

Medidas Dinámicas de Riesgo y Selección de Portafolios

Trabajo de Tesis
Presentado al
Departamento de Matemáticas

por

Nicolás Corredor Matiz

Para optar al Título de
Maestro en Matemáticas

Matemáticas
Universidad de los Andes
February 2007

Medidas Dinámicas de Riesgo y Selección de Portafolios

Aprobado por:

Álvaro Riascos y Diego Escobar, Asesor

Fecha de Aprobación _____

TABLA DE CONTENIDO

LISTA DE TABLAS	v
LISTA DE FIGURAS	vi
I. INTRODUCCIÓN	1
II. MEDIDAS ESTÁTICAS DE RIESGO	6
2.1. Medidas de riesgo	6
2.2. Medidas coherentes	10
2.3. Value at Risk (VaR)	11
2.4. Conditional Value at Risk (CVaR)	15
2.5. Caracterización de las medidas coherentes	20
III. MEDIDAS DINÁMICAS DE RIESGO	22
3.1. Introducción	22
3.2. Propiedades deseables	26
3.2.1. Independencia del pasado	26
3.2.2. Monotonidad y subaditividad dinámicas	28
3.2.3. Invarianza dinámica bajo traslación	29
3.2.4. Homogeneidad positiva dinámica	33
3.2.5. Medidas coherentes dinámicas	35
3.2.6. Consistencia temporal	36
3.2.7. Monotonidad temporal	36
3.3. Ejemplos	38
3.3.1. Extensiones triviales de medidas dinámicas	38

3.3.2.	Medida de riesgo iterada	38
3.3.3.	Construcción de medidas dinámicas de riesgo	39
IV.	OPTIMIZACIÓN DE PORTAFOLIOS	47
4.1.	Optimización con restricciones de CVaR	47
4.1.1.	Enfoque	47
4.1.2.	Ejemplos	54
4.2.	Optimización dinámica de portafolios	56
4.3.	Aplicación	59
4.3.1.	Definición del problema	59
4.3.2.	Primer enfoque	61
4.3.3.	Segundo enfoque	65
4.3.4.	Primeros resultados	66
4.3.5.	Segundo enfoque completado	70
4.3.6.	Pérdidas esperadas, realizadas y comparación de las medidas	73
V.	CONCLUSIONES	77
Apéndice A.	— ESPERANZA CONDICIONAL	80
REFERENCIAS		83

LISTA DE TABLAS

1.	Precios Opciones TAPO call	62
2.	Precios Opciones TAPO Put	62
3.	Resultados Segundo Enfoque Completado	72
4.	Pérdidas esperadas (Esp) y realizadas (Real)	74

LISTA DE FIGURAS

1.	2%-cuantiles para una variable discreta	12
2.	VaR para una variable continua	13
3.	CVaR para una variable continua	20
4.	Promedios mensuales de cobre desde el 1 de enero de 91	60
5.	Promedios mensuales de aluminio desde el 1 de enero de 91	61
6.	Modelo GARCH para el cobre	67
7.	Modelo GARCH para el aluminio	68
8.	Promedios mensuales de cobre realizados	75
9.	Promedios mensuales de aluminio realizados	76

Capítulo I

INTRODUCCIÓN

En el presente trabajo se estudian dos temas relacionados y de mucha importancia para las finanzas. Estos son: medidas dinámicas de riesgo y optimización de portafolios.

En la primera parte de este trabajo (Capítulos 2 y 3) se presenta una propuesta sobre la definición y propiedades que deben cumplir las medidas dinámicas de riesgo y se presentan ejemplos de diferentes medidas con diferentes propiedades. Este es un tema importante, ya que, aunque las medidas estáticas de riesgo y sus propiedades deseables han sido ampliamente estudiadas y caracterizadas en trabajos como el de Artzner et al. (1999), se ha observado la necesidad de extender estos conceptos a un contexto dinámico, para así tener medidas que se ajusten más a la realidad de los inversionistas, empresas, especuladores, etc., para los cuales el riesgo evoluciona en el tiempo debido a nueva información que pueda aparecer, la posibilidad de acciones o flujos intermedios y el posible cambio en el tiempo de la distribución de las pérdidas aleatorias a las cuales están expuestos.

Específicamente, la propuesta sobre la definición y propiedades que deben cumplir las medidas dinámicas de riesgo se hace recopilando, analizando y criticando la definición y propiedades deseables que proponen diferentes autores y proponiendo cambios en algunas estas propiedades de la siguiente manera.

La discusión sobre las propiedades de las medidas dinámicas de riesgo está estrechamente relacionada con la teoría de las medidas estáticas de riesgo, ya que la primera es una extensión de la segunda. Por lo tanto, en el Capítulo 2 se exponen diferentes temas fundamentales de las medidas estáticas de riesgo basado en

McNeil y Embrechts (2004), Artzner et al. (1999), Delbaen (2000), Delbaen (2002) y RiskmetricsTM (1996).

Primero, se define una medida de riesgo en el contexto estático y se expone el concepto de medida *coherente*. Entre otras cosas, una medida coherente asegura que la diversificación no aumente el riesgo, que el riesgo de pagar una cantidad fija de dinero sea igual a dicha cantidad, que el riesgo sea invariante ante un cambio de unidad de conteo y que si una posición generará no menos pérdidas que otra en cualquier caso, entonces el riesgo de la primera posición no sea menor que el de la segunda.

A continuación se define y analiza una medida que es bastante utilizada en la práctica pero que no es coherente en general, llamada “*Value at Risk*” (VaR_α). Se presenta entonces una medida relacionada, que sí es coherente, llamada “*Conditional Value at Risk*” (CVaR_α). Para finalizar el Capítulo 2 se exponen unos interesantes teoremas de caracterización de las medidas estáticas coherentes de riesgo.

En el Capítulo 3 se presentan diferentes temas de la teoría de medidas dinámicas de riesgo, basado principalmente en Artzner et al. (2003), Artzner et al. (2004), Balbás et al. (2002), Cheridito et al. (2004), Detlefsen y Scandolo (2005) y Riedel (2004). Todos estos autores extienden el concepto de medida coherente estática a las medidas dinámicas y también proponen propiedades deseables relacionadas con la evolución de la medida en el tiempo. Por esto se afirma que esta teoría es una extensión de la teoría de medidas de riesgo en el contexto estático. Entonces se busca que las definiciones y propiedades de las medidas dinámicas de riesgo sean una extensión natural de dicha teoría, de manera similar a como se extiende el concepto de variables aleatorias a procesos estocásticos.

Se tiene también que las extensiones de estos autores de la definición de medida de riesgo estática a medida de riesgo dinámica y las propiedades que ésta debe cumplir no son iguales entre ellas. Se presenta entonces una definición de medidas dinámicas de riesgo que busca capturar las ideas más importante de los autores, sin ser mucho más complicada que la definición de medida estática. Similarmente, se presentan las propiedades que debe cumplir una medida coherente de riesgo, discutiendo las propuestas de los diferentes autores estudiados y se propone que una

medida dinámica debe cumplir con unas propiedades más fuertes que las presentadas por todos éstos para encontrar las necesidades de los inversionistas. Por esto se propone el concepto de medidas *fuertemente coherentes*. Esto se hace porque una medida dinámica coherente que no sea fuertemente coherente puede subestimar el riesgo, sobre todo en fechas intermedias. En particular, en una inversión sucede que una medida así, al subestimar pérdidas potencialmente altas en las fechas intermedias, puede generar que la posición deba ser cerrada por el inversor antes de la fecha final (ver Sección 3.2), lo cual debe buscar evitar una buena medida dinámica de riesgo.

Para finalizar el capítulo se presentan diferentes medidas dinámicas de riesgo que son extensiones de medidas coherente estáticas. Solamente una de éstas (llamada $\psi^{\text{máx}}$) es fuertemente coherente. Dicha medida tiene como característica que es bastante fácil de entender, ya que es simplemente el máximo valor de una medida coherente de riesgo que mide en cada fecha intermedia hasta la fecha final las pérdidas acumuladas, además cumple con las extensiones directas realizadas por la mayoría de los autores estudiados del concepto de coherencia en el contexto estático y cumple con otras propiedades que hacen que no se subestimen los riesgos en las fechas intermedias. Esta medida, junto con otra medida dinámica (coherente, pero no fuertemente) serán utilizadas en la segunda parte de este trabajo como herramientas para la selección de portafolios, como se describe más adelante. La otra medida utilizada es $\bar{\psi}$, que es un promedio ponderado de los valores de la misma medida coherente estática que define $\psi^{\text{máx}}$, calculado en las diferentes fechas.

En la segunda parte de este trabajo (Capítulo 4) se presenta una propuesta para la selección de portafolios de manera dinámica, utilizando medidas dinámicas de riesgo que han sido presentadas y discutidas en la primera parte. Este tema es también bastante estudiado e importante en esta época, ya que desde el trabajo sobre selección de portafolios de Markowitz (1952) se han hecho propuestas para mejorar las metodologías de selección utilizando medidas de riesgo con mejores propiedades que la varianza del portafolio, pero se ha visto también la necesidad de utilizar medidas dinámicas, ya que los portafolios presentan en general flujos de caja intermedios y la distribución de sus pérdidas cambia en el tiempo. Por esto, incorporar medidas

dinámicas de riesgo es una necesidad que tienen los inversores.

De manera similar a lo que se expone en la primera parte, la propuesta hecha para la segunda parte, es decir, la selección de portafolios, es una extensión al contexto dinámico de un enfoque del contexto estático. Por esto, en la Sección 4.1 se expone de manera general un enfoque para optimizar o cubrir¹ un portafolio en el contexto estático. En este enfoque se basa la propuesta de este trabajo y su exposición se basa en los trabajos de Uryasev (2000), Andersson et al. (2001) y Rockafellar y Uryasev (2000 y 2002).

En este enfoque se busca minimizar el CVaR_α del portafolio, lo que también reduce en general el VaR_α del mismo. Entre las ventajas de hacer esto está que la medida de riesgo CVaR_α es una medida estática coherente de riesgo, a diferencia de VaR_α . También se tiene que el enfoque propuesto encuentra el portafolio con mínimo CVaR_α y simultáneamente calcula el VaR_α del mismo. Además, si se combina con métodos basados en escenarios, se puede optimizar el portafolio solucionado un problema de programación lineal.

Basado en esto se presenta una metodología que minimiza el valor de la medida $\psi^{\text{máx}}$ del portafolio en el caso en que $\psi^{\text{máx}}$ es el máximo CVaR_α de las pérdidas acumuladas en las fechas intermedias hasta la final. La medida es llamada MaxCVaR_α . Se presenta una aplicación de esta metodología de donde se observa que minimizar esta medida no es óptimo desde el punto de vista de selección de portafolios, ya que al minimizar el MaxCVaR_α se obtienen en general muchas soluciones posibles y entre estas soluciones hay unas que son mejores, porque se puede también minimizar el CVaR_α de otras fechas en las que éste no sea máximo.

Por esto se utiliza la medida $\bar{\psi}$ en una segunda etapa de optimización para minimizar entre las soluciones que minimizan MaxCVaR_α (en el caso en que la medida estática coherente que define a $\bar{\psi}$ es CVaR_α , por lo que esta medida es llamada $\overline{\text{CVaR}_\alpha}$) y así obtener un portafolio con un CVaR_α máximo que ha sido minimizado y con CVaR_α bajos en las demás fechas. No se minimiza simplemente $\overline{\text{CVaR}_\alpha}$ porque, aunque esta es una medida coherente, esto puede generar soluciones

¹De la palabra en inglés “*hedge*”

con CVaR_α muy altos en algunas fechas, como se muestra en la aplicación.

Lo anterior se hace utilizando la metodología propuesta en la Sección 4.1 y extendiéndola al contexto dinámico, logrando que el problema a solucionar no sea muy complicado. Además, los resultados obtenidos en la aplicación muestran que puede ser un enfoque muy atractivo para la selección de portafolios, ya que se obtienen retornos esperados más altos que con el enfoque de la Sección 4.1, sobre todo en las fechas intermedias, con un riesgo no mucho mayor en la fecha final, pero con riesgos menores para algunas fechas intermedias.

Esto muestra además la importancia de analizar los retornos esperados y los riesgos en fechas intermedias, tanto como en la final. Se evita así que flujos intermedios obliguen al inversor a cerrar la posición antes de la fecha final y se logra que el mismo tenga retornos esperados altos desde fechas intermedias. El enfoque propuesto es una metodología para lograr esto, a diferencia del enfoque estudiado, que solamente considera los retornos y riesgos de los flujos de caja acumulados en la fecha final.

Por último, la aplicación es un ejemplo de por qué CVaR_α es una mejor medida del riesgo que VaR_α , ya que esta segunda puede subestimar el riesgo al no medir la severidad de las pérdidas en los peores casos. Además, muestra que la única de las medidas utilizadas capaz de observar riesgos intermedios altos es MaxCVaR_α , ya que se observa que VaR_α , CVaR_α y $\overline{\text{CVaR}}_\alpha$ pueden ignorar riesgos en fechas intermedias, que generen pérdidas (con flujo de caja) intermedias altas. Esto justifica a su vez el por qué el hecho que una medida sea fuertemente coherente es apropiado, ya que por ejemplo, $\overline{\text{CVaR}}_\alpha$ es coherente pero no fuertemente coherente.

Capítulo II

MEDIDAS ESTÁTICAS DE RIESGO

2.1. Medidas de riesgo

Para definir y estudiar las propiedades deseables de las medidas de riesgo se sigue el enfoque de Delbaen (2002) y McNeil, et al. (2004). Se modelan los riesgos como variables aleatorias que representan las pérdidas generadas por variables financieras riesgosas como los precios de acciones, de “*commodities*”, tasas de cambio o interés, etc. Entonces, si la variable tiene como resultado un valor positivo, éste representa una pérdida y un valor negativo representa una ganancia.

Sea Ω un conjunto no vacío y sea \mathcal{F} una σ -álgebra sobre Ω .

Definición 2.1.1. Una variable aleatoria es una función $L : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B})$ medible, donde \mathcal{B} es la σ -álgebra de Borel.

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Para cualquier medida de probabilidad \mathbb{Q} se notará el valor esperado con respecto a ésta por $\mathbb{E}_{\mathbb{Q}}[\cdot]$. En particular se notará $\mathbb{E}[\cdot] \triangleq \mathbb{E}_{\mathbb{P}}[\cdot]$.

Se define el espacio L^0 como el espacio de todas las variables aleatorias (v.a's) sobre (Ω, \mathcal{F}) . El espacio L^0 es dotado de la topología de convergencia en probabilidad. Esto es,

$$L_n \xrightarrow{\mathbb{P}} L \quad \text{ssi} \quad \forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}[|L_n - L| > \epsilon] = 0.$$

También se define por L^∞ el conjunto de las v.a's sobre (Ω, \mathcal{F}) que son acotadas

bajo la norma,¹

$$\|L\|_\infty = \text{ess sup } |L| \triangleq \inf\{c \in \mathbb{R} \mid \mathbb{P}[|L| > c] = 0\}.$$

Esto es, $L^\infty \triangleq \{L \in L^0 \mid \|L\|_\infty < \infty\} = \{L \in L^0 \mid \exists c \in \mathbb{R} : \mathbb{P}(|L| > c) = 0\}$.

Las desigualdades entre variables aleatorias son en el sentido “*casi siempre*” (c.s.). Esto es, $X, Y \in L^0$ son tales que $X \leq Y$ c.s si $X_t(\omega) \leq Y_t(\omega)$ en casi todas partes, es decir, para todo $\omega \in A$, en algún $A \subseteq \Omega$, tal que $\mathbb{P}[A^c] = 0$. Se identifican además las variables aleatorias X y Y si $X = Y$ en casi todas partes, es decir, si $X \leq Y$ y $Y \leq X$ c.s.

Es conveniente trabajar en estos espacios, ya que en finanzas es frecuente reemplazar la medida de probabilidad \mathbb{P} por una medida equivalente. Dos medidas de probabilidad \mathbb{P} y \mathbb{Q} son equivalentes ($\mathbb{P} \equiv \mathbb{Q}$), si $\mathbb{P} \ll \mathbb{Q}$ y $\mathbb{Q} \ll \mathbb{P}$. $\mathbb{P} \ll \mathbb{Q}$ significa que para todo $A \subseteq \Omega$, $\mathbb{Q}(A) = 0 \Rightarrow \mathbb{P}(A) = 0$ y se dice que \mathbb{P} es absolutamente continua con respecto a \mathbb{Q} . Evidentemente L^0 no depende de la medida \mathbb{P} y

$$L^\infty = \{L \in L^0 \mid \exists c \in \mathbb{R} : \mathbb{P}(|L| > c) = 0\} = \{L \in L^0 \mid \exists c \in \mathbb{R} : \mathbb{Q}(|L| > c) = 0\},$$

si $\mathbb{P} \equiv \mathbb{Q}$.

Sea $\mathcal{M} \subseteq L^\infty$, llamado el conjunto de riesgos financieros, constituido por las variables aleatorias $L \in L^\infty$ para las existe algún portafolio tenido durante el periodo $[0, T]$, tal que L representa los pérdidas aleatorias realizadas en la fecha T del mismo. Comúnmente se asume que \mathcal{M} es un *cono convexo*, es decir, para todo $L, L' \in \mathcal{M}$ y $d > 0$ se tiene que $dL \in \mathcal{M}$ y $L + L' \in \mathcal{M}$.²

Se define también el espacio $L^1(\Omega, \mathcal{F}, \mathbb{P})$ o L^1 si no hay confusión, como el conjunto de las variables aleatorias integrables en $(\Omega, \mathcal{F}, \mathbb{P})$. Esto es, las $L \in L^0$ tales que

$$\mathbb{E}[|L|] = \int_\Omega |L| d\mathbb{P} < \infty.$$

Sea r la tasa de retorno libre de riesgo (o de referencia) compuesta continuamente. A continuación se presentan diferentes ejemplos de riesgos financieros (ver McNeil et al., 2004).

