$	$2.169039e - 05$	$1.569530e - 04$

Posteriormente comprobamos supuestos sobre los residuos.

- Portmanteau-Test (correlación entre los residuos) La prueba aplicada al modelo ajustado del portafolio con 3 activos nos devuelve un p-value = 0.07578, por lo que no existe correlación entre los residuos. La prueba aplicada al modelo ajustado del portafolio con 9 activos nos devuelve un p-value = 0.2161, por lo que no existe correlación entre los residuos.
- Heterosedasticidad (varianza constante) La prueba aplicada al modelo ajustado del por-

tafolio con 3 activos nos devuelve un p-value =0.2169, por lo tanto, la varianza de los residuos es constante. La prueba aplicada al modelo ajustado del portafolio con 9 activos nos devuelve un p-value =3.595e-09, por lo tanto, la varianza de los residuos no es constante.

- Test de normalidad multivariada. La prueba aplicada al modelo ajustado del portafolio con 3 activos nos devuelve un p-value <2.2e-16, por lo que no se cumple el supuesto de normalidad de los residuos. La prueba aplicada al modelo ajustado del portafolio con 9 activos nos devuelve un p-value <2.2e-16, por lo que no se cumple el supuesto de normalidad de los residuos.

Los ajustes realizados no son los adecuados ya que no se cumplen todos los supuestos del modelo, lo ideal sería ajustar otro modelo, sin embargo, esto queda fuera del alcance de esta tesis y se llevará a cabo en trabajos posteriores.

Una vez ajustados los modelos, procedemos a realizar la optimización en t períodos del C-VaR.

Para el primer portafolio, realizamos la optimización en $t = 10$ días, con $n = 100$ escenarios.

Los resultados de la optimización están presentados en la Tabla 4-5 y el comportamiento de las ponderaciones de cada activo se muestran en la Figura 4-4.

t	CVaR-95 %	π_1	π_2	π_3
1	0.01022437	0.15808820	0.3700410	0.47187083
2	0.01390324	0.35786875	0.4741624	0.16796882
3	0.01622675	0.14203093	0.4909282	0.36704091
4	0.01564306	0.00000000	0.5803126	0.41968740
5	0.01381699	0.34018936	0.3756882	0.28412240
6	0.01438617	0.09184933	0.8571061	0.05104457
7	0.01567627	0.34214065	0.2803439	0.37751546
8	0.01669442	0.11956333	0.5103655	0.37007122
9	0.01306342	0.23189109	0.4941629	0.27394596
10	0.01552463	0.18007881	0.5746858	0.24523540

Tabla 4-5: C-VaR mínimo y ponderaciones en cada período para un portafolio de 3 activos

Notemos que la dinámica de las ponderaciones del portafolio cambian diariamente y nos

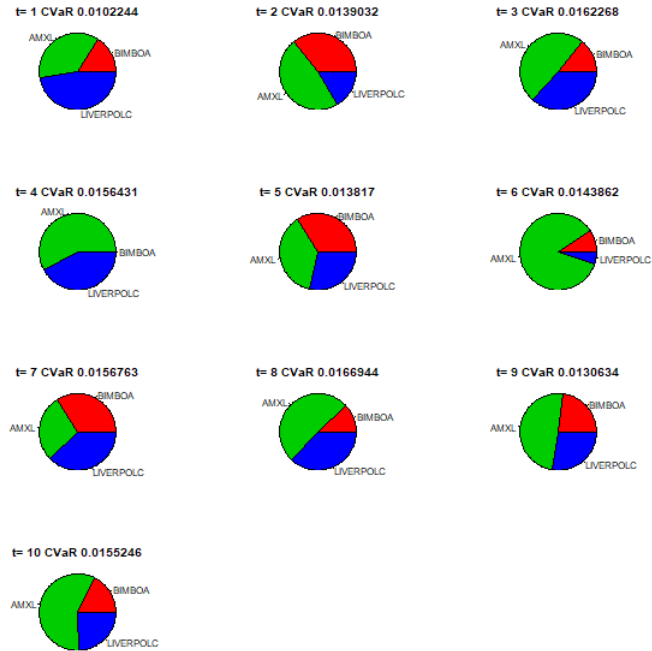


Figura 4-4: Dinámica de las ponderaciones del portafolio de tres activos durante 10 días.

indica la cantidad de activo que debemos comprar o vender según sea el caso para minimizar el C-VaR a un nivel de confianza del 95 % en cada período.

Para el segundo portafolio, realizamos la optimización en $t = 10$ días, con $n = 100$ escenarios. La dinámica de dicho portafolio se muestra en la Tabla 4-7 y gráficamente en la Figura 4-5.