¹El símbolo “ess sup” se utiliza para denotar el supremo esencial de una variable aleatoria

² \mathcal{M} es un *cono convexo* porque los portafolios compuestos por combinaciones lineales de elementos de \mathcal{M} deben también pertenecer a \mathcal{M}

Ejemplo 2.1 (Portafolio de acciones). *Suponga que existen m acciones disponibles para invertir en ellas. Suponga que el precio de la acción i -ésima (llámela S^i) en el momento $t = 0, 1, \dots, T$ es P_t^i . Suponga que R_t^i es el retorno de S^i en el tiempo t y $R_t^i[k]$ es el retorno k periodos adelante, entonces*

$$1 + R_t^i = \frac{P_t^i}{P_{t-1}^i}, \quad 1 + R_t^i[k] = \frac{P_{t+k-1}^i}{P_{t-1}^i} = (1 + R_{t+k-1}^i)(1 + R_{t+k-2}^i) \dots (1 + R_t^i).$$

Por otro lado, modelar los retornos logarítmicos $r_t^i \triangleq \ln(1 + R_t^i) = \ln(P_t^i) - \ln(P_{t-1}^i)$ tiene las ventajas que un aumento en δ en el precio P_t^i , seguido por una disminución en el precio de δ genera retornos en valor absoluto iguales pero con signos opuestos. Entonces, los retornos son simétricos y además

$$r_t^i[k] = \ln(1 + R_t^i[k]) = \ln\left(\frac{P_{t+k-1}^i}{P_{t-1}^i}\right) = r_{t+k-1}^i + r_{t+k-2}^i + \dots + r_t^i.$$

En general se asume que los retornos r_t^i se distribuyen como una normal para todo t (o con una distribución elíptica), por lo que la distribución de $r_t^i[k]$ se obtiene fácilmente. El valor de un portafolio de α_i acciones invertidas en S^i en t es $V_t \triangleq \sum_{i=1}^m \alpha_i P_t^i$, por lo que las pérdidas en el periodo $[0, T]$ están dadas por

$$\mathbb{L}_T \triangleq -(V_T - V_0) = -\sum_{i=1}^m \alpha_i (P_T^i - P_0^i) = -\sum_{i=1}^m \alpha_i P_0^i (e^{r_T^i} - 1).$$

En general se utiliza la aproximación lineal de estas pérdidas dada por

$$l_T \triangleq -\sum_{i=1}^m \alpha_i P_0^i r_T^i.$$

Ejemplo 2.2 (“Forwards”). *Un “forward” es un acuerdo entre dos partes de vender o comprar una cantidad dada k de una variable financiera (moneda extranjera, acción, materias primas, etc.), en una fecha futura (por ejemplo T , asumiendo que la fecha actual es 0), a un valor dado. Estos acuerdos pueden ser meramente financieros o con entrega (entrega efectiva de la moneda extranjera, acciones o materia prima, etc.).*

Si es financiero el acuerdo y si el precio en la fecha T de la variable financiera menos el precio pactado en el acuerdo es positivo, el comprador del “forward” entrega

al vendedor del mismo una cantidad de dinero igual a este número positivo, de lo contrario, el vendedor entrega al comprador una cantidad de dinero igual a esta diferencia en valor absoluto. Es decir, para el comprador del “forward”, si el precio pactado fue F_0 y el precio real en T fue P_T ,

$$\mathbb{L}_T = F_0 - P_T.$$

En general, para asegurar que el comprador del “forward” pueda cubrir los costos del mismo cuando llegue la fecha de maduración T , se pide a éste que deposite en una cuenta de margen un dinero determinado en el contrato. Si el valor de la variable financiera baja, se puede pedir al comprador del “forward” que deposite en esta cuenta un dinero extra o de lo contrario se puede permitir a éste retirar un porcentaje del dinero de la cuenta. Entonces se pueden presentar flujos de caja en el periodo $[0, T)$, lo cual no se toma en cuenta en las medidas estáticas de riesgo.

Ejemplo 2.3 (Portafolio de Derivados). Suponga que la fecha actual es 0. Considere sobre el precio P_t en el tiempo t de una acción dada (puede ser moneda extranjera, materia prima, etc.), dos instrumentos derivados con fecha de ejercicio T . El primero (derivado A) paga al tenedor 1000 si P_T es mayor que un U dado y nada de lo contrario. El segundo (derivado B) paga al tenedor 1000 si P_T es menor que un L dado y nada de lo contrario. Suponga además $U > L$ y que el precio en 0 del derivado A es u y del derivado B es l . Las funciones de pérdidas de un vendedor del derivado A y B respectivamente son,

$$\mathbb{L}_T^A \triangleq \begin{cases} -u e^{rT}, & P_T \leq U \\ 1000 - u e^{rT}, & P_T > U \end{cases}, \quad \mathbb{L}_T^B \triangleq \begin{cases} -l e^{rT}, & P_T \geq L \\ 1000 - l e^{rT}, & P_T < L \end{cases}.$$

En vendedor de los derivado puede también hacer un portafolio vendiendo una opción A y una opción B . En este caso su función de pérdidas es

$$\mathbb{L}_T^{A+B} \triangleq \begin{cases} -(u + l) e^{rT}, & L \leq P_T \leq U \\ 1000 - (u + l) e^{rT}, & \text{de lo contrario} \end{cases}.$$

Se presenta a continuación la definición de una medida de riesgo en este contexto.

Definición 2.1.2. Una medida estática de riesgo es una función³ $\rho : \mathcal{M} \rightarrow \mathbb{R}$.

Se dice que una posición L es aceptable si $\rho(L) \leq 0$. Para una posición financiera con pérdidas dadas por L , $\rho(L) > 0$ es interpretado como la cantidad de dinero que se debe agregar a la posición en $t = 0$ para que ésta sea aceptable para un controlador de riesgo dado. Si $\rho(L) < 0$ se puede interpretar esto como, por ejemplo, la cantidad de dinero que se podría retirar de la cuenta de margen en el caso de una inversión en “forwards”.

2.2. Medidas coherentes

Se presentan a continuación los axiomas que debe cumplir una medida para que ésta sea llamada *coherente*.

Axioma 1 (Invarianza bajo traslación). Para todo $L \in \mathcal{M}$ y $l \in \mathbb{R}$ se tiene que

$$\rho(L + l e^{rT}) = \rho(L) + l.$$

Este axioma dice que si se le agrega a la posición una cantidad fija l en 0 (si $l > 0$, representa una pérdida o, en este caso, un préstamo y si $l < 0$, representa una ganancia) y se descuenta a la tasa libre de riesgo, entonces los requerimientos de capital se alteran exactamente en esa cantidad. En particular

$$\rho(L - \rho(L) e^{rT}) = \rho(L) - \rho(L) = 0.$$

Axioma 2 (Subaditividad). Para todo $L, L' \in \mathcal{M}$ se tiene que

$$\rho(L + L') \leq \rho(L) + \rho(L').$$

La subaditividad busca que la medida de riesgo refleje la idea que diversificar no aumenta el riesgo y, entre otras razones, busca evitar dar incentivos a las firmas para dividirse en subsidiarias más pequeñas para enmascarar el riesgo.

³Se trabaja sobre L^∞ porque, como lo demuestra Delbaen (2000), si \mathbb{P} no tiene átomos (es decir, $\forall \omega \in \Omega, \mathbb{P}(\omega) = 0$), sobre L^0 no existen medidas coherentes que tomen solo valores finitos (ver Definición 2.2.1)

Axioma 3 (Homogeneidad positiva). *Para todo $L \in \mathcal{M}$ y $d > 0$ se tiene que*

$$\rho(dL) = d\rho(L).$$

Se pide que la medida del riesgo sea invariante ante cambio de unidad de conteo. Además, junto con la subaditividad se tiene que el funcional ρ es *convexo* en \mathcal{M} .

Axioma 4 (Monotonicidad). *Para todo $L, L' \in \mathcal{M}$ con $L \leq L'$ c.s, se tiene que*

$$\rho(L) \leq \rho(L').$$

Este axioma es bastante natural y si una medida de riesgo satisface los primeros tres axiomas, éste es equivalente a pedir que $\rho(L) \leq 0$, si $L \leq 0$ c.s.

Definición 2.2.1. *Una medida de riesgo ρ cuyo dominio incluye \mathcal{M} es llamada coherente (sobre \mathcal{M}) si satisface los axiomas 1, 2, 3 y 4.*

Artzner, et al. (1999) también definen el axioma de *Relevancia* que es necesario (mas no suficiente) para prevenir que la concentración de riesgos no sea detectada. El axioma dice que para todo $L \in \mathcal{M}$, tal que $L \geq 0$ y $L \neq 0$ c.s, se tiene que $\rho(L) > 0$. Como una posición L es aceptable si $\rho(L) \leq 0$, entonces este axioma dice que si los posibles resultados de un riesgo tienen una probabilidad de generar ganancias igual a cero y no es exactamente igual a cero (la probabilidad de generar pérdidas es mayor que cero), entonces dicho riesgo financiero no debe ser aceptable.

2.3. Value at Risk (VaR)

Si $\alpha \in [0, 1]$ se llama a q un α -cuantil de la variable aleatoria X si

$$\mathbb{P}[X < q] \leq \alpha \leq \mathbb{P}[X \leq q].$$

El α -cuantil más grande y el más pequeño son,

$$\begin{aligned} q_\alpha(X) &= \inf\{x \mid \mathbb{P}[X \leq x] \geq \alpha\} = \inf\{x \mid \mathbb{P}[X > x] \leq 1 - \alpha\}, \\ q_\alpha^-(X) &= \inf\{x \mid \mathbb{P}[X \leq x] \geq \alpha\} = \inf\{x \mid \mathbb{P}[X > x] < 1 - \alpha\}. \end{aligned}$$

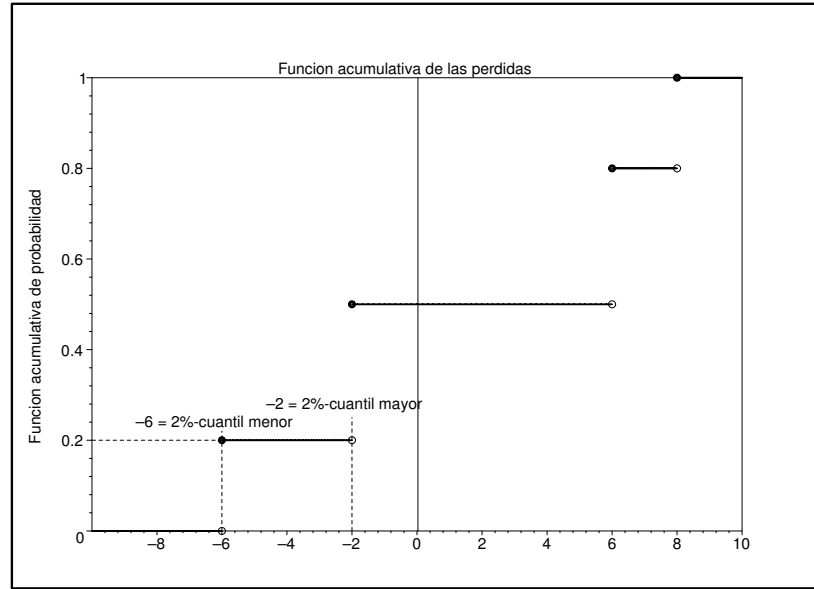


Figura 1: 2%-cuantiles para una variable discreta

Naturalmente se tiene que $q_{\alpha}^{-} \leq q \leq q_{\alpha}$ y q es un α -cuantil ssi $q_{\alpha}^{-} \leq q \leq q_{\alpha}$ (ver Figura 1).

Para $L \in \mathcal{M}$, el valor en riesgo o VaR de sus siglas en inglés (Value at Risk) se define como el α -cuantil menor de L , como lo hacen Balbás et al. (2002).

Definición 2.3.1 (VaR). Sea $\alpha \in (0, 1)$ un nivel de confianza dado. El VaR_{α} de un portafolio al nivel de confianza α es el menor $\lambda \in \mathbb{R}$, tal que la probabilidad que la pérdida $L e^{-rT}$ con $L \in \mathcal{M}$ exceda λ no es mayor que $1 - \alpha$. Esto es

$$\text{VaR}_{\alpha}(L) = q_{\alpha}^{-}(L e^{-rT}) = \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\}. \quad (1)$$

Para un α dado (por ejemplo 99% o 95%, ver Figura 2), se dice que la posición L es VaR_{α} -aceptable si $\text{VaR}_{\alpha}(L) \leq 0$. Entonces se puede pensar en VaR_{α} como la cantidad de capital extra que se debe añadir a una posición para que la probabilidad de quiebra se reduzca a $1 - \alpha$ (Delbaen, 2000). Esto es, como $\text{VaR}_{\alpha}(L - \text{VaR}_{\alpha}(L)) = \text{VaR}_{\alpha}(L) - \text{VaR}_{\alpha}(L) = 0$, la posición $L - \text{VaR}(L)$ es aceptable y con probabilidad no menor que $1 - \alpha$ es menor o igual que 0.

Proposición 2.3.1. VaR_{α} cumple los axiomas 1, 3 y 4.

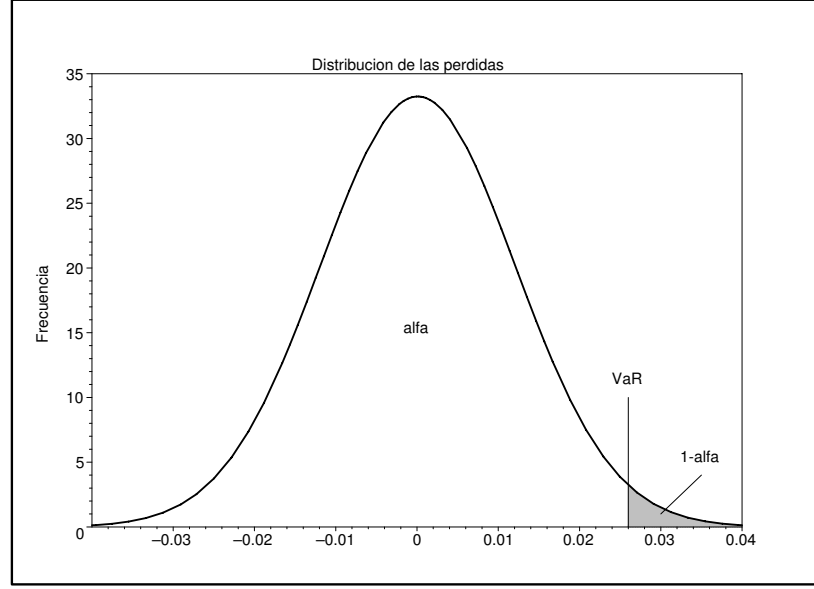


Figura 2: VaR para una variable continua

Demostración. En efecto, suponga que $L, L' \in \mathcal{M}$, $l \in \mathbb{R}$ y $d > 0$. Entonces

$$\begin{aligned} \text{VaR}_\alpha(L + l e^{rT}) &= \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} + l \leq \lambda] \geq \alpha\} \\ &= \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\} + l = \text{VaR}_\alpha(L) + l, \end{aligned}$$

porque c es tal que $\mathbb{P}[L e^{-rT} + l \leq c] \geq \alpha$ ssi $c' = c - l$ es tal que $\mathbb{P}[L e^{-rT} \leq c']$, luego $\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} + l \leq \lambda] \geq \alpha\} = \{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\} + \{l\}$. Además, $\inf(\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\} + \{l\}) = \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\} + l$.

Similarmente,

$$\begin{aligned} \text{VaR}_\alpha(dL) &= \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[dL e^{-rT} \leq \lambda] \geq \alpha\} \\ &= d \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\} = d \text{VaR}_\alpha(L), \end{aligned}$$

porque c es tal que $\mathbb{P}[dL e^{-rT} \leq c] \geq \alpha$ ssi $c' = c/d$ es tal que $\mathbb{P}[L e^{-rT} \leq c']$, luego $\{\lambda \in \mathbb{R} \mid \mathbb{P}[dL e^{-rT} \leq \lambda] \geq \alpha\} = d \{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\}$. Además, $\inf(d \{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\}) = d \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\}$.

Ahora, si $L \leq L'$ entonces $\{\omega \in \Omega \mid L(\omega) e^{-rT} \leq \lambda\} \supseteq \{\omega \in \Omega \mid L'(\omega) e^{-rT} \leq \lambda\}$. Luego $\mathbb{P}[L e^{-rT} \leq \lambda] \geq \mathbb{P}[L' e^{-rT} \leq \lambda]$, por lo que se sigue que

$$\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\} \supseteq \{\lambda \in \mathbb{R} \mid \mathbb{P}[L' e^{-rT} \leq \lambda] \geq \alpha\}$$

y entonces

$$\begin{aligned} \text{VaR}_\alpha(L) &= \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[L e^{-rT} \leq \lambda] \geq \alpha\} \\ &\leq \inf\{\lambda \in \mathbb{R} \mid \mathbb{P}[L' e^{-rT} \leq \lambda] \geq \alpha\} = \text{VaR}_\alpha(L'). \quad \square \end{aligned}$$

Sin embargo, VaR_α no cumple con el axioma de *subaditividad*, como se observa en el siguiente ejemplo, tomado de Artzner, et al. (1999).

Ejemplo 2.4. *Suponga que el tiempo actual es 0. Tome S_t y las dos opciones con fecha de ejercicio T , opciones A y B , como en el ejemplo 2.3. Suponga además que $\mathbb{P}[S_T < L] = \mathbb{P}[S_T > U] = 0,008$ y que $r = 0$. Entonces el vendedor de una opción A tiene una función de pérdidas,*

$$\mathbb{L}_T^A \triangleq \begin{cases} -u, & \text{con probabilidad } 0.992 \\ 1000 - u, & \text{con probabilidad } 0.008 \end{cases}.$$

Por lo tanto se tiene que, para L_A la v.a. que representa dichas pérdidas,

$$\text{VaR}_{0,99}(L_A) = -u.$$

Similarmente, para el vendedor de una opción B se tiene que $\text{VaR}_{0,99}(L_B) = -l$.

Por otro lado, para un vendedor de $A + B$ se tiene que la función de pérdidas L_{A+B} es tal que $\mathbb{L}_T^{A+B} = -u - l$ con probabilidad 0,984 y $\mathbb{L}_T^{A+B} = 1000 - u - l$ con probabilidad 0,016, así que en este caso

$$\text{VaR}_{0,99}(L_{A+B}) = 1000 - u - l.$$

Se sigue entonces que $\text{VaR}_{0,99}(L_{A+B}) > \text{VaR}_{0,99}(L_A) + \text{VaR}_{0,99}(L_B)$.

No obstante, si la función de pérdidas sigue una distribución normal o, más generalmente, una distribución elíptica, el VaR_α es una medida coherente de riesgo.

Además, aunque $L \geq 0$ implica $\text{VaR}_\alpha(L) \geq 0$, VaR_α no es *relevante*. Para ver esto es necesario únicamente considerar, para un $\alpha \in (0, 1)$ dado, L discreta, tal que $\mathbb{P}[L = 0] = \beta$ y $\mathbb{P}[L = 1] = 1 - \beta$ con $\beta \in (\alpha, 1)$. Entonces $\text{VaR}_\alpha(L) = 0$.

2.4. *Conditional Value at Risk* (CVaR)

Dado que VaR no es en general una medida coherente de riesgo, ésta puede ser manipulada para esconder riesgos importantes como se vio en el ejemplo 2.4; no mide la magnitud de la pérdida que se experimentará en los casos en que las pérdidas sean grandes, motivando así la venta de riesgos de poca frecuencia pero de alta severidad, Artzner, et al. (1999) proponen la medida “*Conditional Value at Risk*” (CVaR).⁴

Antes de definir CVaR se definirá una medida coherente de riesgo más general. Sea \mathcal{P} un conjunto arbitrario de medidas de probabilidad absolutamente continuas con respecto a \mathbb{P} . Si \mathbb{Q} es una medida de probabilidad y $\mathbb{Q} \ll \mathbb{P}$ entonces existe la derivada de Radon-Nikodým (ver Apéndice, Teorema A.0.1) de \mathbb{Q} respecto a \mathbb{P} , $\frac{d\mathbb{Q}}{d\mathbb{P}}$. Naturalmente, se tiene que $\frac{d\mathbb{Q}}{d\mathbb{P}} \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Luego, se puede identificar el conjunto \mathcal{P} con un subconjunto de L^1 , ya que $\frac{d\mathbb{Q}}{d\mathbb{P}}$ es única \mathbb{P} -casi siempre. Los elementos de \mathcal{P} son llamados escenarios generalizados o probabilidades de prueba sobre Ω .⁵

Definición 2.4.1. *Sea \mathcal{P} un conjunto arbitrario de escenarios generalizados. Se define la medida de riesgo*

$$\rho_{\mathcal{P}}(L) = \sup\{\mathbb{E}_{\mathbb{Q}}[L e^{-rT}] \mid \mathbb{Q} \in \mathcal{P}\}, \quad (2)$$

Proposición 2.4.1. *La medida de riesgo (2) es coherente.*

⁴Los autores la llaman “*tail conditional expectation*” y también conocida como “*expected short-fall*” (McNeil et al., 2004)

⁵Generalized scenarios o test probabilities, Artzner, et al. (2003)

Demostración. Para demostrar esto suponga $L, L' \in \mathcal{M}$, $l \in \mathbb{R}$, $d > 0$, entonces

$$\begin{aligned}
 \rho_{\mathcal{P}}(L + l e^{rT}) &= \sup\{\mathbb{E}_{\mathbb{Q}}[L e^{-rT} + l] \mid \mathbb{Q} \in \mathcal{P}\} = \sup\{\mathbb{E}_{\mathbb{Q}}[L e^{-rT}] + l \mid \mathbb{Q} \in \mathcal{P}\} \\
 &= \sup\{\mathbb{E}_{\mathbb{Q}}[L e^{-rT}] \mid \mathbb{Q} \in \mathcal{P}\} + l = \rho_{\mathcal{P}}(L) + l, \\
 \rho_{\mathcal{P}}(L + L') &= \sup\{\mathbb{E}_{\mathbb{Q}}[L e^{-rT} + L' e^{-rT}] \mid \mathbb{Q} \in \mathcal{P}\} \\
 &= \sup\{\mathbb{E}_{\mathbb{Q}}[L e^{-rT}] + \mathbb{E}_{\mathbb{Q}}[L' e^{-rT}] \mid \mathbb{Q} \in \mathcal{P}\} \\
 &\leq \sup\{\mathbb{E}_{\mathbb{Q}}[L e^{-rT}] \mid \mathbb{Q} \in \mathcal{P}\} + \sup\{\mathbb{E}_{\mathbb{Q}}[L' e^{-rT}] \mid \mathbb{Q} \in \mathcal{P}\} \\
 &= \rho_{\mathcal{P}}(L) + \rho_{\mathcal{P}}(L'), \\
 \rho_{\mathcal{P}}(dL) &= \sup\{\mathbb{E}_{\mathbb{Q}}[dL e^{-rT}] \mid \mathbb{Q} \in \mathcal{P}\} \\
 &= d \sup\{\mathbb{E}_{\mathbb{Q}}[L e^{-rT}] \mid \mathbb{Q} \in \mathcal{P}\} = d \rho_{\mathcal{P}}(dL)
 \end{aligned}$$

y si $L \leq L'$ c.s.,

$$\forall \mathbb{Q} \in \mathcal{P}, \mathbb{E}_{\mathbb{Q}}[L e^{-rT}] \leq \mathbb{E}_{\mathbb{Q}}[L' e^{-rT}] \Rightarrow \rho_{\mathcal{P}}(L) \leq \rho_{\mathcal{P}}(L'). \quad \square$$

Dado esto, para diferentes escenarios generalizados (diferentes conjuntos \mathcal{P}), se obtienen diferentes medidas coherentes de riesgo y entonces a continuación se presentan unos ejemplos.

Ejemplo 2.5. Sea $\mathcal{P} = \{\mathbb{Q} \mid \text{es una medida de probabilidad, } \mathbb{Q} \ll \mathbb{P}\}$, entonces se obtiene que $\rho_{\mathcal{P}}(L) = \text{ess sup}(L e^{-rT})$. Además, como $\text{ess sup } L \leq 0$ ssi $L \leq 0$ c.s., la medida es muy severa, dado que la familia \mathcal{P} es muy grande. A continuación se demostrarán estas afirmaciones.

En efecto, sea $L' = L e^{-rT}$. Si $c > \text{ess sup } L'$, sea c' tal que $\text{ess sup } L' < c' < c$, entonces $\forall \mathbb{Q} \in \mathcal{P}$,

$$\mathbb{E}_{\mathbb{Q}}[L] = \int_{\{L' \leq c'\}} L d\mathbb{Q} + \int_{\{L' > c'\}} L d\mathbb{Q} = \int_{\{L' \leq c'\}} L d\mathbb{Q} \leq c' \mathbb{Q}[L' \leq c'] \leq c' < c.$$

Por lo tanto $c > \text{ess sup } L' \Rightarrow \rho(L) < c$, luego $\rho(L) \leq \text{ess sup } L'$.

Ahora, si $c < \text{ess sup } L'$, como $\mathbb{P}(L' > c) > 0$, sea \mathbb{Q} dado por

$$\mathbb{Q}(A) = \int_{\Omega} \frac{\mathbb{I}_{(c, \infty)}}{\mathbb{P}[L' > c]} d\mathbb{P}.$$

Trivialmente $\mathbb{Q} \in \mathcal{P}$ y se tiene además que

$$\mathbb{E}_{\mathbb{Q}}[L] = \frac{1}{\mathbb{P}[L' > c]} \int_{\{L' > c\}} L' d\mathbb{P} \geq \frac{\mathbb{P}[L' > c]c}{\mathbb{P}[L' > c]} = c.$$

Luego $c < \text{ess sup } L' \Rightarrow \rho(L) \geq c$, por lo que se concluye que $\rho(L) \geq \text{ess sup } L'$ y entonces $\rho(L) = \text{ess sup } L'$.

Para demostrar que $\text{ess sup } L \leq 0$ ssi $L \leq 0$ c.s, simplemente observe que

$$L \leq 0 \text{ c.s} \Leftrightarrow \forall \epsilon > 0, \mathbb{P}[L > \epsilon] = 0 \Leftrightarrow \text{ess sup } L \leq 0.$$

Ejemplo 2.6. La medida de riesgo que se obtiene si se analiza la medida $\rho_{\mathcal{P}_k}$, donde

$$\mathcal{P}_k \triangleq \left\{ \mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \leq k \right\},$$

para $k > 1$ es muy importante en la práctica. Esta medida es conocida como “conditional value at risk”. Si $k \leq 1$, la medida que se obtiene es poco interesante, ya que $\mathcal{P}_1 = \{\mathbb{P}\}$. Esto es porque en este caso $\frac{d\mathbb{Q}}{d\mathbb{P}} = 1$, \mathbb{P} -c.s. Suponga $\frac{d\mathbb{Q}}{d\mathbb{P}} \leq 1$. Sea $A = \{\omega \mid \frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) < 1\}$, $A^c = \{\omega \mid \frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) = 1\}$, donde $\mathbb{P}(A) > 0$, entonces

$$\int_{\Omega} d\mathbb{Q} = \int_{\Omega} \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} = \int_A \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} + \int_{A^c} \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} < \mathbb{P}(A) + \mathbb{P}(A^c) = 1, \quad (3)$$

lo cual es una contradicción y se tiene que $\mathbb{P}(A) = 0$. Entonces $\frac{d\mathbb{Q}}{d\mathbb{P}} = 1$, \mathbb{P} -c.s. Ahora, $\frac{d\mathbb{Q}}{d\mathbb{P}} < 1$ no es posible porque se tendría la contradicción (3) con $A = \Omega$.

Luego para cualquier $A \subseteq \Omega$,

$$\mathbb{Q}(A) = \int_A \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} = \int_A 1 d\mathbb{P} = \mathbb{P}(A),$$

por lo que $\mathcal{P}_1 = \{\mathbb{P}\}$ y la medida $\rho_{\mathcal{P}_1}$ diría que L es aceptable si

$$\rho_{\mathcal{P}_1}(L) = \mathbb{E}[L e^{-rT}] \leq 0,$$

lo cual es demasiado tolerante.

Observe que en el caso en que $k < 1$ se tiene que $\mathcal{P}_k = \emptyset$, ya que también se cumple naturalmente que $\mathbb{P} \notin \mathcal{P}_k$ porque $\frac{d\mathbb{P}}{d\mathbb{P}} = 1$.

Por esto se considera el caso en que $k > 1$ y se tiene el siguiente resultado que será utilizado para definir “conditional value at risk” (Delbaen, 2002).

Teorema 2.4.1. Sea $k > 1$ y $\alpha = 1 - \frac{1}{k} \in (0, 1)$, entonces

$$\rho_{\mathcal{P}_k}(L) = \frac{1}{1 - \alpha} \left(\int_{\{L e^{-rT} > q_\alpha^-\}} L e^{-rT} d\mathbb{P} + (1 - \alpha - \mathbb{P}[L e^{-rT} > q_\alpha^-]) q_\alpha^- \right).$$

Demostración. Sea $q_\alpha^-(L e^{-rT}) = q_\alpha^-$ y $L' = L e^{-rT}$. Defina \mathbb{Q}_0 por

$$\mathbb{Q}_0(A) = \int_A k \mathbb{I}_{\{L' > q_\alpha^-\}} + k \beta \mathbb{I}_{\{L' = q_\alpha^-\}} d\mathbb{P},$$

con $\beta = \frac{1 - \alpha - \mathbb{P}[L' > q_\alpha^-]}{\mathbb{P}[L' = q_\alpha^-]}$ si $\mathbb{P}[L' = q_\alpha^-] \neq 0$ y $\beta = 0$ de lo contrario. La función \mathbb{I}_A es la función indicadora del conjunto A , esto es,

$$\mathbb{I}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A. \end{cases}$$

Como q_α^- es un α -cuantil se tiene que $\mathbb{P}[X < q_\alpha^-] \leq \alpha \leq \mathbb{P}[X \leq q_\alpha^-]$ o $\mathbb{P}[X \geq q_\alpha^-] \geq 1 - \alpha \geq \mathbb{P}[X > q_\alpha^-]$. Luego $1 - \alpha - \mathbb{P}[L' > q_\alpha^-] \geq 0$ y

$$1 - \alpha - \mathbb{P}[X > q_\alpha^-] = \mathbb{P}[X \leq q_\alpha^-] - \alpha \leq \mathbb{P}[X \leq q_\alpha^-] - \mathbb{P}[X < q_\alpha^-] = \mathbb{P}[X = q_\alpha^-],$$

por lo que $0 \leq \beta \leq 1$. Luego $0 \leq \frac{d\mathbb{Q}_0}{d\mathbb{P}} = k \mathbb{I}_{\{L' > q_\alpha^-\}} + k \beta \mathbb{I}_{\{L' = q_\alpha^-\}} \leq k$, y se sigue que $\mathbb{Q}_0 \in \mathcal{P}_k$. Ahora,

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}_0}[L'] &= \int_{\Omega} L' d\mathbb{Q}_0 = \int_{\Omega} L' \left(k \mathbb{I}_{\{L' > q_\alpha^-\}} + k \beta \mathbb{I}_{\{L' = q_\alpha^-\}} \right) d\mathbb{P} \\ &= k \int_{\Omega} L' \mathbb{I}_{\{L' > q_\alpha^-\}} d\mathbb{P} + k \beta \int_{\Omega} L' \mathbb{I}_{\{L' = q_\alpha^-\}} d\mathbb{P} \\ &= k \int_{\{L' > q_\alpha^-\}} L' d\mathbb{P} + k \beta q_\alpha^- \mathbb{P}[L' = q_\alpha^-] \\ &= \frac{1}{1 - \alpha} \left(\int_{\{L' > q_\alpha^-\}} L' d\mathbb{P} + (1 - \alpha - \mathbb{P}[L' > q_\alpha^-]) q_\alpha^- \right). \end{aligned}$$

Para ver que la igualdad se tiene, sea $\mathbb{Q} \in \mathcal{P}_k$ cualquiera. Entonces se tiene que,

para $A_1 = \{L' > q_\alpha^-\}$, $A_2 = \{L' = q_\alpha^-\}$ y $A_3 = \{L' < q_\alpha^-\}$,

$$\begin{aligned}
 \mathbb{E}_{\mathbb{Q}}[L'] &= \int_{\Omega} L' d\mathbb{Q} \\
 &= k \int_{A_1} L' d\mathbb{P} + k\beta \int_{A_2} L' d\mathbb{P} + \int_{A_1} L' \left(\frac{d\mathbb{Q}}{d\mathbb{P}} - k \right) d\mathbb{P} \\
 &\quad + \int_{A_2 \cup A_3} L' d\mathbb{Q} - k\beta \int_{A_2} L' d\mathbb{P} \\
 &\leq k \int_{A_1} L' d\mathbb{P} + k\beta \int_{A_2} L' d\mathbb{P} + \int_{A_1} q_\alpha^- \left(\frac{d\mathbb{Q}}{d\mathbb{P}} - k \right) d\mathbb{P} \\
 &\quad + \int_{A_2 \cup A_3} q_\alpha^- d\mathbb{Q} - k\beta \int_{A_2} q_\alpha^- d\mathbb{P} \\
 &= k \int_{A_1} L' d\mathbb{Q} + k\beta \int_{A_2} L' d\mathbb{P} + q_\alpha^- [\mathbb{Q}(\Omega) - k \{ \mathbb{P}(A_1) + \beta \mathbb{P}(A_2) \}] \\
 &= k \int_{A_1} L' d\mathbb{P} + k\beta \int_{A_2} L' d\mathbb{P} \\
 &= \mathbb{E}_{\mathbb{Q}_0}[L'],
 \end{aligned}$$

ya que $\mathbb{Q}(\Omega) - k \{ \mathbb{P}(A_1) + \beta \mathbb{P}(A_2) \} = 1 - k \{ \mathbb{P}[L > q_\alpha^-] + (1 - \alpha - \mathbb{P}[L > q_\alpha^-]) \} = 1 - k(1 - \alpha) = 0$. Entonces, $\rho_{\mathcal{P}_k}(L) = \sup\{\mathbb{E}_{\mathbb{P}}[L'] \mid \mathbb{P} \in \mathcal{P}\} = \mathbb{E}_{\mathbb{Q}_0}[L']$. \square

Corolario 2.4.1. Sea $k > 1$ y $\alpha = 1 - \frac{1}{k}$, si L es continua en $q_\alpha^-(L)$, entonces

$$\rho_{\mathcal{P}_k}(L) = \mathbb{E}_{\mathbb{P}}[L e^{-rT} \mid L e^{-rT} \geq q_\alpha^-(L e^{-rT})] \geq q_\alpha^-(L e^{-rT}) = \text{VaR}_\alpha(L).$$

Demostración. Si L es continua en q_α^- , entonces $\mathbb{P}[L' \leq q_\alpha^-] = \mathbb{P}[L' < q_\alpha^-] = \alpha$ y $\mathbb{P}[L' = q_\alpha^-] = 0$. Luego $\beta = 0$ y

$$\rho_{\mathcal{P}_k}(L) = \mathbb{E}_{\mathbb{Q}_0}[L'] = \frac{1}{1 - \alpha} \int_{\{L' \geq q_\alpha^-\}} L' d\mathbb{P} = \frac{\mathbb{E}_{\mathbb{P}}[L' \mathbb{I}_{\{L' \geq q_\alpha^-\}}]}{\mathbb{P}[L' \geq q_\alpha^-]} = \mathbb{E}_{\mathbb{P}}[L' \mid L' \geq q_\alpha^-]. \square$$

Definición 2.4.2 (CVaR). Dado un nivel de confianza $\alpha \in (0, 1)$, la medida de riesgo “conditional value at risk” (CVaR) está dada por,

$$\text{CVaR}_\alpha(L) = \frac{1}{1 - \alpha} \left(\int_{\{L e^{-rT} > q_\alpha^-\}} L e^{-rT} d\mathbb{P} + (1 - \alpha - \mathbb{P}[L e^{-rT} > q_\alpha^-]) q_\alpha^- \right). \quad (4)$$

En particular, en el caso que L sea continua,

$$\text{CVaR}_\alpha(L) = \mathbb{E}_{\mathbb{P}}[L e^{-rT} \mid L e^{-rT} \geq q_\alpha^-(L e^{-rT})]. \quad (5)$$

Observe que la medida CVaR es coherente por el Teorema 2.4.1 y el Corolario 2.4.1. Similarmente a como se hace con la medida VaR, se utilizan en general valores de α cercanos a 95% y 99% (ver Figura 3).

McNeil, et al. (2004) definen (4) como “generalized expected shortfall” y (5) como “expected shortfall”.