Análogo para el caso cuando tenemos un portafolio con tres acciones, las ponderaciones del segundo portafolio cambian diariamente y nos indica la cantidad de activo que debemos comprar o vender según sea el caso para minimizar el C-VaR a un nivel de confianza del 95 % en cada periodo.

t	CVaR-95 %	π_1	π_2	π_3	π_4	π_5
1	0.00808	0.06611	0.07058	0.21157	0.23171	0.05072
2	0.01177	0.06948	0.00000	0.18379	0.01250	0.20327
3	0.01118	0.18538	0.00000	0.18244	0.09447	0.13324
4	0.01239	0.02777	0.06803	0.08295	0.20937	0.28608
5	0.01192	0.13859	0.00000	0.16091	0.12558	0.05741
6	0.01158	0.00322	0.00000	0.09088	0.39734	0.03561
7	0.01121	0.25854	0.00000	0.23130	0.10163	0.25854
8	0.01125	0.00000	0.04648	0.19592	0.07964	0.26314
9	0.00940	0.00000	0.00000	0.02462	0.10273	0.17811
10	0.01096	0.00000	0.02151	0.07686	0.34843	0.29600

Tabla 4-6: C-VaR mínimo y ponderaciones en cada período para un portafolio con 9 acciones

t	π_6	π_7	π_8	π_9
1	0.02975	0.00000	0.113126	0.22640
2	0.04804	0.16170	0.186171	0.13501
3	0.07756	0.01305	0.120420	0.19340
4	0.28455	0.02355	0.000000	0.01768
5	0.00000	0.19573	0.288446	0.03331
6	0.09032	0.08171	0.005361	0.29553
7	0.00000	0.03460	0.189059	0.07451
8	0.14673	0.16225	0.094894	0.01091
9	0.14053	0.15397	0.194664	0.20534
10	0.08065	0.00000	0.005852	0.17065

Tabla 4-7: C-VaR mínimo y ponderaciones en cada período para un portafolio con 9 acciones

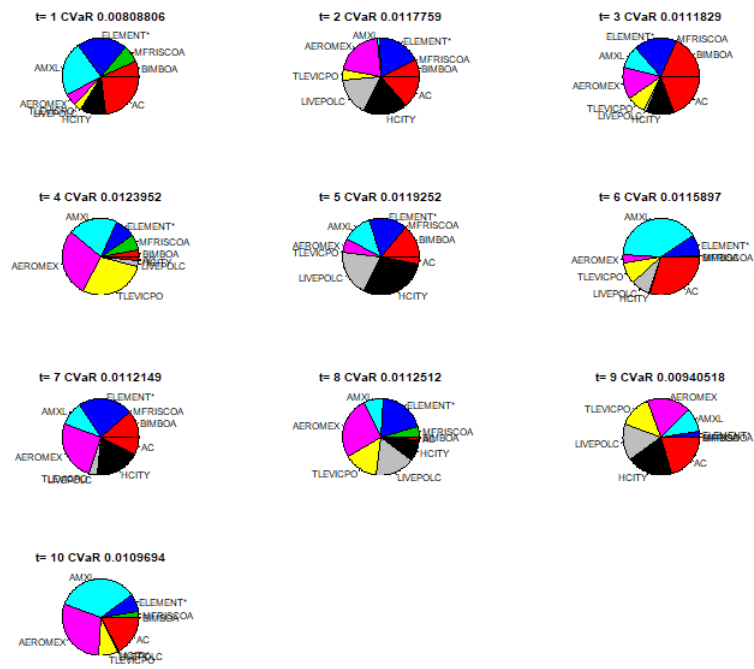


Figura 4-5: Dinámica de las ponderaciones del portafolio de nueve activos durante 10 días.

Conclusiones

Se mostraron las condiciones en las que se preserva el primer orden estocástico, el orden convexo y el orden convexo creciente cuando medimos riesgos, para ello fue importante demostrar la extensión de la caracterización de L^∞ a L^1 de una medida de riesgo invariante en ley y coherente con la propiedad de Fatou. Fue necesario estudiar algunos conceptos de integrabilidad uniforme que se muestran en el Apéndice A, así como, estudiar la propiedad de Fatou para medidas de riesgo.

La optimización del C-VaR a un nivel de confianza α permite que cualquier inversionista cree políticas para mitigar el riesgo que pueden generar los comportamientos inusuales de los activos, así como la oportuna actuación para evitar pérdidas significativas. Como se mencionó en los resultados, la dinámica de las ponderaciones del portafolio van cambiando diariamente lo que nos indica si debemos vender o comprar el diferencial que arroja cada iteración.

Es recomendable que se simulen más escenarios para considerar más casos del comportamiento de los rendimientos que constituyen el portafolio. En el caso práctico presentado en este trabajo, el modelo ajustado para las simulaciones no fue el adecuado pues no se cumplieron los supuestos de normalidad y varianza constante. Por lo que es recomendable cambiar de modelo para que las simulaciones sean más exactas y así tener mayor precisión en las ponderaciones de cada portafolio y de la optimización.