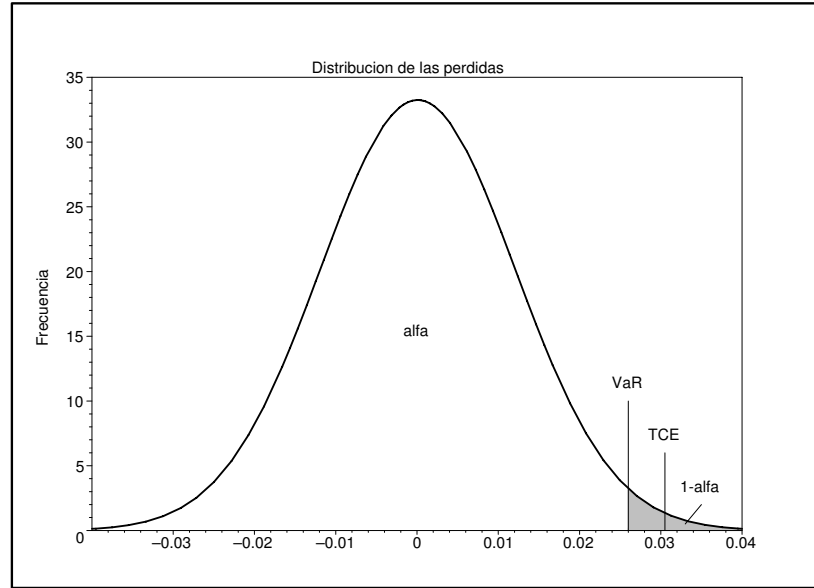


Figura 3: CVaR para una variable continua

2.5. Caracterización de las medidas coherentes

Se mostró que las medidas de riesgo construidas como se hizo en (2) son coherentes. Más aún, ésta es la única manera de construir medidas coherentes de riesgo, si la medida cumple también con una condición de continuidad que se presentará a continuación. Se presentan los siguientes resultados de caracterización por completitud y dado que son resultados muy interesantes, pero no se presentan las demostraciones (interesados ver Delbaen (2000)), ya que estos resultados no serán utilizados en este trabajo.

Primero, siguiendo la notación de Delbaen (2000), sea $\mathbf{ba}(\Omega, \mathcal{F}, \mathbb{P})$ (o $\mathbf{ba}(\mathbb{P})$ si no hay confusión) el espacio de las medidas tales que $\mu \in \mathbf{ba}(\mathbb{P})$ ssi μ es acotada,

finitamente aditiva en (Ω, \mathcal{F}) , y $\mathbb{P}(A) = 0$ implica $\mu(A) = 0$. Los elementos $\mu \in \mathbf{ba}(\Omega, \mathcal{F}, \mathbb{P})$ no negativos (como funciones, es decir, $\mu(\omega) \geq 0$ para todo $\omega \in \Omega$), con $\mu(\Omega) = 1$ también son llamados *probabilidades finitamente aditivas*.

Teorema 2.5.1. *Suponga que $\rho : L^\infty \rightarrow \mathbb{R}$ es coherente. Entonces existe un conjunto $\mathcal{P} \subseteq \mathbf{ba}(\mathbb{P})$ de probabilidades finitamente aditivas, convexo, cerrado en la topología $\sigma(\mathbf{ba}(\mathbb{P}), L^\infty(\mathbb{P}))$, tal que*

$$\rho(L) = \sup_{\mu \in \mathcal{P}} \mathbb{E}_\mu[L e^{-rT}].$$

El anterior teorema caracteriza las medidas coherentes de riesgo con conjuntos de probabilidades finitamente aditivas. Para obtener una caracterización con probabilidades σ -aditivas es necesario un supuesto adicional. Se presenta entonces la siguiente propiedad (Delbaen, 2000).

Axioma 5. *La medida coherente $\rho : L^\infty \rightarrow \mathbb{R}$ cumple con la propiedad de Fatou si para toda secuencia $(L_n)_{n \geq 1}$ de funciones uniformemente acotadas por 1 ($\|L_n\|_\infty \leq 1$), que convergen en probabilidad a L ($L_n \xrightarrow{\mathbb{P}} L$), se tiene que*

$$\rho(L) \leq \liminf \rho(L_n).$$

Entonces se tiene el siguiente teorema de caracterización.

Teorema 2.5.2. *Suponga que $\rho : L^\infty \rightarrow \mathbb{R}$ es coherente. Si ρ satisface la propiedad de Fatou, entonces existe un conjunto \mathcal{P} , de medidas de probabilidad absolutamente continuas con respecto a \mathbb{P} , tales que, para todo $L \in L^\infty$,*

$$\rho(L) = \sup_{\mathbb{Q} \in \mathcal{P}} \mathbb{E}_\mathbb{Q}[L e^{-rT}].$$

Capítulo III

MEDIDAS DINÁMICAS DE RIESGO

3.1. *Introducción*

Las anteriores medidas de riesgo son estáticas en el sentido que en el tiempo 0 se mide el riesgo de una posición analizando la función L que representa las pérdidas de ésta en el tiempo T y no se permite que cambien en el tiempo. No se tiene en cuenta ninguna información nueva que pueda surgir o cambios que puedan hacerse en la posición en el intervalo $(0, T)$. Más concretamente:

1. No se considera en el análisis las pérdidas potenciales o flujos de caja en momentos entre 0 y T . Por ejemplo, si se está midiendo el riesgo de la posición para evitar una posible quiebra, estas medidas de riesgo no considerarían la posibilidad que ocurran sucesos que lleven a la quiebra al dueño de la posición en un tiempo anterior a T . Tampoco considera la acumulación de flujos de caja que se pueden presentar, por ejemplo, en un portafolio de futuros.
2. No se consideran cambios en la posición en momentos intermedios. Por ejemplo, se puede tener un portafolio por seis meses en el que cada mes se harán modificaciones.
3. No se tiene en cuenta que el riesgo de la posición se debe medir en momentos intermedios.
4. Si se utiliza la medida de riesgo en un problema de optimización dinámico se debe trabajar preferiblemente con una medida dinámica (Wang, 2000).

Sea $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{N}}, \mathbb{P})$ un espacio *filtrado* de probabilidad, donde $(\mathcal{F}_t)_{t \in \mathbb{N}}$ es una secuencia de σ -álgebras $\mathcal{F}_0, \mathcal{F}_1, \dots$, tales que $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$, para todo $t \in \mathbb{N}$. Dicha secuencia se llama una *filtración*. Se supone que $\mathcal{F}_0 = \{\emptyset, \Omega\}$ y que cada \mathcal{F}_t es una sub- σ -álgebra de \mathcal{F} , que representa la información que se conoce hasta la fecha t .

Siguiendo la notación de los autores Cheridito, et al. (2004), sea \mathcal{R}^0 el conjunto de procesos estocásticos (\mathcal{F}_t) -adaptados sobre el espacio $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{N}}, \mathbb{P})$. Esto es, para $X = (X_t)_{t \in \mathbb{N}} \in \mathcal{R}^0$, cada X_t es \mathcal{F}_t -medible o $\sigma(X_t) \subseteq \mathcal{F}_t$.¹

Se define también el conjunto de los procesos estocásticos acotados, dado por

$$\mathcal{R}^\infty \triangleq \{X \in \mathcal{R}^0 \mid \|X\|_{\mathcal{R}^\infty} < \infty\},$$

donde

$$\|X\|_{\mathcal{R}^\infty} \triangleq \inf \left\{ m \in \mathbb{R} \mid \sup_{t \in \mathbb{N}} \|X_t\|_\infty \leq m \right\} = \inf \left\{ m \in \mathbb{R} \mid \sup_{t \in \mathbb{N}} \{\text{ess sup } |X_t|\} \leq m \right\}.$$

Las desigualdades entre procesos estocásticos son en el sentido “*casi siempre*” (c.s). Esto es, $X, Y \in \mathcal{R}^0$ son tales que $X \leq Y$ c.s si para todo $t \in \mathbb{N}$, $X_t(\omega) \leq Y_t(\omega)$ c.s. Se identifican los procesos estocásticos X y Y si $X = Y$ (es decir, si $X \leq Y$ c.s y $Y \leq X$ c.s).

Para modelar las pérdidas de una inversión de la fecha S hasta T , con S y T , tales que $0 \leq S < \infty$ y $S \leq T \leq \infty$, se utilizará la proyección $\pi_{S,T} : \mathcal{R}^0 \rightarrow \mathcal{R}^0$ definida por²

$$(\pi_{S,T}(X))_t \triangleq \mathbb{I}_{\{t \geq S\}} X_{t \wedge T}, \quad t \in \mathbb{N},$$

donde $a \wedge b \triangleq \min\{a, b\}$. Esto es, $(\pi_{S,T}(X))_t = (0, \dots, 0, X_S, X_{S+1}, \dots, X_{T-1}, X_T, X_T, \dots)$.

Se define entonces el conjunto

$$\mathcal{R}_{S,T}^\infty \triangleq \pi_{S,T}(\mathcal{R}^\infty) = \{\pi_{S,T}(X) \mid X = (X_t)_{t \in \mathbb{N}} \in \mathcal{R}^\infty\}.$$

Se define el subconjunto $\mathcal{M}_S^T \subseteq \mathcal{R}_{S,T}^\infty$ de los elementos de $\mathcal{R}_{S,T}^\infty$ que representan las *pérdidas acumuladas* de alguna posición desde la fecha S hasta T , donde se

¹ $\sigma(X_t)$ o la σ -álgebra generada por X_t , es la σ -álgebra más pequeña que hace a X_t medible, esto es, la familia de todos los conjuntos de la forma $X^{-1}(A) = \{\omega \in \Omega \mid X_t(\omega) \in A\}$, donde $A \subseteq \mathbb{R}$ es medible

²En este caso $\mathbb{I}_{\{t \geq S\}}$ es igual a 1 en las fechas t del conjunto $\{t \mid S \leq t\}$ y 0 de lo contrario

identifican los procesos que sean iguales c.s. Como se modelan pérdidas, se hablará de flujos de caja en un sentido diferente al tradicional, ya que un flujo de caja positivo es una pérdida y un flujo negativo es una ganancia. Además se asume que dichos flujos acumulados están descontados con respecto a la tasa de referencia r . Los elementos de estudio son los procesos estocásticos $L = (L_t)_{t \in \mathbb{N}} \in \mathcal{M}_S^T$. Si un proceso $(U_t)_{t \in \mathbb{N}} \in \mathcal{R}_{S,T}^\infty$ representa los flujos de caja de S a T de una posición a la que se le desea medir el riesgo, se estudia el proceso $L = (L_t) \in \mathcal{M}_S^T$, con $L_t = \sum_{s=0}^t U_s e^{r(t-s)}$, el cual a su vez mide los flujos de caja acumulados.

No se modelan las pérdida como variables aleatorias (acotadas), interpretadas como el pago final de la posición, como lo proponen Detlefsen y Scandolo (2005), ya que en el contexto de este trabajo se desea medir también en las fechas intermedias el valor de la misma, buscando que ésta sea aceptable durante todo el horizonte de inversión.

Se debe asumir que \mathcal{M}_S^T es un *cono convexo* por la misma razón que se \mathcal{M} se asume de esta manera. No obstante, se propone una condición más fuerte. Esto es porque la composición de los portafolios puede cambiar en el tiempo y la manera en que cambien en el futuro puede no ser conocida en el presente. Sea

$$\mathcal{N}_S^T \triangleq \{X \in \mathcal{R}_{S,T}^\infty \mid X \text{ no negativo}\},$$

el conjunto de procesos estocásticos en $\mathcal{R}_{S,T}^\infty$, no negativos (puede ser igual a cero porque la posición se puede cerrar en algún momento). Entonces la condición que se impone sobre \mathcal{M}_S^T es la siguiente. Para todo $\lambda = (\lambda_t)_{t \in \mathbb{N}} \in \mathcal{N}_S^T$, $\delta = (\delta_t)_{t \in \mathbb{N}} \in \mathcal{N}_S^T$ y $L, L' \in \mathcal{M}_S^T$,

$$\lambda L + \delta L' = (\lambda_t L_t + \delta_t L'_t)_{t \in \mathbb{N}} \in \mathcal{M}_S^T.$$

Se presenta a continuación la definición de medidas dinámicas de riesgo, adaptada de las presentadas por Riedel (2004), Cheridito, et al. (2004) y Balbás et al. (2002).

Definición 3.1.1. *Una medida dinámica de riesgo ψ sobre $[S, T]$ es un funcional*

$$\psi : \mathcal{M}_S^T \rightarrow \mathcal{R}^\infty. \quad (6)$$

La razón por la cual la imagen de la medida ψ está contenida en \mathcal{R}^∞ es porque la medida de riesgo en la fecha $\tau \in \mathbb{N}$ depende de la información disponible hasta el

momento en que se haga la medición del riesgo. Por ejemplo, en el momento τ no se sabe en general con exactitud el valor que tendrá la medida de riesgo en $s > \tau$, sino que éste valor es la variable aleatoria $\psi_s(L)$. Se pide entonces que $\psi(L)$ sea (\mathcal{F}_t) -adaptado para que para cada t , $\psi_t(L)$ sea \mathcal{F}_t -medible.

No se define la medida de riesgo dinámica de la manera como lo hacen Artzner, et al. (2003), en donde el dominio es el mismo pero la imagen de la medida está contenida en \mathbb{R} . En este caso la medida tiene en cuenta la evolución de los riesgos en el tiempo, pero no la evolución de la medida misma, cuando se mide en las diferentes fechas $t \in \mathbb{N}$. A diferencia de esto, la definición 3.1.1 sí permite directamente postular condiciones deseables para la evolución de los valores de las medidas de riesgo en el tiempo.

Por otro lado, la definición de medida dinámica de riesgo propuesta por Riedel (2004)³ es $\phi = (\phi_t)_{t=0, \dots, T-1}$, con $\phi_t : \mathcal{M}_0^T \times \Omega \rightarrow \mathbb{R}$. Asume naturalmente que para todo t y $L \in \mathcal{M}_S^T$ fijo, $\phi_t(L, \cdot)$ es \mathcal{F}_t -medible, así que $(\phi_t(L, \cdot))_{t=0, \dots, T}$ es un proceso estocástico (\mathcal{F}_t) -adaptado, lo cual se tiene directamente en la definición acá presentada. Además, todas las propiedades deseables de las medidas propuestas por este autor se pueden traducir a la definición acá propuesta directamente. Por esto se definen las medidas dinámicas de riesgo como en la Definición 3.1.1 y no como lo propone Riedel, ya que también se considera que la Definición 3.1.1 es más sencilla y guarda más similitudes con la definición de medida de riesgo estática.

Esta definición es fundamentalmente la definición propuesta por Balbás et al. (2002). Dicho autor define una medida de riesgo como $M : \mathbb{R}^m \rightarrow \mathcal{R}^\infty$, donde a cada $x \in \mathbb{R}^m$ le asigna el proceso ρ_x . El dominio de la medida es \mathbb{R}^m por lo siguiente. En su trabajo el autor supone que existen m instrumentos financieros (los activos I_1, \dots, I_m) e identifica cada $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ con una inversión inicial en el portafolio compuesto por la combinación lineal de x_i unidades invertidas en el instrumento financiero I_i . Observe entonces que, identificando adecuadamente y suponiendo que existieran m instrumentos financieros, las definiciones coinciden. De nuevo, se considera que la manera en que se presenta la definición es una extensión

³Trabaja sobre las fechas $\{0, 1, \dots, T\}$

que guarda más similitudes con la definición de medida de riesgo estática y es más general.

Otra definición bastante interesante es debida a Cheridito, et al. (2004). Dichos autores definen las medidas de riesgo⁴ como un funcional

$$\phi_{S,T} : \mathcal{R}_{S,T}^{\infty} \rightarrow L^{\infty}(\mathcal{F}_S),$$

donde $L^{\infty}(\mathcal{F}_S)$ es el conjunto de las variables aleatorias en L^{∞} que son \mathcal{F}_S -medibles. Después extienden los axiomas de coherencia estáticos al contexto dinámico sobre esta medida y para considerar preguntas de consistencia temporal de las medidas de riesgo consideran para una inversión en un intervalo $[S, T] \cap \mathbb{N}$ el proceso $(\phi_{t,T})_{t \in [S, T] \cap \mathbb{N}}$. Esto permite, como se verá más adelante, que la extensión de los axiomas de coherencia del contexto estático al dinámico sea más fuerte que la extensión directa hecha de la manera como lo hacen otros autores. El proceso $(\phi_{t,T}(L))_{t \in [S, T] \cap \mathbb{N}}$ es esencialmente el mismo proceso que se obtiene de la definición acá presentada, restringiendo para cada $L \in \mathcal{R}_{S,T}^{\infty}$, $(\psi_t(L))_{t \in [S, T] \cap \mathbb{N}}$. Así el dominio no sea \mathcal{R}_S^T , por la definición de \mathcal{M}_S^T no se pierde generalidad, sino que no se consideran procesos que no son relevantes. No se define de esta manera las medidas dinámicas de riesgo porque se considera que la manera en que se presenta en este trabajo es más sencilla y (de nuevo) guarda más similitudes con las medidas estáticas.

3.2. Propiedades deseables

Se trabaja a continuación en \mathcal{M}_S^T , para $S, T \in \mathbb{N}$, $0 \leq S < T < \infty$. Es decir, se analiza el riesgo de posiciones que se abren en la fecha S (o después) y se cierran en la fecha T (o antes). Luego, las pérdidas estocásticas que se analizan son elementos de este conjunto.

3.2.1. Independencia del pasado

Riedel (2004) propone que el pasado no debe afectar el valor de la medida de riesgo, una propiedad que es consistente con el hecho que para muchos modelos se

⁴Los autores usan la terminología “*monetary utility functional*”, que está estrechamente ligada con lo que ellos llaman medidas de riesgo

supone que los precios de las acciones y otras variables financieras no dependen del pasado sino solamente del precio de la misma en el presente, como lo hace el modelo de Black-Scholes-Merton para valorar opciones.

El axioma propuesto por Riedel (2004) dice que para todo $L, L' \in \mathcal{M}_S^T$ y $\tau \in \mathbb{N}$, si $L_t = L'_t$ c.s para todo $t \geq \tau$, entonces $\psi_\tau(L) = \psi_\tau(L')$, c.s.