Apéndice A

Apéndice

A.1. Tipos de convergencia

Definición A.1.1. Sea X una variable aleatoria y $\{X_n, n \geq 0\}$ una sucesión de variables aleatorias en $\{\Omega, \mathcal{F}, P\}$, decimos que X_n converge en probabilidad a X ($X_n \rightarrow_P X$) si $\forall \epsilon > 0$

$$\lim_{n \rightarrow \infty} P(\{\omega \mid |X_n(\omega) - X(\omega)| > \epsilon\}) = 0. \quad (\text{A-1})$$

Definición A.1.2. Sea X una variable aleatoria y $\{X_n, n \geq 0\}$ una sucesión de variables aleatorias en $\{\Omega, \mathcal{F}, P\}$, decimos que X_n converge en L^1 a X ($X_n \rightarrow_{L^1} X$) si

$$\lim_{n \rightarrow \infty} E[|X_n - X|] = 0. \quad (\text{A-2})$$

A.2. Convergencia

Teorema A.2.1. Si $\{a_n\}$ es una sucesión monótona de números reales, es decir, si $a_n \leq a_{n+1}$ o $a_{n+1} \leq a_n$ para $n \geq 1$. La sucesión $\{a_n\}$ converge si y sólo si $\{a_n\}$ es acotada.

A.3. Integrabilidad uniforme

Definición A.3.1. Sea una familia $\chi \subseteq L^0$ de variables aleatorias, decimos que χ es uniformemente integrable (UI) si

$$\lim_{K \rightarrow \infty} \left(\sup_{X \in \chi} E[|X|1_{\{|X| \geq K\}}] \right) = 0. \quad (\text{A-3})$$

Comenzaremos por la caracterización de las familias uniformemente integrables.

Proposición A.3.1. Una familia $\chi \subseteq L^0$ de variables aleatorias es uniformemente integrable si y sólo si,

1. existe $C \geq 0$ tal que $E[|X|] \leq C$, para toda $X \in \chi$,
2. para todo $\epsilon > 0$ existe $\delta > 0$ tal que para cualquier $A \in \mathcal{F}$, tenemos que

$$P[A] \leq \delta \rightarrow \sup_{X \in \chi} E[|X|1_A] \leq \epsilon. \quad (\text{A-4})$$

Demostración.

UI \Rightarrow 1,2 Asumimos que χ es UI, elegimos $K > 0$ tal que

$$\sup_{X \in \chi} E[|X|1_{\{|X| > K\}}] \leq 1.$$

Ahora

$$E[|X|] = E[|X|1_{\{|X| \leq K\}}] + E[|X|1_{\{|X| > K\}}] \leq K + E[|X|1_{\{|X| > K\}}],$$

Tomando supremo en ambos lados, para todo X , sabemos que $\sup_{X \in \chi} E[|X|] \leq K + 1$, por lo que probamos 1.

Para probar 2, sea $\epsilon > 0$ y utilizamos el supuesto de integrabilidad uniforme de χ para encontrar una constante $K > 0$ tal que $\sup_{X \in \chi} E[|X|1_{\{|X| > K\}}] \leq \epsilon/2$ Para $\delta = \frac{\epsilon}{2K}$ y $A \in \mathcal{F}$, la

condición $P[A] \leq \delta$ implica

$$\begin{aligned} E[|X|1_A] &= E[|X|1_A 1_{\{|X| \leq K\}}] + E[|X|1_A 1_{\{|X| > K\}}] \\ &\leq KP[A] + E[|X|1_{\{|X| > K\}}] \leq \epsilon \end{aligned} \tag{A-5}$$

Dado que lo anterior se cumple para todo $X \in \chi$ en particular se cumple para el supremo.

1, 2 \Rightarrow UI Sea $C > 0$ para que se cumpla 1, sea $\epsilon > 0$ y $\delta > 0$ para que se cumpla 2. Para $K = \frac{C}{\delta}$, por la desigualdad de Markov, tenemos que

$$P[|X| \geq K] \leq \frac{1}{K} E[|X|] \leq \delta,$$

para todo $X \in \chi$. Por lo tanto, por el inciso 2, tenemos que $E[|X|1_{\{|X| \geq K\}}] \leq \epsilon$ para toda $X \in \chi$ y así $\sup_{X \in \chi} E[|X|1_{\{|X| \geq K\}}] \leq \epsilon$ y finalmente

$$\lim_{K \rightarrow \infty} \left(\sup_{X \in \chi} E[|X|1_{\{|X| \geq K\}}] \right) \leq 0.$$

■

Teorema A.3.1. Convergencia Acotada Sea $\{X_n : n \geq 0\}$ una sucesión tal que $X_n \leq K < +\infty$ y $X_n \rightarrow_P X$ entonces

$$E[|X_n - X|] \rightarrow 0.$$

Demostración.

Notemos que $|X| - K \leq |X| - |X_n| \leq |X_n - X|$, entonces;

$$P(|X| \geq K + m^{-1}) \leq P(|X_n - X| \geq m^{-1}),$$

por la convergencia en probabilidad se sigue que $P(|X| \leq K) = 1$. Sea $\epsilon > 0$

$$\begin{aligned}
E[|X_n - X|] &= E[|X_n - X| \mid |X_n - X| > \epsilon/2] + E[|X_n - X| \mid |X_n - X| \leq \epsilon/2] \\
&\leq 2KP(|X_n - X| > \epsilon/2) + \epsilon/2 < \epsilon
\end{aligned} \tag{A-6}$$

para n suficientemente grande. ■

Teorema A.3.2. *Sea $\{X_n, n \geq 0\}$ una sucesión de variables aleatorias. Lo siguiente es equivalente:*

1. $X_n \in L^1$ para toda n , $X \in L^1$ y $X_n \rightarrow_{L^1} X$,
2. $\{X_n, n \geq 0\}$ es UI y $X_n \rightarrow_P X$.

Demostración.

1) \Rightarrow 2) Suponga que $X_n \rightarrow_{L^1} X$, sea $\epsilon > 0$, por la desigualdad de Chebyshev;

$$P(|X_n - X| > \epsilon) \leq \frac{E[|X_n - X|]}{\epsilon} \leq E[|X_n - X|]. \tag{A-7}$$

Por la convergencia en L^1 , tenemos que

$$\begin{aligned}
\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) &\leq \lim_{n \rightarrow \infty} E[|X_n - X|] = 0 \\
\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) &= 0.
\end{aligned} \tag{A-8}$$

Por lo tanto, $X_n \rightarrow_P X$. Además, dado $\epsilon > 0$ existe N tal que $E[|X_n - X|] \leq \epsilon/2$ cuando $n \geq N$. Podemos encontrar $\delta > 0$ tal que $P(A) \leq \delta$ implique:

$$E[|X|1_A] \leq \epsilon/2, \quad E[|X_n|1_A] \leq \epsilon, \quad n=1, \dots, N. \tag{A-9}$$

Entonces, para $n \geq N$ y $P(A) \leq \delta$,

$$E[|X_n|1_A] \leq E[|X_n - X|] + E[|X|1_A] \leq \epsilon.$$

Por lo tanto, por el Teorema A.3.2 tenemos que $\{X_n : n \geq 0\}$ es UI.

2) \Rightarrow 1) Existe una subsucesión (n_k) tal que $X_{n_k} \rightarrow X$ casi seguramente. Entonces, por el Lema de Fatou, $E[|X|] \leq \liminf_k E[|X_{n_k}|] < \infty$. Ahora, sea $\epsilon > 0$, existe $K < \infty$ tal que, para toda n ,

$$E[|X_n|1_{|X_n| \geq K}] < \epsilon/3, \quad E[|X|1_{|X| \geq K}] < \epsilon/3.$$

Consideremos la sucesión uniformemente acotada $X_n^K = (-K) \vee X_n \wedge K$ y el conjunto $X^K = (-K) \vee X \wedge K$. Entonces $X_n^K \rightarrow_P X^K$, entonces, por la convergencia acotada, existe N tal que, para $n \geq N$,

$$E[|X_n^K - X^K|] < \epsilon/3.$$

Por lo tanto, para toda $n \geq N$,

$$E[|X_n - X|] \leq E[|X_n|1_{|X_n| > k}] + E[|X_n^K - X^K|] + E[|X|1_{|X| > k}] < \epsilon.$$

Dado que $\epsilon > 0$ fue arbitrario,

$$\lim_{n \rightarrow \infty} E[|X_n - X|] = 0,$$

Por lo tanto, $X_n \rightarrow_{L^1} X$. ■

Apéndice B

Código

B.1. Minimizar CVaR de un portafolio con distribución Normal Multivariada

En esta subsección presentamos el algoritmo que se utilizó para resolver el ejemplo presentado en la sección 4.4, subsección 4.5.

```
library(MASS)
```

```
library(plot3D)
```

```
Min_CVar<-function(M,S,q,beta)
```

```
{
```

```
  En z guardaremos los valores que cada  
  iteracion nos dara.
```

```
  z=NULL
```

```
  Generamos la muestra de tamaño q de los activos.
```

```
  Y<-mvrnorm(q,M,S)
```

```
  Obtenemos el número de activos con el que estamos  
  trabajando.
```

```
m=length(M)
```

Generamos la matriz asociada al problema.

```
Q<-matrix(rep(0,((q+1)*(m+2*q+2))),
```

```
nrow = (q+1),ncol=(m+2*q+2))
```

Generamos los coeficientes.

```
c<-c(rep(0,m),1,rep(1/(q*(1-beta)),q),rep(0,q+1))
```

Llenamos la matriz con los valores de Y generados,

con los coeficientes de las variables U_k

y las variables de holgura V_k .

```
for(i in 1:q)
```

```
{
```

```
  Q[i,]=c(Y[i,],1,rep(0,(m+2*q+2-m-1)))
```

```
  Q[i,m+i+1]=1
```

```
  Q[i,m+q+i+1]=-1
```

```
}
```

La restriccion de la suma de $x_1+x_2+\dots+x_q=1$.

```
Q[(q+1),]=c(rep(1,m),rep(0,(m+2*q+2-m)))
```

Generamos el vector B.

```
b=rep(0,q+1)
```

```
b[q+1]=1
```

Como el sistema formado por las restricciones no

es canonico, sumamos la ultima restriccion

($x_1+x_2+\dots+x_m=1$) a todas las demas.

```
for (i in 1:q)
```

```

{
  Q[i,]=Q[i,]+Q[(q+1),]
  b[i]=b[i]+b[q+1]
}

```

Asignaremos una **variable** artificial para asi poder tener un sistema canonico.

```

Q[(q+1),m+2*q+2]=1
coltot=seq(1,length(Q[1,]))

```

Llamaremos a B, la matriz asociada a la base de nuestro sistema, cuyas variables seran u_1, \dots, u_q , a (**variable** artificial).