Se propone el Axioma de una manera más fuerte, ya que se logra de esta manera también reflejar la idea propuesta por los autores Detlefsen y Scandolo (2005), que proponen que una medida dinámica de riesgo debe cumplir la condición de *regularidad*. Dicha condición asegura que si se sabe que un evento $A \in \mathcal{F}_\tau$ cualquiera va a prevalecer en el futuro, entonces el nivel de riesgo de las pérdidas estocásticas debe depender solamente de lo que en realidad pueda ocurrir. Esto es, si $A \in \mathcal{F}_{\tau+1}$, entonces $\psi_\tau(L \mathbb{I}_A) = \psi_\tau(L) \mathbb{I}_A$. Se tiene que si $\forall u, \psi_u(0) = 0$, el siguiente axioma refleja esto, como se demuestra en la Proposición 3.2.1.

Axioma 6 (Independencia del pasado). *Para todo $L, L' \in \mathcal{M}_S^T$ y $\tau, s \in \{0, 1, \dots, T\}$, con $\tau \leq s$, si $A \in \mathcal{F}_\tau$ y $L_t \mathbb{I}_A = L'_t \mathbb{I}_A$ c.s para todo $t \geq s$, entonces*

$$\psi_{s-1}(L) \mathbb{I}_A = \psi_{s-1}(L') \mathbb{I}_A, \text{ c.s.}$$

Proposición 3.2.1. *Si $\forall u, \psi_u(0) = 0$, el axioma de Independencia del pasado es equivalente a lo siguiente. Para todo $L \in \mathcal{M}_S^T$ y $\tau, s \in \{0, 1, \dots, T\}$, con $\tau \leq s$, si $A \in \mathcal{F}_\tau$,*

$$\psi_{s-1}(L \mathbb{I}_A) = \psi_{s-1}(L) \mathbb{I}_A, \text{ c.s.}$$

Demostración. (Ver Detlefsen y Scandolo, 2005) (\Rightarrow) Se tiene que $L \mathbb{I}_A = (L \mathbb{I}_A) \mathbb{I}_A$ y $0 \mathbb{I}_{A^c} = (L \mathbb{I}_A) \mathbb{I}_{A^c}$. Entonces se tiene que $\psi_{s-1}(L) \mathbb{I}_A = \psi_{s-1}(L \mathbb{I}_A) \mathbb{I}_A$ y $0 = \psi_{s-1}(0) \mathbb{I}_{A^c} = \psi_{s-1}(L \mathbb{I}_A) \mathbb{I}_{A^c}$ c.s, por lo que

$$\psi_{s-1}(L \mathbb{I}_A) = \psi_{s-1}(L \mathbb{I}_A) \mathbb{I}_A + \psi_{s-1}(L \mathbb{I}_A) \mathbb{I}_{A^c} = \psi_{s-1}(L) \mathbb{I}_A, \text{ c.s.}$$

(\Leftarrow) Si $L_t \mathbb{I}_A = L'_t \mathbb{I}_A$ c.s para todo $t \geq s$, entonces

$$\psi_{s-1}(L) \mathbb{I}_A = \psi_{s-1}(L \mathbb{I}_A) = \psi_{s-1}(L' \mathbb{I}_A) = \psi_{s-1}(L') \mathbb{I}_A, \text{ c.s.} \quad \square$$

Cheridito, et al. (2004) proponen la condición de esta Proposición. Más exactamente, proponen para cada $\phi_{t,T}$ del proceso $(\phi_{t,T})_{t \in [S,T] \cap \mathbb{N}}$, que si $L \in \mathcal{M}_t^T$ y $A \in \mathcal{F}_S$, entonces $\phi_{t,T}(L \mathbb{I}_A) = \phi_{t,T}(L) \mathbb{I}_A$. El axioma de independencia del pasado presentado es más general, ya que A no pertenece a \mathcal{F}_s sino a algún \mathcal{F}_τ y si se cumple el axioma de homogeneidad positiva dinámica (ver sección 3.2.4) se cumple que $\psi_u(0) = \psi_u(2 \times 0) = 2\psi_u(0)$, así que en este caso $\forall u, \psi_u(0) = 0$.

3.2.2. Monotonidad y subaditividad dinámicas

El axioma de monotonidad en el contexto dinámico propuesto por los autores Riedel (2004) y Balbás et al. (2002) dice que $L \geq L'$ c.s implica $\psi_t(L) \geq \psi_t(L')$ c.s. Se propone la extensión de este axioma al contexto dinámico un poco más general, de manera similar a como lo hacen Cheridito, et al. (2004).

Axioma 7 (Monotonidad dinámica). *Para todo $s \in \{0, \dots, T\}$, si $L_t \geq L'_t$ c.s para todo $t \geq s$, entonces*

$$\psi_t(L) \geq \psi_t(L') \text{ c.s.} \quad (7)$$

para todo $t \geq s - 1$.

La desigualdad (7) se tiene solo para $t \geq s - 1$ porque en las fechas entre 0 y $s - 1$ no se tiene información sobre los flujos de caja acumulados. Esto es, no hay ninguna restricción para L_0, \dots, L_{s-1} y L'_0, \dots, L'_{s-1} y asumir la desigualdad en estas fechas puede llevar a conclusiones erradas. Para ver esto se presenta el siguiente ejemplo.

Ejemplo 3.1. *Suponga $S = 0$ y $T = 2$. Sean $L = (L_1, L_2)$, con $L_1 \geq 0, L_2 \leq 0$ y $L' = (L'_1, L'_2)$, con $L'_1 = 0$ y $L'_2 = \frac{1}{2}L_2$. Como $L_2 \leq L'_2$, se debe cumplir $\psi_1(L) \leq \psi_1(L')$. No obstante, como $L' \leq 0$, se tiene que $\psi_0(L') \leq 0$, pero $L_1 \geq 0$, por lo que $\psi_0(L) \geq 0$. Entonces, $\psi_0(L) \geq \psi_0(L')$. Ver Figura ??.*

Los autores Cheridito, et al. (2004) definen esta propiedad para cada $\phi_{t,T}$ del proceso $(\phi_{t,T})_{t \in [S,T] \cap \mathbb{N}}$ de la siguiente manera. Si $L, L' \in \mathcal{M}_t^T$ son tales que $L \leq L'$ c.s, entonces $\phi_{t,T}(L) \leq \phi_{t,T}(L')$. Este es un axioma evidentemente más fuerte que el propuesto por los otros autores nombrados, pero menos fuerte que el presentado porque en el presente trabajo los elementos L y L' no son restringidos a \mathcal{M}_t^T .

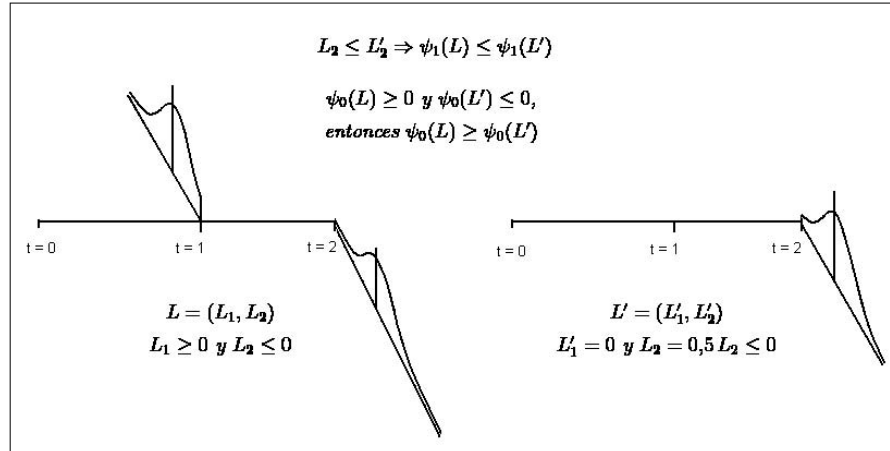


Figura 4: Ejemplo monotonicidad

La subaditividad tiene una extensión más natural al contexto dinámico que la monotonicidad. Esta extensión es la siguiente.

Axioma 8 (Subaditividad dinámica). *Para todo $L, L' \in \mathcal{M}_S^T$ y todo $t \in \{0, \dots, T\}$,*

$$\psi_t(L + L') \leq \psi_t(L) + \psi_t(L'), \text{ c.s.}$$

3.2.3. Invarianza dinámica bajo traslación

El axioma de invarianza bajo traslación tiene en este contexto diferentes extensiones posibles. La primera, propuesta por Balbás et al. (2002), está relacionada con la medida del proceso $L + l \mathbb{I}_{\{t \geq S\}}$, con $l \in \mathbb{R}$. Es decir, con la medida de riesgo de un proceso si se le añade una cantidad l fija de dinero durante $t = S, S + 1, \dots, T$. Esta extensión se llama invarianza *débil* por razones que se verán adelante.

Axioma 9 (Invarianza dinámica débil bajo traslación). *Suponga que $L \in \mathcal{M}_S^T$ y $l \in \mathbb{R}$, entonces, para todo $t \in \{0, 1, \dots, T\}$,*

$$\psi_t(L + l \mathbb{I}_{\{t \geq S\}}) = \psi_t(L) + l, \text{ c.s.}$$

Ahora, a pesar de que en el contexto estático es equivalente añadir l a la posición en el momento 0 o añadir una pérdida aleatoria que represente un flujo de caja en T , predecible e igual a l . Esto no es igual en el contexto dinámico porque los riesgos

predecibles pueden aparecer en diferentes instantes de tiempo y ser diferentes entre ellos.

Por esto, otra extensión posible propuesta por Riedel (2004) está relacionada con flujos de caja predecibles.

Axioma 10 (Invarianza bajo traslación predecible). *Sea $s \geq \tau$ y $Z \in L^\infty(\mathcal{F}_\tau)$ y defina $\lambda = (\lambda_t)_{t \in \mathbb{N}} \in \mathcal{M}_S^T$ por $\lambda_t \triangleq Z \mathbb{I}_{\{t \geq s\}}$. Entonces, para todo $L \in \mathcal{M}_S^T$,*

$$\psi_\tau(L + \lambda) = \psi_\tau(L) + Z, \text{ c.s.} \quad (8)$$

Observe que este axioma implica la invarianza débil tomando $\tau = s = 0$.

Discusión sobre el axioma invarianza bajo traslación predecible: El axioma invarianza bajo traslación predecible no es considerado en este trabajo como un axioma deseable para las medidas de riesgo dinámicas por lo siguiente. Se desea con las medidas dinámicas de riesgo asegurar que la posición que se toma en el lapso de tiempo que va desde S a T pueda efectivamente ser sostenida durante todo ese tiempo, evitando una posible quiebra en fechas intermedias, lo cual puede no ser el caso si se utiliza una medida que cumpla dicho axioma. Para ver esto se presenta el siguiente ejemplo.

Ejemplo 3.2. *Suponga $S = 0$ y $T = 12$. Sea $L = (L_t)_{t=1, \dots, 12}$, con $L_1 = L_2 = \dots = L_5 = 0$ y $L_6 = L_7 = \dots = L_{12} \geq 0$. Entonces L representa una posición que no dará ganancias, con un solo flujo de caja no negativo en $t = 6$. Suponga además que $\psi_0(L) = 150$. Por otro lado, sea $\lambda = -150 \mathbb{I}_{\{t > 11\}}$, una posición que representa una ganancia de 150 en $t = 12$. Por lo tanto, si se cumple el axioma de invarianza bajo traslación predecible $\psi_0(L + \lambda) = \psi_0(L) - 150 = 0$, pero evidentemente se debe tener que $\psi_0(L + \lambda) > 0$, ya que $L + \lambda = (0, \dots, 0, L_6, \dots, L_{11}, L_{12} - 150)$ y el riesgo en las fechas 6 a 11 de $L + \lambda$ es igual que el riesgo de L . Entonces hay un riesgo de quiebra en las fechas intermedias que estaría siendo ignorado. Ver Figura ??.*

Similarmente a como lo hacen los autores Cheridito, et al. (2004), la siguiente es una extensión al contexto dinámico del axioma de invarianza bajo traslación.

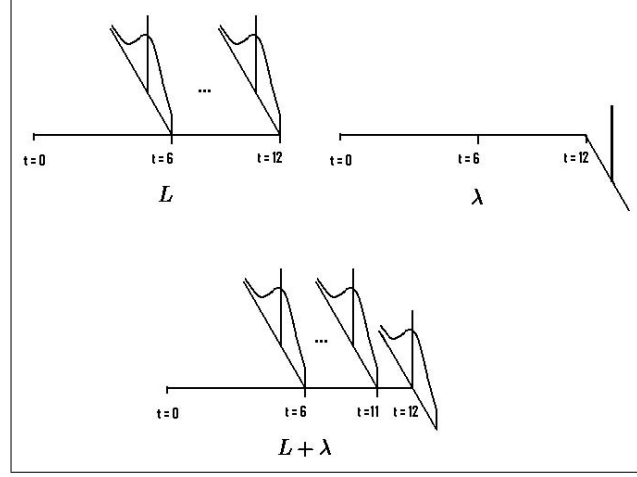


Figura 5: Ejemplo invarianza bajo traslación predecible

Axioma 11 (Invarianza dinámica bajo traslación). Sean $s, \tau \in \mathbb{N}$, con $S \leq s \leq T$ y $0 \leq \tau \leq T$. Entonces, para todo $L \in \mathcal{M}_S^T$, todo $\lambda = (\lambda_t)$ dado por $\lambda_t = Z \mathbb{I}_{\{t \geq s\}}$, con $Z \in L^\infty(\mathcal{F}_\tau)$ y todo $t \geq s - 1$,

$$\psi_t(L + \lambda) = \psi_t(L) + Z, \text{ c.s.} \quad (9)$$

Este axioma dice el valor del riesgo al incluir un flujo de caja predecible debe afectar el valor del riesgo del portafolio de manera similar que en el contexto estático, pero solamente desde la fecha $s - 1$ en adelante, ya que en este caso no hay riesgo de quebrar antes que se realice efectivamente el flujo de caja.

Cheridito, et al. (2004) proponen este axioma de la siguiente manera. Para cada $\phi_{t,T}$ del proceso $(\phi_{t,T})_{t \in [S,T] \cap \mathbb{N}}$, $\phi_{t,T}(L + Z \mathbb{I}_{[t,\infty)}) = \phi_{t,T}(L) + Z$, para todo $L \in \mathcal{R}_{t,T}^\infty$ y $Z \in L^\infty(\mathcal{F}_t)$.

Continuación de la discusión sobre el axioma de invarianza bajo traslación: El axioma de invarianza dinámica bajo traslación propuesto es evidentemente estrictamente más fuerte que la invarianza débil bajo traslación y no tiene el problema expuesto del axioma invarianza bajo traslación predecible. No obstante, es interesante buscar cómo extraer más información para la medida de riesgo al añadir una posición con flujos de caja predecibles. Esto es importante porque si se pide

únicamente a una medida de riesgo que cumpla invarianza dinámica bajo traslación, ésta puede cumplir con la invarianza bajo traslación predecible, lo que puede traer consecuencias negativas como se observa en el Ejemplo 3.2. Para evitar esto se propone el siguiente axioma.

Axioma 12 (Invarianza dinámica fuerte bajo traslación). Sean $s, \tau \in \mathbb{N}$, con $S \leq s \leq T$ y $0 \leq \tau \leq T$. Entonces, para todo $L \in \mathcal{M}_S^T$, todo $\lambda = (\lambda_t) \in \mathcal{M}_S^T$ dado por $\lambda_t = Z \mathbb{I}_{\{t \geq s\}}$, con $Z \in L^\infty(\mathcal{F}_\tau)$ y todo $t \geq s - 1$,

$$\psi_t(L + \lambda) = \psi_t(L) + Z, \quad c.s, \quad (10)$$

y si $\tau \leq t < s - 1$,

$$\psi_t(L + \lambda) = \max\{\psi_t(\pi_{s,T}(L)) + Z, \psi_t(\pi_{0,s-1}(L))\}, \quad c.s, \quad (11)$$

recordando que $\pi_{s,T}(L) = (\mathbb{I}_{\{t \geq s\}} X_{t \wedge T})_{t \in \mathbb{N}}$ y $\pi_{0,s-1}(L) = (L_{t \wedge (s-1)})_{t \in \mathbb{N}}$.

El proceso $\pi_{s,T}(L)$ modela los flujos de caja acumulados de L desde s hasta la fecha T y el proceso $\pi_{0,s-1}(L)$ modela los flujos de caja acumulados de L hasta la fecha $s - 1$. Entonces, $\psi_t(\pi_{s,T}(L))$ (resp. $\psi_t(\pi_{0,s-1}(L))$) es el capital necesario para que la posición L sea aceptable desde s hasta T (resp. desde t hasta cuando se realiza el flujo de caja en s), dada la información hasta t . Si

$$(\psi_t(\pi_{s,T}(L)) + Z)(\omega) \geq \psi_t(\pi_{0,s-1}(L))(\omega),$$

se cumple $\psi_t(L + \lambda)(\omega) = (\psi_t(\pi_{s,T}(L)) + Z)(\omega) = \psi_t(\pi_{s,T}(L))(\omega) + Z(\omega)$, similar al axioma de invarianza bajo traslación predecible, pero se asegura que la posición es aceptable hasta que se realice efectivamente el flujo de caja. Entonces, como la posición está asegurada hasta s , no se pide que sea igual que en dicho axioma, es decir, no se pide que $\psi_t(L + \lambda)(\omega) = \psi_t(L)(\omega) + Z(\omega)$.

Ahora, si

$$\psi_t(\pi_{s,T}(L))(\omega) + Z(\omega) \leq \psi_t(\pi_{0,s-1}(L))(\omega),$$

es necesario un capital extra para asegurar la posición hasta que se realice el flujo de caja. El capital necesario es justamente $\psi_t(\pi_{0,s-1}(L))(\omega)$.⁵

⁵Esto no se pide para todo $\omega \in \Omega$, sino que el conjunto donde esto no se cumple debe tener medida 0

Ejemplo 3.3. Retomando el Ejemplo 3.2, observe que $L + \lambda$ tiene exactamente el mismo riesgo que L en las fechas $0, 1, \dots, 11$ y es generado por el flujo de caja en $t = 6$. Además λ disminuye el riesgo, pero solamente desde la fecha $t = 12$. Por lo tanto, en $t = 0$ no se debe reducir el riesgo de la posición $L + \lambda$ respecto a la posición L , entonces $\psi_0(L + \lambda)$ debe ser 150.