Tambien construimos el **vector de** coeficientes **de** la base.

```

col=NULL
B=matrix(rep(0,(q+1)*(q+1)),(q+1),(q+1))
cb=NULL
for(i in 1:q)
{
  B[,i]=Q[,m+1+i]
  cb=c(cb,c[m+1+i])
  col=c(col,c[m+1+i])
}
B[, (q+1)]=Q[,m+2*q+2]
col=c(col,c[m+2*q+2])

cb=c(cb,c[m+2*q+2])

```

Obtenemos la inversa **de** P.

```
B1=ginv(B)
```

Obtenemos una solucion basica asociada a P.

```
Xb=B1 %* %b
```

Obtenemos la solucion factible.

```
z=c(z,cb %* %Xb)
```

Calculos **de** optimalidad.

```
aux=cb %* %B1 %* %Q[,-col]-c[-col]
```

```
valoresx=rep(0,m)
```

```
while(length(which(aux>0))>0 )
```

```
{
```

```
  Seleccionamos las entradas del vector aux  
  que sean mayores que 0 y obtenemos el  
  maximo.
```

```
  mx=max(aux[which(aux>0)])
```

```
  entrada=which(aux==mx)
```

```
  if(length(entrada)>1)
```

```
  {
```

```
    entrada=entrada[1]
```

```
  }
```

```
  entrada=coltot[-col][entrada]
```

Calculos **de** factibilidad.

```
xaux<-B1 %* %Q[,entrada]
```

```
W=Xb/xaux
```

```
v=min(W[which(W>0)])
```

Obtenemos el cociente **de** los b y los coeficientes **de** la **variable de** entrada y la

variable de salida.

```
salida=which(W==v)
```

Modificamos la matriz **de** la base y del **vector de** coeficientes.

```
B[,salida]=Q[,entrada]
```

```
cb[salida]=c[entrada]
```

```
col[salida]=entrada
```

Obtenemos la inversa **de** B.

```
B1=ginv(B)
```

Modificamos Xb.

```
Xb=B1 %* %b
```

Obtenemos la solucion factible.

```
z=c(z,cb %* %Xb)
```

```
if(z[length(z)]>z[length(z)-1])
```

```
break
```

Calculos **de** optimalidad.

```

aux=cb %* %B1 %* %Q[,-col]-c[-col]

for(i in 1:m)
{
  if(length(which(col==i))!=0)
  {
    valores=c(valores,
              Xb[which(col==i)])
  }
  else{
    valores=c(valores,0)
  }
}
}

```

Guardamos el valor **de** las variables correspondientes a las ponderaciones **de** cada uno **de** los activos y su valor asociado, como tambien el CVaR optimo

(guardado en la **variable** 4).

```
variables=NULL
```

```
valores=NULL
```

```

for (i in 1:(m+1)) {
  if(length(which(col==i))!=0)
  {
    valores=c(valores ,col[which(col==i)])
    variables=c(variables ,Xb[which(col==i)])
  }
}

```

```

}

return(c(z[(length(z)-1):length(z)], valores, variables))
}

```

B.2. Ajuste de Series de tiempo y simulación de datos

En esta subsección presentaremos el código que utilizamos para el ajuste de los modelos de series de tiempo multivariadas.

RENDIMIENTOS

Cargamos la base que contiene los 9 rendimientos.

```

Rendimientos <- read_csv("Rendimientos.csv",
col_types = cols(Dates = col_skip()))

```

Guardamos los nombres de los rendimientos.

```

acciones<-names(Rendimientos)

```

Convertimos los datos en una serie de tiempo.

```

rendimientos<-ts(Rendimientos)

```

Quitamos la información de JAVER ya que la mayoría de sus rendimientos eran 0.

```

rendimientos<-rendimientos[,-c(6)]
acciones<-acciones[-c(6)]

```


Graficamos el comportamiento de los

datos en conjunto

```
par(mfrow=c(2,5))
for (i in 1:9) {
  if(i==9)
  {
    plot(rendimientos[,i],col="violetred",
         ylim=c(-0.4,0.2),xlim=c(1,450),ylab="",main=acciones[i])
  }
  else{
    plot(rendimientos[,i],col=i,
         ylim=c(-0.4,0.2),xlim=c(1,450),ylab="",main=acciones[i])
  }
}
```

Crearemos un portafolio con 3 acciones y

para su construccion nos basaremos en

la correlacion que existe entre ellas.

Para asi crear un portafolio diversificado.

```
par(mfrow=c(1,1))
corrplot(cor(rendimientos),method = "number")
```

Seleccionamos las acciones ,

BIMBO 1,

AMXL 4,

LIVERPOL 8.

```
rend=rendimientos[,c(1,4,8)]  
corrplot(cor(rend),method = "number")
```

Graficamos las series **de** tiempo.

```
MTSplot(rendimientos)  
MTSplot(rend)
```

VARselect nos indica el mejor ajuste **de** un modelo VAR que podemos crear segun la informacion AIC y BIC.

```
VARselect(rendimientos,lag.max = 5)  
VARselect(rend,lag.max = 5)
```

Ajuste **de** un modelo VAR(1) para el portafolio **de** 3 acciones.

```
model.1<-VAR(rend, p=1, type="const" ,  
season=NULL, exogen=NULL)
```

Ajustaremos el mismo modelo pero con otro comando pues VARMA nos genera la matriz **de** coeficientes y la matriz **de** correlaciones que utilizaremos para la simulacion.

```
model<-VARMA(rend,1,0,include.mean = F)
```

Parametros del primer modelo (3 acciones).

```
model$ARorder  
c1<-model$coef
```

```
c1
s1<-model$Sigma
s1
```

Ajuste **de** un modelo VAR(1) para el portafolio **de** 9 acciones.

```
model1.1<-VAR(rendimientos , p=1, type="const" ,
season=NULL, exogen=NULL)
```

Ajustaremos el mismo modelo pero con otro comando pues VARMA nos genera la matriz **de** coeficientes y la matriz **de** correlaciones que utilizaremos para la simulacion.

```
model1<-VARMA(rendimientos ,1,0 ,include.mean = F)
```

Parametros del segundo modelo (9 acciones).

```
model1$ARorder
model1$MAorder
c2<-model1$coef
c2
s2<-model1$Sigma
s2
```

Comprobar supuestos

1. Portmanteau-Test. (Correlacion entre los residuales)
Rend (3 activos).

```

model.1.serial<-serial.test(model.1, lags.pt=16,
  type = "BG")
model.1.serial
plot(model.1.serial)

```

Rendimientos (9 activos).

```

modell.1.serial<-serial.test(modell.1, lags.pt=16,
  type = "BG")
modell.1.serial
plot(modell.1.serial)

```

2. Testing heteroscedasticity (varianza constante)

Rend (3 activos).

```

model.1.arch<-arch.test(model.1, lags.multi=5,
  multivariate.only=TRUE)
model.1.arch

```

Rendimientos (9 activos).

```

modell.1.arch<-arch.test(modell.1, lags.multi=5,
  multivariate.only=TRUE)
modell.1.arch

```

3. Testing **for** normality (normalidad de los residuos)

#Rend (3 activos)

```

model.1.norm <- normality.test(model.1,
  multivariate.only=TRUE)
model.1.norm

```

```

#Rend (9 activos)
modell.1.norm <- normality.test(modell.1,
  multivariate.only=TRUE)
modell.1.norm

```

SIMULACION

Obtenemos el ultimo valor **de** las series **de** tiempo.

```

x_rend<-rend[length(rend[,1]),]
x_rendimientos<-rendimientos[length(rendimientos[,1]),]

```

Simulacion **de** n escenarios en un periodo.

Dado que los ajustes **de** series **de** tiempo que **se** realizaron corresponden a un VAR(1) para la simulacion necesitaremos el ultimo dato observado asi como la matriz coeficientes del polinomio **de** retraso **c**, la matriz **de** correlacion **s**.

```

sim_n<-function(ult_obs,n,c,s)
{
  A<-matrix(NA,nrow=n,ncol=length(ult_obs))
  for(i in 1:n)
  {
    A[i,]<-c %*% %ult_obs+
    mvrnorm(1,rep(0,length(ult_obs)),s)
  }
  return(A)
}

```

}

B.3. Optimización del CVaR

En esta subsección presentamos el algoritmo que se utilizó para la optimización del CVaR de un portafolio, cuando los escenarios fueron simulados una vez ajustado un modelo de series de tiempo multivariado.

Minimización del C-VaR con rendimientos antes simulados.

```
Min_CVar1<-function(A,beta)
```

```
{
```

```
  En z guardaremos los valores que
```

```
  cada iteración nos dará.
```

```
  z=NULL
```

```
  Asignamos a la variable Y, los escenarios.
```

```
  antes simulados
```

```
  Y<-A
```

```
  Obtenemos el número de escenarios simulados.
```

```
  q<-length(A[,1])
```

```
  El número de activos que estamos considerando.
```

```
  m=length(A[1,])
```

```
  Generamos la matriz asociada al problema.
```

```
  Q<-matrix(rep(0,((q+1)*(m+2*q+2))),
```

```
  nrow = (q+1),ncol=(m+2*q+2))
```

```
  Generamos los coeficientes.
```

```
  c<-c(rep(0,m),1,rep(1/(q*(1-beta)),q),rep(0,q+1))
```

```
  Llenamos la matriz con los valores de Y generados,
```

con los coeficientes **de** las variables U_k
y las variables **de** holgura V_k .

```
for(i in 1:q)
{
  Q[i,]=c(Y[i,],1,rep(0,(m+2*q+2-m-1)))
  Q[i,m+i+1]=1
  Q[i,m+q+i+1]=-1
}
```