Se tiene que $\pi_{0,11}(L) = (0, \dots, 0, L_6, \dots, L_{11}, L_{11}) = (0, \dots, 0, L_6, \dots, L_6) = L$ y $\pi_{12,12}(L) = (0, \dots, 0, L_{12}) = (0, \dots, 0, L_6)$ (ver Figura ??). Como $\pi_{12,12}(L)$ tiene exactamente el mismo flujo de caja que L , conocido en $t = 6$, pero se hace efectivo en $t = 12$ y dado que los flujos están descontados con respecto a la tasa de referencia dada, se debe tener que $\psi_0(\pi_{12,12}(L)) = 150$. Por lo tanto, el axioma de invarianza dinámica fuerte bajo traslación dice que $\psi_0(L + \lambda) = \psi_0(\pi_{0,11}(L)) = 150 > \psi_0(\pi_{12,12}(L)) - 150 = 0$.

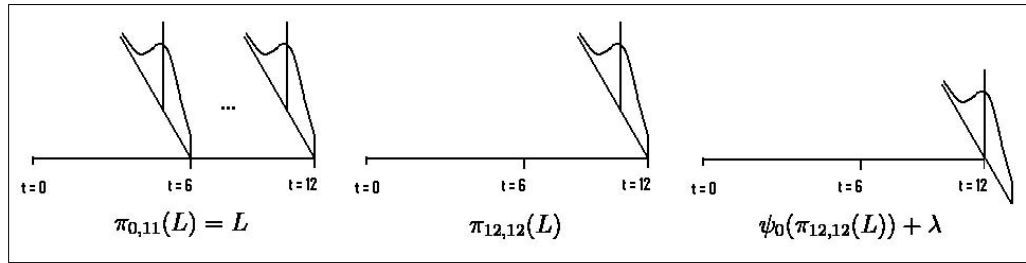


Figura 6: Continuación ejemplo invarianza bajo traslación predecible

Observe que si ψ cumple este axioma implica que no cumple el axioma de invarianza bajo traslación predecible. Esto se puede ver tomando L y λ como en los Ejemplos 3.2 y 3.3.

3.2.4. Homogeneidad positiva dinámica

Un problema similar al que se presenta con la extensión al contexto dinámico del axioma de invarianza bajo traslación ocurre con el axioma de homogeneidad positiva. La primera extensión es la siguiente (Riedel (2004), Balbás et al. (2002) y Hardy y Wirch (2004)).

Axioma 13 (Homogeneidad positiva dinámica débil). *Para todo $L \in \mathcal{M}_S^T$, $d \in \mathbb{R}$, $d > 0$ y $t \in \{0, \dots, T\}$,*

$$\psi_t(dL) = d\psi_t(L), \text{ c.s.}$$

En el contexto dinámico no se desea solamente que se cumpla la homogeneidad positiva dinámica débil, sino que se debe cumplir una condición más fuerte. Como motivación de ésta se presenta el siguiente ejemplo.

Ejemplo 3.4. *Suponga que el perfil de riesgo $L \in \mathcal{M}_S^T$ representa las pérdidas de una posición que se tendrá por un año y se monitoreará cada mes el riesgo de ésta. Ahora suponga que se quiere medir el valor de la misma posición pero con el siguiente cambio. En el mes 3 y no antes se sabrá con exactitud si a los seis meses se conservará la posición como está o si se multiplicará el tamaño de ésta por única vez en el año por una cantidad aleatoria positiva d . La cantidad d es en $t = 0$ una variable aleatoria cuyo valor afectará la posición en el mes 6 hasta que se cierre la misma, pero desde el mes 3 su valor será conocido. Por esto, el riesgo de la posición medido en cualquier fecha desde el mes 5 debe ser exactamente d veces el valor de la posición L como en el contexto estático. Para los meses 3 y 4 se debe tener un cuidado extra para garantizar la aceptabilidad de la posición hasta que se efectúe el cambio a la misma en el mes 6, de manera similar que para el axioma de invarianza bajo traslación.*

Axioma 14 (Homogeneidad positiva dinámica y Homogeneidad positiva dinámica fuerte). *Sean $0 \leq \tau \leq s \leq T$ y $\delta = (\delta_t)$, con $\delta_t = d\mathbb{I}_{\{t \geq s\}} + \mathbb{I}_{\{t < s\}}$, para $d \in L^\infty(\mathcal{F}_\tau)$, $d \geq 0$. Entonces ψ cumple con el axioma de homogeneidad positiva dinámica si para todo $L \in \mathcal{M}_S^T$ y todo $t \geq s - 1$,*

$$\psi_t(\delta L) = d\psi_t(L), \text{ c.s.} \tag{12}$$

Además, ψ cumple con el axioma de homogeneidad positiva dinámica fuerte si también es tal que para $\tau \leq t < s - 1$,

$$\psi_t(\delta L) = \max\{d\psi_t(\pi_{s,T}(L)), \psi_t(\pi_{0,s-1}(L))\}, \text{ c.s.} \tag{13}$$

Cheridito, et al. (2004) proponen el axioma de homogeneidad positiva dinámica de la siguiente manera. Para cada $\phi_{t,T}$ del proceso $(\phi_{t,T})_{t \in [S,T] \cap \mathbb{N}}$ y para todo $L \in \mathcal{R}_{t,T}^\infty$ y $d \in L^\infty(\mathcal{F}_t)$, $d \geq 0$ se tiene que, $\phi_{t,T}(dL) = d\phi_{t,T}(L)$.

Note que homogeneidad positiva dinámica fuerte implica homogeneidad positiva dinámica y homogeneidad positiva dinámica implica homogeneidad positiva dinámica débil, pero las implicaciones contrarias no se cumplen.

De nuevo, el proceso $\pi_{0,s-1}(L)$ modela los flujos de caja acumulados de L hasta la fecha $s - 1$ y entonces $\psi_t(\pi_{0,s-1}(L))$ es el capital necesario para que la posición L sea aceptable desde t hasta $s - 1$. Si

$$d\psi_t(\pi_{s,T}(L))(\omega) \geq \psi_t(\pi_{0,s-1}(L))(\omega),$$

se cumple que $\psi_t(\delta L)(\omega) = d\psi_t(\pi_{s,T}(L))$, similarmente al contexto estático, pero se asegura que la posición sea aceptable hasta que se realice efectivamente el cambio en la misma. No obstante, para evitar sobreestimar el riesgo no se pide que $\psi_t(\delta L)(\omega) = d\psi_t(L)(\omega)$. Ahora, si

$$d\psi_t(\pi_{s,T}(L))(\omega) \leq \psi_t(\pi_{0,s-1}(L))(\omega),$$

es necesario asegurar la posición hasta que se realice el cambio en ésta y el capital necesario es $\psi_t(\pi_{0,s-1}(L))(\omega)$.⁶

En este caso ocurre algo similar a lo discutido sobre la invarianza bajo traslación. Se puede definir un axioma análogo a la invarianza bajo traslación predecible, pero también se puede fácilmente crear un ejemplo análogo al 3.2, que muestre que un axioma así no es deseable.

La diferenciación entre homogeneidad positiva y homogeneidad positiva débil no la hacen los autores Artzner, et al. (2003), Riedel (2004), Balbás et al. (2002) y Hardy y Wirch (2004). Dichos autores trabajan con el axioma definido en este trabajo como homogeneidad positiva débil.

⁶Esto no se pide para todo $\omega \in \Omega$, sino que el conjunto donde esto no se cumple debe tener medida 0

3.2.5. Medidas coherentes dinámicas

Finalmente se puede extender la definición de medidas coherentes al contexto dinámico. La definición es la siguiente.

Definición 3.2.1 (Medida coherente dinámica). *Una medida $\psi : \mathcal{M}_S^T \rightarrow \mathcal{R}^0$ sobre $[S, T]$ es coherente dinámica si satisface los siguientes axiomas:*

1. *Independencia del pasado, Axioma 6*
2. *Monotonidad dinámica, Axioma 7*
3. *Subaditividad dinámica, Axioma 8*
4. *Invarianza dinámica bajo traslación, Axioma 11*
5. *Homogeneidad positiva, Axioma 14*

La medida es fuertemente coherente si además cumple los axiomas:

1. *Invarianza dinámica fuerte bajo traslación, Axioma 12*
2. *Homogeneidad positiva fuerte, Axioma 14*

Se propone la definición de medida fuertemente coherente, ya que se observó que una medida coherente pero no fuertemente coherente puede cumplir con propiedades no deseables.

3.2.6. Consistencia temporal

Además de los anteriores axiomas básicos existen diferentes propuestas de axiomas sobre consistencia temporal. Detlefsen y Scandolo (2005) proponen un axioma de consistencia temporal de la siguiente manera. Recordando que en su trabajo los autores modelan solo los pagos finales de la pérdida aleatoria, si $X, X' \in L^\infty$, entonces el axioma de consistencia temporal que proponen dice que (abusando de la notación), para $0 \leq t \leq T - 1$,

$$\psi_{t+1}(X) = \psi_{t+1}(X') \Rightarrow \psi_t(X) = \psi_t(X').$$

Este axioma tiene sentido si se modelan los pagos finales de la posición porque si, independientemente de los estados de la naturaleza, se tiene en $t + 1$ en particular que los pagos finales de la posición dada son aceptables, entonces deben serlo en t . En el contexto de este trabajo este axioma no es apropiado porque $\psi_{t+1}(\cdot)$ considera solo los pagos L_{t+2}, \dots, L_T y no considera lo que pase en L_{t+1} , a diferencia de $\psi_t(\cdot)$.

Otra definición de consistencia temporal es la que se encuentra en Cheridito, et al. (2004). La definición es la siguiente.

Axioma 15 (Consistencia temporal). *Para todo $L \in \mathcal{M}_S^T$ y todos $0 \leq t < s \leq T$,*

$$\psi_t(L) = \psi_t(L \mathbb{I}_{[0,s]} + \psi_s(L) \mathbb{I}_{(s,\infty)}). \quad (14)$$

Esta es una propiedad deseable para las medidas dinámicas de riesgo. El axioma dice que el valor del riesgo de la posición L , observado en la fecha t , debe ser igual al riesgo de analizar el proceso con pérdidas aleatorias $L_0, \dots, L_s, \psi_s(L), \psi_s(L), \dots$. Esto es adecuado porque la variable aleatoria $\psi_s(L)$ debe reflejar el riesgo que generan las pérdidas $L_{s+1}, L_{s+2}, \dots, L_T, L_T, \dots$.

Cheridito, et al. (2004) proponen este axioma de la siguiente manera. Para cada $\phi_{t,T}$ del proceso $(\phi_{t,T})_{t \in [S,T] \cap \mathbb{N}}$, si s es tal que $t \leq s \leq T$ y si $L \in \mathcal{R}_t^T$, entonces

$$\phi_{t,T}(L) = \phi_{t,T}(L \mathbb{I}_{[t,s]} + \phi_{s,T}(L) \mathbb{I}_{[s,\infty)}).$$

3.2.7. Monotonidad temporal

Se propone que las medidas dinámicas de riesgo además cumplan la siguiente propiedad. Dado que los elementos de \mathcal{M}_S^T son flujos de caja acumulados, se pide que al observar los valores de la medida de riesgo de cualquier L en la fecha t y compararla con la medida de riesgo en la fecha $s > t$, como los conjuntos $A \in \mathcal{F}_t$ contienen la información que se conoce de hasta t , se debe tener que el promedio en cualquier $A \in \mathcal{F}_t$ de la medida de riesgo en t no es menor que el promedio de la medida en s sobre el mismo conjunto. Esto es, $\mathbb{E}[\mathbb{I}_A \psi_t(L)] \geq \mathbb{E}[\mathbb{I}_A \psi_s(L)]$, para todo $A \in \mathcal{F}_t$.

El motivo de esto es el hecho que la medida en s considera, los flujos de caja acumulados $L_{s+1}, L_{s+2}, \dots, L_T$ para evitar una posible quiebra desde la fecha $s + 1$

hasta T , mientras que la medida en t hace lo mismo pero desde $t + 1$ hasta T , considerando los mismos flujos de caja acumulados y además los flujos L_{t+1}, \dots, L_s . En otras palabras, el valor $\psi_t(L)$ busca asegurar la viabilidad de la inversión por más tiempo y considerando no menos riesgos que $\psi_s(L)$.

Por razones que serán explicadas a continuación, para asegurar esto se pide que para todo L , $\mathbb{E}[\psi_t(L) \mid \mathcal{F}_t] \geq \mathbb{E}[\psi_s(L) \mid \mathcal{F}_t]$. Además, como $\psi_t(L)$ es \mathcal{F}_t -medible se utiliza el hecho que $\mathbb{E}[\psi_t(L) \mid \mathcal{F}_t] = \psi_t(L)$. Entonces se propone el siguiente axioma.

Axioma 16 (Monotonidad temporal). *Para todo $0 \leq t < s \leq T$ y todo $L \in \mathcal{M}_S^T$,*

$$\psi_t(L) \geq \mathbb{E}[\psi_s(L) \mid \mathcal{F}_t], \text{ c.s.}$$

Se propone el axioma de esta manera por lo siguiente. Primero, ψ cumple el axioma si y solo si para todo $A \in \mathcal{F}_t$,

$$\int_A \psi_t(L) d\mathbb{P} \geq \int_A \mathbb{E}[\psi_s(L) \mid \mathcal{F}_t] d\mathbb{P}. \quad (15)$$

En efecto, sea $B \triangleq \{\omega \in \Omega \mid \psi_t(L)(\omega) < \mathbb{E}[\psi_s(L) \mid \mathcal{F}_t](\omega)\}$. Como $\mathbb{E}[\psi_s(L) \mid \mathcal{F}_t]$ y $\psi_t(L)$ son \mathcal{F}_t -medibles, entonces $B \in \mathcal{F}_t$. Luego $\int_B \psi_t(L) d\mathbb{P} < \int_B \mathbb{E}[\psi_s(L) \mid \mathcal{F}_t] d\mathbb{P}$. Por lo tanto, (15) se cumple si y solo si $\mathbb{P}[B] = 0$, es decir, si el axioma se cumple.

Además, por la “propiedad de promedio parcial” se cumple que

$$\int_A \mathbb{E}[\psi_s(L) \mid \mathcal{F}_t] d\mathbb{P} = \int_A \psi_s(L) d\mathbb{P}.$$

Entonces ψ cumple el axioma si y solo si para todo $A \in \mathcal{F}_t$, $\mathbb{E}[\mathbb{I}_A \psi_t(L)] \geq \mathbb{E}[\mathbb{I}_A \psi_s(L)]$, como se quería.

3.3. Ejemplos

3.3.1. Extensiones triviales de medidas dinámicas

Sea ρ una medida estática de riesgo (como CVaR_α) tal que existe $\mathbb{Q} \ll \mathbb{P}$, tal que $\rho(\cdot) = \mathbb{E}_{\mathbb{Q}}[\cdot]$. Se puede considerar $\rho(\cdot)$ como una medida dinámica trivial, como lo hacen Hardy y Wirch (2004). La medida de L no cambia en el tiempo y su valor

es igual a la medida estática en el tiempo 0 de la variable L_T . Esto es,⁷

$$\psi_t^E(L) \triangleq \mathbb{E}_{\mathbb{Q}}[L_T \mid \mathcal{F}_0], \quad (16)$$

para todo τ . Se nota ψ^E ya que es la extensión trivial de una medida estática.

Dado lo anterior, se puede también naturalmente definir otra medida de riesgo que pueda capturar la nueva información que se revele al pasar el tiempo. Esta medida está dada por la siguiente expresión,

$$\psi_t^A(L) \triangleq \mathbb{E}_{\mathbb{Q}}[L_T \mid \mathcal{F}_t]. \quad (17)$$

Se nota ψ^A porque es la extensión de una medida estática que permite que la información sea actualizada.

3.3.2. Medida de riesgo iterada

A continuación se expone la metodología propuesta por Hardy y Wirch (2004) para construir medidas dinámicas de riesgo utilizando una medida estática coherente por medio de *recursión* hacia atrás. En particular definen la medida dinámica “*Iterated Tail Conditional Expectation*” (ITCE) partiendo de la medida CVaR o “*Tail Conditional Expectation*” definida anteriormente.

Sea ρ una medida estática de riesgo, tal que existe $\mathbb{Q} \ll \mathbb{P}$, tal que $\rho(\cdot) = \mathbb{E}_{\mathbb{Q}}[\cdot]$. Se construye entonces la medida de riesgo dinámica iterada ψ^I de la siguiente manera,

$$\psi_{T-1}^I(L) = \mathbb{E}_{\mathbb{Q}}[e^{-r} L_T \mid \mathcal{F}_{T-1}], \quad (18)$$

para el paso final y recursivamente hacia atrás se define

$$\psi_t^I(L) = \mathbb{E}_{\mathbb{Q}}[e^{-r} \psi_{t+1}^I(L) \mid \mathcal{F}_t], \quad 0 \leq t \leq T. \quad (19)$$

Cuando en particular se toma ρ como CVaR_{α} , esta medida dinámica de riesgo es llamada “*Iterated Tail Conditional Expectation*” o ITCE. Entre las propiedades que cumple esta medida, las autoras demuestran que cumple con el axioma de invarianza

⁷Ver el Apéndice A para la definición y algunas propiedades de esperanza condicional

bajo traslación predecible, por lo que no es una medida apropiada en el contexto de este trabajo. También se tiene que la medida solo considera los flujos de caja acumulados en la fecha final. Además, las autoras también muestran que si aumenta la cantidad de iteraciones, la medida de riesgo iterada puede aumentar dramáticamente comparada con el valor que tendría la medida CVaR_α .

3.3.3. Construcción de medidas dinámicas de riesgo

De nuevo, basado en medidas coherentes de riesgo se construye a continuación una medida dinámica. Suponga que $r = 0$. Abusando de la notación, una medida estática de riesgo ρ_t que mida la pérdida de una posición financiera en el intervalo $[0, t]$ puede ser considerada como una función $\rho_t : \mathcal{M}_S^t \rightarrow \mathbb{R}$, donde $\rho_t(L) \triangleq \rho_t(L_t)$. Utilizando este abuso de notación se propone lo siguiente.