La restriccion **de** la suma **de** $x_1+x_2+\dots+x_q=1$.

```
Q[(q+1),]=c(rep(1,m),rep(0,(m+2*q+2-m)))
```

Generamos el **vector** B .

```
b=rep(0,q+1)
```

```
b[q+1]=1
```

Como el sistema formado por las restricciones no es canonico, sumamos la ultima restriccion ($x_1+x_2+\dots+x_m=1$) a todas las demas.

```
for (i in 1:q)
{
  Q[i,]=Q[i,]+Q[(q+1),]
  b[i]=b[i]+b[q+1]
}
```

Asignaremos una **variable** artificial para asi poder tener un sistema canonico.

```
Q[(q+1),m+2*q+2]=1
```

```
coltot=seq(1,length(Q[1,]))
```

Llamaremos a B, la matriz asociada a la base de nuestro sistema, cuyas variables seran u_1, \dots, u_q , a (variable artificial).

Tambien construimos el vector de coeficientes de la base.

```
col=NULL
```

```
B=matrix(rep(0,(q+1)*(q+1)),(q+1),(q+1))
```

```
cb=NULL
```

```
for(i in 1:q)
```

```
{
```

```
  B[,i]=Q[,m+1+i]
```

```
  cb=c(cb,c[m+1+i])
```

```
  col=c(col,m+1+i)
```

```
}
```

```
B[,(q+1)]=Q[,m+2*q+2]
```

```
col=c(col,m+2*q+2)
```

```
cb=c(cb,c[m+2*q+2])
```

Obtenemos la inversa de P.

```
B1=ginv(B)
```

Obtenemos una solucion basica asociada a P.

```
Xb=B1 %* %b
```

Obtenemos la solucion factible.


```
z=c(z , cb %* %Xb)
```

Calculos **de** optimalidad .

```
aux=cb %* %B1 %* %Q[,-col]-c[-col]
```

```
valoresx=rep(0,m)
```

```
while(length(which(aux>0))>0 )
```

```
{
```

Seleccionamos las entradas del **vector** aux
que sean mayores que 0 y obtenemos el
maximo .

```
mx=max(aux[which(aux>0)])
```

```
entrada=which(aux==mx)
```

```
if(length(entrada)>1)
```

```
{
```

```
  entrada=entrada[1]
```

```
}
```

```
entrada=coltot[-col][entrada]
```

Calculos **de** factibilidad .

```
xaux<-B1 %* %Q[,entrada]
```

```
W=Xb/xaux
```

```
v=min(W[which(W>0)])
```

Obtenemos el cociente **de** los b y los coeficientes
de la **variable de** entrada y la **variable de** salida .

```
salida=which(W==v)
```

Modificamos la matriz **de** la base y del **vector**.

de coeficientes

```
B[, salida]=Q[, entrada]
```

```
cb[salida]=c[entrada]
```

```
col[salida]=entrada
```

Obtenemos la inversa **de** B.

```
B1=ginv(B)
```

Modificamos Xb.

```
Xb=B1 %* %b
```

Obtenemos la solución factible.

```
z=c(z, cb %* %Xb)
```

```
if(z[length(z)]>z[length(z)-1])
```

```
break
```

Calculos **de** optimalidad.

```
aux=cb %* %B1 %* %Q[,-col]-c[-col]
```

Guardamos el valor asociado a cada una **de**

las variables **de**

cada activo.

```
for(i in 1:m)
```

```
{
```

```
  if(length(which(col==i))!=0)
```

```
  {
```

```

        valoresx=c(valoresx ,
        Xb[which(col==i)])
    }
    else{
        valoresx=c(valoresx ,0)
    }
}
}
}

```

Guardamos el valor **de** las variables correspondientes a las ponderaciones **de** cada uno **de** los activos y su valor asociado, como tambien el CVaR optimo (guardado en la **variable** 4).

```

variables=NULL
valores=NULL
for (i in 1:(m+1)) {
    if (length(which(col==i)) !=0)
    {
        valores=c(valores ,col[which(col==i)])
        variables=c(variables ,Xb[which(col==i)])
    }
}
}

```

Para dar un mejor formato **de** salida , almacenamos los datos anteriores en un **data.frame**.

```

resvar<-c("z_opt" , variables)
resval<-c(z[length(z)] , valores)
res<-data.frame(var=resvar , val=resval ,

```

```

stringsAsFactors = F)
return(res)
}

```

B.4. Optimización en t periodos

En esta subsección presentamos el algoritmo que se utilizó para la optimización, en t periodos, del CVaR. Para ello, primero necesitamos que se generaran n escenarios iniciales para así, ir iterando y generar los datos con la dependencia de los datos inmediatos anteriores (VAR(1)).

Optimizacion del CVaR en t periodos ,
 con n escenarios en cada periodo
 a un nivel **beta**. Para poder correr
 esta funcion , necesitamos previamente
 generar n escenarios en $t=1$ con la funcion
 sim_n.

```

sim_nt<-function(nsim,t,c,s,beta)
{
  Obtenemos el numero de escenarios.
  n<-length(nsim[,1])
  Obtenemos el numero de activos.
  m<-length(nsim[1,])
  Optimizamos el CVaR en el primer
  periodo.
  optim<-Min_CVar1(nsim,beta)
  Guadramos las variables que arroja
  la optimizacion.
  variables<-optim$var

```

```

Guardamos los valores de las variables
antes mencionadas.
valores<-optim$val
Llevamos acabo el ajuste en los t-1
periodos restantes.
for (i in 2:t) {
  naux<-matrix(NA,nrow=n,ncol=m)
  for(j in 1:n)
  {
    Simulamos para cada escenario , los rendimientos
    del siguiente periodo
    naux[j ,]<-sim_n(nsim[j ,] ,1 ,c ,s)
  }
  nsim<-naux
  Optimizamos el CVaR para cada periodo.
  optim<-Min_CVar(nsim ,beta)
  Guardamos las variables y los valores que
  se obtienen en la optimizacion .
  variables<-c(variables ,optim$var)
  valores<-c(valores ,optim$val)
}
Regresamos la dinamica del comportamiento de las
politicas y del CVaR.
Res<-data.frame(var=variables ,val=valores ,
stringsAsFactors = F)
return(Res)
}

```

Grafica de la optimizacion

```
grafica_t<-function(optim,m,t,nom)
```

```
{
```

La optimizacion la guardamos en una matriz, en donde cada renglon contiene la politica optima que minimiza el CVAR en cada periodo (desde 1 hasta t).

```
M<-matrix(rep(0,t*m),nrow=t,ncol=m)
```

```
n<-1
```

```
cvar<-optim[which(optim[,1]==(m+1)),2]
```

```
for(i in 2:length(optim[,1]))
```

```
{
```

```
  if(optim[i,1]=="z_opt")
```

```
  {
```

```
    n=n+1
```

```
  }
```

```
  for(j in 1:m)
```

```
  {
```

```
    if(optim[i,1]==j)
```

```
    {
```

```
      M[n,j]=optim[i,2]
```

```
    }
```

```
  }
```

```
}
```

```
par(mfrow=c(2,floor(t/2)))
```

Graficamos la dinamica de las politicas en cada tiempo.

```
for(i in 1:t)
```

```
{
```

```

    pie (M[i , ] , col=2:(m+1)
        , labels=nom, radius=2,
        main=paste ( " t=" , i , "CVaR" ,
        signif ( cvar [ i ] )))
    }
}

```

En este último apartado mostraremos como se corrieron cada una de las funciones antes mostradas.

Para las primeras 3 acciones

```

beta=0.95
t<-10
nom3<-c ( "BIMBOA" , "AMXL" , "LIVERPOLC" )

```

```

simulaciones<-sim_n ( x_rend , 100 , c1 , s1 )
simulaciones

```

```

beta=0.95
t<-10
optimizaciont<-sim_nt ( simulaciones , t , c1 , s1 , beta )
optimizaciont
grafica_t ( optimizaciont , 3 , t , nom3 )

```

Para 9 acciones

```

beta=0.95
t1<-10
simulaciones_1<-sim_n ( x_rendimientos , 100 , c2 , s2 )
optimizaciont_1<-sim_nt ( simulaciones_1 , t1 , c2 , s2 , beta )

```

```
grafica_t(optimizaciont_1,9,t1,acciones)
```


Bibliografía

- [1] ARTZNER, P. DELBAEN, F. EBER, J.-M. and HEATH, D. (1998) Coherent measure of risk. *Mathematical Finance* **9**, 203-228.
- [2] BÄUERLE, N. and MÜLLER, A.(2005) Stochastic Orders and Risk Measures: Consistency and Bounds, 1-16.
- [3] ELTON, J. and HILL, T. P. (1992) Fusions of a probability distribution. *The Annals of Probability* **20**, 421-454.
- [4] EMBRECHTS, P. and WANG, R. (2005) Seven Proofs for the Subadditivity of Expected Shortfall, 8-9.
- [5] FÖLMEER, H. and SCHIED, A. (2002) Convex measures of risk and trading constraints. *Finance Stoch* **6**, 429-447.
- [6] HAUGH, M.(2010) Risk Measures, Risk Aggregation and Capital Allocation. *Springer*, 1-2.
- [7] KUSUOKA, S. (2001) On Law Invariant Coherent Risk Measure. *Adv. Math. Econ.* **3**, 158-168.
- [8] PFAFF, B. (2008) Analysis of Integrated and Cointegrated Time Series with R. *Springer* **1**.
- [9] ROCKAFELLAR, R. and URYASEV, S.(2000) Optimization of conditional value-at-risk. *Journal of Risk*, 21-42.

- [10] STRASSEN. V. (1965) The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 423-439.
- [11] TSAY, R. (2005) Analysis of Financial Time Series. *Wiley* **2**.
- [12] TSAY, R. (2014) Multivariate Time Series Analysis. *Wiley*.