Definición 3.3.1. *Dadas las medidas coherentes estáticas de riesgo $\rho_t : \mathcal{M}_S^t \rightarrow \mathcal{R}^\infty$ (donde ρ_t es una medida de riesgo estática en el intervalo $[0, t]$), para $t = 1, \dots, T$, tal que, para \mathbb{Q} una probabilidad absolutamente continua con respecto a \mathbb{P} se cumple que para todo t , $\rho_t(\cdot) = \mathbb{E}_{\mathbb{Q}}[\cdot]$, se define la medida dinámica $\psi^{\text{máx}} = (\psi_t^{\text{máx}})$ por*

$$\psi_t^{\text{máx}}(L) = \max_{i=t+1, \dots, T} \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t]. \quad (20)$$

Observe que por la definición de esperanza condicional, para todo $L \in \mathcal{M}_S^T$, $\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t]$ es \mathcal{F}_t -medible para todo i , luego $\psi_t^{\text{máx}}(L)$ lo es. Además, $\psi_t^{\text{máx}}(L) \in \mathcal{R}^\infty$, porque para todo $i = t+1, \dots, T$, $\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] \in \mathcal{R}^\infty$. Por lo tanto, la medida está bien definida. A continuación se verifican las propiedades de $\psi^{\text{máx}}$.

Proposición 3.3.1. *La medida dinámica $\psi^{\text{máx}}$ de la Definición 3.3.1 es fuertemente coherente.*

Demostración. Sean $L, L' \in \mathcal{M}_S^T$ y $0 \leq t \leq s \leq T$.

Independencia del pasado: Sea $A \in \mathcal{F}_t$. Trivialmente se tiene que $\forall u, \psi_u^{\text{máx}}(0) = 0$, entonces se verificará que $\psi_s^{\text{máx}}(L \mathbb{I}_A) = \psi_s^{\text{máx}}(L) \mathbb{I}_A$, c.s. Como \mathbb{I}_A es una variable aleatoria \mathcal{F}_t -medible, es \mathcal{F}_s -medible, entonces

$$\forall i \geq s+1, \mathbb{E}_{\mathbb{Q}}[L_i \mathbb{I}_A | \mathcal{F}_s] = \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_s] \mathbb{I}_A.$$

Luego,

$$\begin{aligned}\psi_s^{\max}(L) \mathbb{I}_A &= \max_{i=s+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_s]\} \mathbb{I}_A = \max_{i=s+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_s] \mathbb{I}_A\} \\ &= \max_{i=s+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i \mathbb{I}_A | \mathcal{F}_s]\} = \psi_s^{\max}(L \mathbb{I}_A), \text{ c.s.}\end{aligned}$$

Monotonidad dinámica: Suponga $L_u \geq L'_u$ c.s, para todo $u \geq s$. Se tiene que

$$\forall u, \forall i \geq s, \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_u] \geq \mathbb{E}_{\mathbb{Q}}[L'_i | \mathcal{F}_u], \text{ c.s.}$$

Por esto se cumple c.s lo siguiente, si $t \geq s - 1$,

$$\psi_t^{\max}(L) = \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t]\} \geq \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L'_i | \mathcal{F}_t]\} = \psi_t^{\max}(L').$$

Subaditividad dinámica:

$$\begin{aligned}\psi_t^{\max}(L + L') &= \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i + L'_i | \mathcal{F}_t]\} = \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + \mathbb{E}_{\mathbb{Q}}[L'_i | \mathcal{F}_t]\} \\ &\leq \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t]\} + \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L'_i | \mathcal{F}_t]\} \\ &= \psi_t^{\max}(L) + \psi_t^{\max}(L'), \text{ c.s.}\end{aligned}$$

Invarianza dinámica fuerte bajo traslación: Sean $s, \tau \in \mathbb{N}$, con $S \leq s \leq T$ y $0 \leq \tau \leq T$. Sea $Z \in L^\infty(\mathcal{F}_\tau)$ y $\lambda = (\lambda_t) \in \mathcal{M}_S^T$ definido por $\lambda_t = Z \mathbb{I}_{\{t \geq s\}}$. Si $t \geq s - 1$,

$$\begin{aligned}\psi_t^{\max}(L + \lambda) &= \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i + Z | \mathcal{F}_t]\} = \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + Z\} \\ &= \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t]\} + Z = \psi_t^{\max}(L) + Z, \text{ c.s.}\end{aligned}$$

Si $\tau \leq t < s - 1$,

$$\begin{aligned}\psi_t^{\max}(L + \lambda) &= \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i + Z \mathbb{I}_{\{i \geq s\}} | \mathcal{F}_t]\} \\ &= \max_{i=t+1, \dots, T} \begin{cases} \mathbb{E}_{\mathbb{Q}}[L_i + Z | \mathcal{F}_t] & \text{si } i \geq s \\ \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] & \text{si } i < s. \end{cases}\end{aligned}$$

Entonces, si para $\omega \in \Omega$ el argumento máximo es $j < s$, se tiene que

$$\begin{aligned}\psi_t^{\max}(L + \lambda)(\omega) &= \mathbb{E}_{\mathbb{Q}}[L_j | \mathcal{F}_t](\omega) \\ &= \max \{\mathbb{E}_{\mathbb{Q}}[L_{t+1} | \mathcal{F}_t], \dots, \mathbb{E}_{\mathbb{Q}}[L_{s-1} | \mathcal{F}_t]\}(\omega) = \psi_t^{\max}(\pi_{0, s-1}(L))(\omega).\end{aligned}$$

Si el argumento máximo es $j \geq s$, se tiene que

$$\begin{aligned}\psi_t^{\max}(L + \lambda)(\omega) &= \mathbb{E}_{\mathbb{Q}}[L_j + Z \mid \mathcal{F}_t](\omega) = \left(\max_{i=s, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i + Z \mid \mathcal{F}_t]\} \right) (\omega) \\ &= \left(\max_{i=s, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t]\} + Z \right) (\omega) = (\psi_t^{\max}(\pi_{s,T}(L)) + Z) (\omega).\end{aligned}$$

Homogeneidad positiva dinámica fuerte: Suponga que $0 \leq \tau \leq s$ y $\delta = (\delta_t)$ está dado por $\delta_t = d \mathbb{1}_{\{t \geq s\}} + \mathbb{1}_{\{t < s\}}$, para $d \in L^\infty(\mathcal{F}_\tau)$, $d \geq 0$. Si $t \geq s - 1$, entonces

$$\psi_t^{\max}(\delta L) = \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[d L_i \mid \mathcal{F}_t]\} = d \max_{i=t+1, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t]\} = d \psi_t^{\max}(L).$$

Si $\tau \leq t < s - 1$,

$$\psi_t^{\max}(\delta L) = \max_{i=t+1, \dots, T} \begin{cases} \mathbb{E}_{\mathbb{Q}}[d L_i \mid \mathcal{F}_t] & \text{si } i \geq s \\ \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] & \text{si } i < s. \end{cases}$$

Entonces, si para $\omega \in \Omega$ el argumento mínimo j es tal que $j < s$,

$$\begin{aligned}\psi_t^{\max}(d L)(\omega) &= \mathbb{E}_{\mathbb{Q}}[L_j \mid \mathcal{F}_t](\omega) \\ &= \max \{\mathbb{E}_{\mathbb{Q}}[L_{t+1} \mid \mathcal{F}_t], \dots, \mathbb{E}_{\mathbb{Q}}[L_{s-1} \mid \mathcal{F}_t]\} (\omega) = \psi_t^{\max}(\pi_{0,s-1}(L)) (\omega).\end{aligned}$$

Si el argumento máximo es $j \geq s$,

$$\begin{aligned}\psi_t^{\max}(d L)(\omega) &= \mathbb{E}_{\mathbb{Q}}[d L_j \mid \mathcal{F}_t](\omega) = \left(\max_{i=s, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[d L_i \mid \mathcal{F}_t]\} \right) (\omega) \\ &= \left(d \max_{i=s, \dots, T} \{\mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t]\} \right) (\omega) = (d \psi_t(\pi_{s,T}(L))) (\omega). \quad \square\end{aligned}$$

Un caso particular de esta medida será utilizado para optimizar portafolios de una manera dinámica, pero se verá en la Sección 4.3 que esta medida por si sola puede no ser apropiada para esto. Por esto se presenta la siguiente medida dinámica de riesgo.

Definición 3.3.2. Sean $\lambda_i^t \geq 0$, $i, t = 0, 1, \dots, T$, tales que para todo t , $\sum_{i=t+1}^T \lambda_i^t = 1$. Se define la medida dinámica $\bar{\psi} = (\bar{\psi}_t)$ por

$$\bar{\psi}_t(L) = \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t]. \quad (21)$$

Observe que para todo $L \in \mathcal{M}_S^T$ y para todo i , $\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t]$ es \mathcal{F}_t -medible, luego $\bar{\psi}_t(L)$ lo es. Además $\bar{\psi}_t(L) \in \mathcal{R}^\infty$, porque para todo $i = t+1, \dots, T$, $\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] \in \mathcal{R}^\infty$. Entonces la medida está bien definida. Se nota $\bar{\psi}$ ya que es una ponderación de los valores de la medida estática en las diferentes fechas. A continuación se verifican las propiedades de $\bar{\psi}$.

Proposición 3.3.2. *La medida dinámica $\bar{\psi}$ de la Definición 3.3.2 es coherente. $\bar{\psi}$ cumple el axioma de consistencia temporal (Axioma 15) si y solo si $\forall t, \forall s > t, \forall j \geq s, \lambda_j^s \sum_{i=s+1}^T \lambda_i^t = \lambda_j^t$.*

$\bar{\psi}$ no cumple en general con los axiomas de Invarianza dinámica fuerte bajo traslación (Axioma 12) y Homogeneidad positiva fuerte (Axioma 14) (no es fuertemente coherente).

Demostración. Sean $L, L' \in \mathcal{M}_S^T$ y $0 \leq t \leq s \leq T$.

Independencia del pasado: Sea $A \in \mathcal{F}_t$, trivialmente $\forall u, \bar{\psi}_u(0) = 0$. Entonces se verificará que $\bar{\psi}_s(L \mathbb{I}_A) = \bar{\psi}_s(L) \mathbb{I}_A$, c.s. \mathbb{I}_A es \mathcal{F}_t -medible, luego

$$\begin{aligned} \bar{\psi}_s(L) \mathbb{I}_A &= \mathbb{I}_A \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] = \sum_{i=t+1}^T \lambda_i^t (\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] \mathbb{I}_A) \\ &= \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mathbb{I}_A | \mathcal{F}_t] = \bar{\psi}_s(L \mathbb{I}_A), \text{ c.s.} \end{aligned}$$

Monotonidad dinámica: Suponga $L_u \geq L'_u$ c.s para todo $u \geq s$. Como $\forall i, \lambda_i^t \geq 0$ se tiene que

$$\bar{\psi}_t(L) = \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] \geq \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L'_i | \mathcal{F}_t] = \bar{\psi}_t(L').$$

Subaditividad dinámica:

$$\begin{aligned} \bar{\psi}_t(L + L') &= \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i + L'_i | \mathcal{F}_t] = \sum_{i=t+1}^T \lambda_i^t (\mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + \mathbb{E}_{\mathbb{Q}}[L'_i | \mathcal{F}_t]) \\ &= \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L'_i | \mathcal{F}_t] = \bar{\psi}_t(L) + \bar{\psi}_t(L'), \text{ c.s.} \end{aligned}$$

Invarianza dinámica bajo traslación: Sean $s, \tau \in \mathbb{N}$, con $S \leq s \leq T$ y $0 \leq \tau \leq T$. Sea $Z \in L^\infty(\mathcal{F}_\tau)$ y $\lambda = (\lambda_t) \in \mathcal{M}_S^T$ definido por $\lambda_t = Z \mathbb{I}_{\{t \geq s\}}$. Si $t \geq s - 1$,

$$\begin{aligned} \bar{\psi}_t(L + \lambda) &= \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i + Z \mid \mathcal{F}_t] = \sum_{i=t+1}^T \lambda_i^t (\mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] + Z) \\ &= \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] + Z \underbrace{\sum_{i=t+1}^T \lambda_i^t}_{=1} = \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] + Z = \bar{\psi}_t(L) + Z. \end{aligned}$$

Se observa que si $\tau \leq t < s - 1$ y $\lambda_{t+1}^t > 0$, como $t + 1 < s$ y entonces $\sum_{i=s}^T \lambda_i^t < 1$, por lo tanto

$$\begin{aligned} \bar{\psi}_t(L + \lambda) &= \sum_{i=t+1}^{s-1} \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] + \sum_{i=s}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i + Z \mid \mathcal{F}_t] \\ &= \sum_{i=t+1}^{s-1} \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] + \sum_{i=s}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] + Z \sum_{i=s}^T \lambda_i^t < \bar{\psi}_t(L) + Z. \end{aligned}$$

Entonces, $\bar{\psi}$ no cumple en general el axioma de invarianza fuerte.

Homogeneidad positiva dinámica: Suponga que $0 \leq \tau \leq s$ y $\delta = (\delta_t)$ es está dado por $\delta_t = d \mathbb{I}_{\{t \geq s\}} + \mathbb{I}_{\{t < s\}}$, para $d \in L^\infty(\mathcal{F}_\tau)$, $d \geq 0$. Si $t \geq s - 1$, entonces

$$\bar{\psi}_t(\delta L) = \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[d L_i \mid \mathcal{F}_t] = d \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] = d \bar{\psi}_t(L).$$

Si $\tau \leq t < s - 1$ y $\lambda_{t+1}^t > 0$ y $d \neq 1$, como $t + 1 < s$,

$$\begin{aligned} \bar{\psi}_t(\delta L) &= \sum_{i=t+1}^{s-1} \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] + \sum_{i=s}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[d L_i \mid \mathcal{F}_t] = \\ &= \sum_{i=t+1}^{s-1} \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] + d \sum_{i=s}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] \neq d \left(\sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t] \right) = d \bar{\psi}_t(L). \end{aligned}$$

Por lo tanto, $\bar{\psi}$ no cumple en general el axioma de homogeneidad positiva fuerte.

Consistencia temporal: Se tiene que $\mathbb{E}_{\mathbb{Q}}[\mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_s] \mid \mathcal{F}_t] = \mathbb{E}_{\mathbb{Q}}[L_i \mid \mathcal{F}_t]$. Suponga que $\bar{\psi}$ cumple el axioma de consistencia temporal, entonces si $\Lambda_s^t \triangleq \sum_{i=s+1}^T \lambda_i^t$,

se tiene que

$$\begin{aligned}
 \bar{\psi}_t(L \mathbb{I}_{[0,s]} + \bar{\psi}_s(L) \mathbb{I}_{(s,\infty)}) &= \sum_{i=t+1}^s \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + \sum_{i=s+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[\bar{\psi}_s(L) | \mathcal{F}_t] \\
 &= \sum_{i=t+1}^s \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + \sum_{i=s+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}} \left[\sum_{j=s+1}^T \lambda_j^s \mathbb{E}_{\mathbb{Q}}[L_j | \mathcal{F}_s] \middle| \mathcal{F}_t \right] \\
 &= \sum_{i=t+1}^s \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + \Lambda_s^t \sum_{j=s+1}^T \lambda_j^s \mathbb{E}_{\mathbb{Q}}[\mathbb{E}_{\mathbb{Q}}[L_j | \mathcal{F}_s] | \mathcal{F}_t] \quad (22) \\
 &= \sum_{i=t+1}^s \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + \sum_{j=s+1}^T \Lambda_s^t \lambda_j^s \mathbb{E}_{\mathbb{Q}}[L_j | \mathcal{F}_t] \\
 &= \sum_{i=t+1}^s \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] + \sum_{i=s+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t] = \bar{\psi}_t(L).
 \end{aligned}$$

La penúltima igualdad se tiene por el supuesto que $\bar{\psi}$ cumple el axioma. Sea $j > s$. Tomando en particular $L = (L_t)$, con $L_t = 0$ si $t \neq j$ y $L_j = 1$, se tiene

$$\bar{\psi}_t(L \mathbb{I}_{[0,s]} + \bar{\psi}_s(L) \mathbb{I}_{(s,\infty)}) = \Lambda_s^t \lambda_j^s = \lambda_j^t = \bar{\psi}_t(L).$$

Luego, como $s > t$ es arbitrario se cumple que

$$\forall t, \forall s > t, \forall j > s, \lambda_j^s \left(\sum_{i=s+1}^T \lambda_i^t \right) = \lambda_j^t. \quad (23)$$

Además, si esto se cumple $\bar{\psi}$ es consistente temporalmente por (22). \square

Por el resultado anterior, si se desea que $\bar{\psi}$ sea consistente temporalmente, se deben construir adecuadamente las ponderaciones λ_i^t . Por ejemplo, dada una ponderación de arbitraria para $\bar{\psi}_0$, tal que $\lambda_1^0, \dots, \lambda_T^0 \geq 0$, $\lambda_T^0 > 0$ y $\sum_{i=1}^T \lambda_i^0 = 1$ se construyen ponderaciones $\lambda_{i+1}^t, \dots, \lambda_T^t \geq 0$, recursivamente con (23). Esto es posible porque si $t_1 < t_2$, $s > t_2$ y $j > s$,

$$\forall i \geq t_1, \lambda_i^{t_2} = \frac{\lambda_i^{t_1}}{\sum_{k=t_2+1}^T \lambda_k^{t_1}} \Rightarrow \lambda_j^s = \frac{\lambda_j^{t_2}}{\sum_{i=s+1}^T \lambda_i^{t_2}} = \frac{\left(\lambda_j^{t_1} / \sum_{k=t_2+1}^T \lambda_k^{t_1} \right)}{\sum_{i=s+1}^T \left(\lambda_i^{t_1} / \sum_{k=t_2+1}^T \lambda_k^{t_1} \right)} = \frac{\lambda_j^{t_1}}{\sum_{i=s+1}^T \lambda_i^{t_1}}.$$

Esta medida no es entonces en general fuertemente coherente, pero se utilizará para la optimización de portafolios por razones que quedarán claras en la Sección 4.3.4.

No es la intención de este trabajo presentar un teorema de caracterización de las medidas *coherentes* dinámicas, pero se tiene que en la fecha 0 estas dos últimas medidas definidas medidas (es decir, $\psi_0^{\text{máx}}$ y $\bar{\psi}_0$) son casos particulares de una caracterización presentada en Artzner et al. (2003). En este trabajo los autores caracterizan las medidas dinámicas de riesgo en la fecha 0 para medidas que cumplan con los axiomas de Monotonidad dinámica (Axioma 7), Subaditividad dinámica (Axioma 8), Invarianza dinámica débil bajo traslación (Axioma 9) y Homogeneidad positiva dinámica débil (Axioma 13).⁸ Una medida que cumpla estos axiomas se dirá que es *coherente* en el sentido *débil* o *débilmente coherente*. Entonces, se presenta a continuación el resultado de caracterización por completitud, sin demostración (ver Artzner et al. (2003) para una demostración) y se muestra por qué las medidas anteriormente definidas en la fecha inicial son casos particulares.

Proposición 3.3.3. *Para toda medida dinámica ψ débilmente coherente, existe un conjunto \mathcal{A} de procesos positivos, no decrecientes, tales que para todo $A \in \mathcal{A}$, $\mathbb{E}[A_T] = 1$ y $A_{-1} = 0$ y para todo $L \in \mathcal{M}$ se cumple que*

$$\psi_0(L) = \sup_{A \in \mathcal{A}} \mathbb{E} \left[\sum_{t=0}^T L_t (A_t - A_{t-1}) \right]. \quad (24)$$

Se verá entonces que $\psi_0^{\text{máx}}$ y $\bar{\psi}_0$ son efectivamente casos particulares de la siguiente manera. Tome \mathbb{Q} como a lo largo de esta sección. Sea $\mathcal{A} = \{A^1, A^2, \dots, A^T\}$, donde cada $A^i = (A_t^i)$ está definido por $A_t^i = 0$, si $t < i$ y $A_t^i = \frac{d\mathbb{Q}}{d\mathbb{P}}$, si $t \geq i$. Se cumple que para todo $i = 1, \dots, T$, $\mathbb{E}[A_T^i] = \mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}}] = \int_{\Omega} \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} = \int_{\Omega} d\mathbb{Q} = \mathbb{E}_{\mathbb{Q}}[1] = 1$ y

$$\begin{aligned} \sup_{A \in \mathcal{A}} \mathbb{E} \left[\sum_{t=0}^T L_t (A_t - A_{t-1}) \right] &= \max_{i=1, \dots, T} \mathbb{E} \left[\sum_{t=0}^T L_t (A_t^i - A_{t-1}^i) \right] \\ &= \max_{i=1, \dots, T} \mathbb{E} \left[L_i \frac{d\mathbb{Q}}{d\mathbb{P}} \right] = \max_{i=1, \dots, T} \mathbb{E}_{\mathbb{Q}} [L_i] = \max_{i=1, \dots, T} \mathbb{E}_{\mathbb{Q}} [L_i | \mathcal{F}_0] = \psi_0^{\text{máx}}(L). \end{aligned}$$

⁸Para estos autores las medidas *coherentes* dinámicas son las que cumplen estos axiomas

Ahora, sea $\mathcal{A} = \{A\}$, donde $A = (A_t)$ está definido por $A_{-1} = A_0 = 0$ y $A_t = \lambda_t^0 \frac{dQ}{dP} + A_{t-1}$, para $t > 0$, con λ_t^0 como en la Definición 3.3.2. Se cumple que

$$\begin{aligned} \mathbb{E}[A_T] &= \mathbb{E}\left[\lambda_T^0 \frac{dQ}{dP} + A_{T-1}\right] = \mathbb{E}\left[\lambda_T^0 \frac{dQ}{dP} + \lambda_{T-1}^0 \frac{dQ}{dP} + A_{T-2}\right] = \dots \\ &= \mathbb{E}\left[\frac{dQ}{dP} \sum_{t=1}^T \lambda_t^0\right] = \sum_{t=1}^T \lambda_t^0 \mathbb{E}\left[\frac{dQ}{dP}\right] = \sum_{t=1}^T \lambda_t^0 \mathbb{E}_{\mathbb{Q}}[1] = \sum_{t=1}^T \lambda_t^0 = 1 \end{aligned}$$

y, como se quería probar,

$$\begin{aligned} \sup_{A \in \mathcal{A}} \mathbb{E}\left[\sum_{t=0}^T L_t(A_t - A_{t-1})\right] &= \mathbb{E}\left[\sum_{t=0}^T L_t(A_t - A_{t-1})\right] \\ &= \mathbb{E}\left[\sum_{t=1}^T L_t \lambda_t^0 \frac{dQ}{dP}\right] = \sum_{t=1}^T \lambda_t^0 \mathbb{E}\left[L_t \frac{dQ}{dP}\right] = \sum_{t=1}^T \lambda_t^0 \mathbb{E}_{\mathbb{Q}}[L_t | \mathcal{F}_0] = \bar{\psi}_0(L). \end{aligned}$$

Capítulo IV

OPTIMIZACIÓN DE PORTAFOLIOS

4.1. Optimización con restricciones de CVaR

4.1.1. Enfoque

Se expone a continuación el enfoque para optimizar o cubrir un portafolio, propuesto en Uryasev (2000), Andersson et al. (2001) y Rockafellar y Uryasev (2000 y 2002). Se busca en dicho enfoque minimizar el CVaR_α del portafolio, lo que también reduce en general el VaR_α del mismo. Esto es apropiado porque, como se ha visto, la medida de riesgo CVaR_α es una medida estática coherente de riesgo, a diferencia del VaR_α . Además, el enfoque propuesto encuentra el portafolio con mínimo CVaR_α y simultáneamente calcula el VaR_α del mismo. Más aún, al combinarlo con métodos basados en escenarios, se puede llevar el problema de optimización a un problema de programación lineal.

Suponga que se utiliza el subconjunto $\{L_1, \dots, L_m\} \subseteq \mathcal{M}$ de riesgos financieros para construir lo que se llamará en general un portafolio (ver ejemplos en la Sección 4.1.2). Sea $\mathbf{L} = (L_1, \dots, L_m)$ y sea $\mathbf{x} \in X \subseteq \mathbb{R}^n$ un vector de decisión, para algún subconjunto X de \mathbb{R}^n que representa las restricciones al vector de decisión. Además, suponga que las pérdidas del portafolio están dadas por la función $\mathbb{L}(\mathbf{x}, \mathbf{L})$. Si se conoce la distribución de probabilidad del vector aleatorio \mathbf{L} , entonces la variable $Z_{\mathbf{x}} = \mathbb{L}(\mathbf{x}, \mathbf{L})$ es una variable aleatoria con distribución dependiente de \mathbf{x} . Por esto, para escoger un valor de \mathbf{x} se debe tener en cuenta el valor esperado de $Z_{\mathbf{x}}$, tanto como el riesgo de ésta.

Para esto se puede utilizar como medida de riesgo la varianza del portafolio, como lo hace Markowitz (1952). También se puede utilizar la medida de riesgo VaR_α (para

un nivel de confianza α), que es una medida bastante utilizada en la práctica. No obstante, dicha medida no es coherente y no mide la severidad de los peores casos. Por esto se puede utilizar la medida de riesgo CVaR_α que es coherente, como se ha visto.

Suponga que la medida de probabilidad \mathbb{P} y \mathbf{x} son independientes, que $\mathbb{L}(\mathbf{x}, \mathbf{L})$ es continua en \mathbf{x} y medible en \mathbf{L} y que $\mathbb{E}[|\mathbb{L}(\mathbf{x}, \mathbf{L})|] < \infty$ para todo $\mathbf{x} \in X$. Es decir, para cada $\mathbf{x} \in X$, $\mathbb{L}(\mathbf{x}, \mathbf{L}) \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, por lo que $\mathbb{L}(\mathbf{x}, \mathbf{L}) \in \mathcal{M}$ (por definición de \mathcal{M}). Sea $f(\mathbf{L})$ la función de densidad de probabilidad de \mathbf{L} . Para cada $\mathbf{x} \in X$ se denota por $\Psi(\mathbf{x}, \cdot)$ la distribución de la pérdida $Z_{\mathbf{x}} = \mathbb{L}(\mathbf{x}, \mathbf{L})$. Esto es

$$\begin{aligned} \Psi(\mathbf{x}, \beta) &= \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta] = \mathbb{E}[\mathbb{I}_{\{\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta\}}] \\ &= \int_{\{\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta\}} d\mathbb{P} = \int_{\{\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta\}} f(\mathbf{L}) d\mathbf{L}. \end{aligned} \quad (25)$$

Se denota $\Psi(\mathbf{x}, \beta^-) \triangleq \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) < \beta]$ y se dice que existe un *átomo de probabilidad* en β , si $\Psi(\mathbf{x}, \beta) - \Psi(\mathbf{x}, \beta^-) = \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) = \beta] > 0$.

Para $r = 0$, el VaR_α (con $\alpha \in (0, 1)$) para la variable $Z_{\mathbf{x}}$ es

$$\begin{aligned} \text{VaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L})) &= \inf\{\beta \in \mathbb{R} \mid \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta] \geq \alpha\} \\ &= \min\{\beta \in \mathbb{R} \mid \Psi(\mathbf{x}, \beta) \geq \alpha\}. \end{aligned}$$

Se alcanza el mínimo porque $\Psi(\mathbf{x}, \beta)$ es continua por la derecha y no decreciente con respecto a β .

Observe que si $\Psi(\mathbf{x}, \cdot)$ es continua y estrictamente creciente, $\text{VaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L}))$ es igual al único β que satisface $\Psi(\mathbf{x}, \beta) = \alpha$. En otro caso esta igualdad puede no tener una solución o puede tener un rango de soluciones.

Se mostrará a continuación cómo el VaR_α y CVaR_α pueden ser calculados simultáneamente al resolver un problema de optimización, que en algunos casos es un problema elemental de optimización convexa y se puede aproximar con un problema lineal. Para esto se define la función (Rockafellar y Uryasev, 2000),

$$F_\alpha(\mathbf{x}, \beta) \triangleq \beta + \frac{1}{1 - \alpha} \mathbb{E} [[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+], \quad \text{donde } [t]^+ \triangleq \max\{0, t\}. \quad (26)$$

Teorema 4.1.1. Como función de $\beta \in \mathbb{R}$, $F_\alpha(\mathbf{x}, \beta)$ es finita, convexa (y por lo tanto continua), con

$$\text{CVaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L})) = \min_{\beta} F_\alpha(\mathbf{x}, \beta). \quad (27)$$

Más aun, el conjunto $\text{argmin}_{\beta} F_\alpha(\mathbf{x}, \beta)$ es no vacío, $\text{VaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L}))$ es el mínimo β del conjunto $\text{argmin}_{\beta} F_\alpha(\mathbf{x}, \beta)$ y

$$\text{CVaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L})) = F_\alpha(\mathbf{x}, \text{VaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L}))). \quad (28)$$

Demostración. Como función de β , $F_\alpha(\mathbf{x}, \beta)$ es finita. Esto es porque para todo \mathbf{x} , $\mathbb{E}[|\mathbb{L}(\mathbf{x}, \mathbf{L})|] < \infty$, lo que implica

$$\begin{aligned} F_\alpha(\mathbf{x}, \beta) &\triangleq \beta + \frac{1}{1-\alpha} \mathbb{E} [|\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta|^+] \\ &= \beta + \frac{1}{1-\alpha} \int_{\{\mathbb{L}(\mathbf{x}, \mathbf{L}) \geq \beta\}} \mathbb{L}(\mathbf{x}, \mathbf{L}) d\mathbb{P} - \frac{1}{1-\alpha} \int_{\{\mathbb{L}(\mathbf{x}, \mathbf{L}) \geq \beta\}} \beta d\mathbb{P} \\ &\leq \beta + \frac{1}{1-\alpha} \int_{\{\mathbb{L}(\mathbf{x}, \mathbf{L}) \geq \beta\}} |\mathbb{L}(\mathbf{x}, \mathbf{L})| d\mathbb{P} - \frac{\beta}{1-\alpha} \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) \geq \beta] < \infty. \end{aligned}$$

Como $|\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta|^+$ es convexa con respecto a β , entonces $F_\alpha(\mathbf{x}, \beta)$ también lo es. Por esto, siendo $F_\alpha(\mathbf{x}, \cdot)$ una función finita y convexa, existen sus derivadas por la derecha y la izquierda en cualquier $\beta \in \mathbb{R}$ ($\frac{\partial^+ F_\alpha(\mathbf{x}, \beta)}{\partial \beta}$ y $\frac{\partial^- F_\alpha(\mathbf{x}, \beta)}{\partial \beta}$, respectivamente).

Más aún, se demostrará a continuación que

$$\frac{\partial^+ F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = \frac{\Psi(\mathbf{x}, \beta) - \alpha}{1-\alpha} \quad \text{y} \quad \frac{\partial^- F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = \frac{\Psi(\mathbf{x}, \beta^-) - \alpha}{1-\alpha}. \quad (29)$$

Se tiene que

$$\begin{aligned} \frac{F_\alpha(\mathbf{x}, \beta') - F_\alpha(\mathbf{x}, \beta)}{\beta' - \beta} &= \frac{\beta' + \frac{1}{1-\alpha} \mathbb{E} [|\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta'|^+] - \beta - \frac{1}{1-\alpha} \mathbb{E} [|\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta|^+]}{\beta' - \beta} \\ &= 1 + \frac{1}{1-\alpha} \mathbb{E} \left[\frac{|\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta'|^+ - |\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta|^+}{\beta' - \beta} \right]. \end{aligned}$$

Ahora, para verificar la primera parte de (29), suponga $\beta' > \beta$ dados, entonces

$$\frac{|\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta'|^+ - |\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta|^+}{\beta' - \beta} \begin{cases} = -1 & \text{si } \mathbb{L}(\mathbf{x}, \mathbf{L}) \geq \beta' \\ = 0 & \text{si } \mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta \\ \in (-1, 0) & \text{si } \beta < \mathbb{L}(\mathbf{x}, \mathbf{L}) < \beta'. \end{cases}$$

Se sigue que

$$\begin{aligned}
 & \mathbb{E} \left[\frac{[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta']^+ - [\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+}{\beta' - \beta} \right] \\
 &= - \int_{\{\mathbb{L}(\mathbf{x}, \mathbf{L}) > \beta'\}} d\mathbb{P} + \int_{\{\beta < \mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta'\}} \frac{[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta']^+ - [\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+}{\beta' - \beta} d\mathbb{P} \\
 &= - (1 - \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta']) - \int_{\{\beta < \mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta'\}} \frac{\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta}{\beta' - \beta} d\mathbb{P} \\
 &= - [1 - \Psi(\mathbf{x}, \beta')] - \rho(\beta', \beta) [\Psi(\mathbf{x}, \beta') - \Psi(\mathbf{x}, \beta)],
 \end{aligned}$$

para algún $\rho(\beta', \beta) \in (0, 1)$ (que depende de β y β'). Dicho $\rho(\beta', \beta)$ existe porque

$$\begin{aligned}
 0 &\leq \int_{\{\beta < \mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta'\}} \frac{\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta}{\beta' - \beta} d\mathbb{P} \leq \int_{\{\beta < \mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta'\}} 1 d\mathbb{P} = \mathbb{P}[\beta < \mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta'] \\
 &= \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta'] - \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta] = \Psi(\mathbf{x}, \beta') - \Psi(\mathbf{x}, \beta).
 \end{aligned}$$

Además, si $\beta' \searrow \beta$ (esto es, $\beta' \rightarrow \beta$ y $\beta' > \beta$) si tiene que $\Psi(\mathbf{x}, \beta') \searrow \Psi(\mathbf{x}, \beta)$, luego

$$\lim_{\beta' \searrow \beta} \mathbb{E} \left[\frac{[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta']^+ - [\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+}{\beta' - \beta} \right] = - [1 - \Psi(\mathbf{x}, \beta)].$$

Por lo tanto,

$$\lim_{\beta' \searrow \beta} \frac{F_\alpha(\mathbf{x}, \beta') - F_\alpha(\mathbf{x}, \beta)}{\beta' - \beta} = 1 + \frac{1}{1 - \alpha} (- [1 - \Psi(\mathbf{x}, \beta)]) = \frac{\Psi(\mathbf{x}, \beta) - \alpha}{1 - \alpha}.$$

Similarmente, si $\beta' < \beta$, entonces

$$\frac{[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta']^+ - [\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+}{\beta' - \beta} \begin{cases} = -1 & \text{si } \mathbb{L}(\mathbf{x}, \mathbf{L}) \geq \beta \\ = 0 & \text{si } \mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta' \\ \in (-1, 0) & \text{si } \beta' < \mathbb{L}(\mathbf{x}, \mathbf{L}) < \beta, \end{cases}$$

y por razones similares, dado que $\Psi(\mathbf{x}, \beta') \nearrow \Psi(\mathbf{x}, \beta^-)$ si $\beta' \nearrow \beta$ (esto es, $\beta' \rightarrow \beta$ y $\beta' < \beta$), se sigue que

$$\lim_{\beta' \nearrow \beta} \mathbb{E} \left[\frac{[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta']^+ - [\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+}{\beta' - \beta} \right] = - [1 - \Psi(\mathbf{x}, \beta^-)].$$

Entonces, finalmente esto implica que

$$\lim_{\beta' \nearrow \beta} \frac{F_\alpha(\mathbf{x}, \beta') - F_\alpha(\mathbf{x}, \beta)}{\beta' - \beta} = 1 + \frac{1}{1 - \alpha} (- [1 - \Psi(\mathbf{x}, \beta^-)]) = \frac{\Psi(\mathbf{x}, \beta^-) - \alpha}{1 - \alpha}.$$

Ahora, por lo anterior se tiene que

$$\lim_{\beta \rightarrow -\infty} \frac{\partial^+ F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = \lim_{\beta \rightarrow -\infty} \frac{\partial^- F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = \frac{-\alpha}{1 - \alpha}$$

y

$$\lim_{\beta \rightarrow \infty} \frac{\partial^+ F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = \lim_{\beta \rightarrow \infty} \frac{\partial^- F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = 1.$$

Además, dichas derivadas por la izquierda y derecha son no decrecientes por la convexidad de $F_\alpha(\mathbf{x}, \beta)$.

Sean $A_c \triangleq \{\beta \in \mathbb{R} \mid F_\alpha(\mathbf{x}, \beta) \leq c\}$ (imagen inversa del conjunto $(-\infty, c]$ por la función continua $F_\alpha(\mathbf{x}, \cdot)$), para $c \in \mathbb{R}$. Entonces los conjuntos A_c son cerrados. Además son compactos porque son acotados. En efecto, como $\lim_{\beta \rightarrow \infty} \frac{\partial^- F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = 1$, existe un $\beta^* \in \mathbb{R}$, tal que para todo $\beta \geq \beta^*$, $\frac{\partial^+ F_\alpha(\mathbf{x}, \beta)}{\partial \beta} > 0,5$ y $\frac{\partial^- F_\alpha(\mathbf{x}, \beta)}{\partial \beta} > 0,5$, por lo que la función $F_\alpha(\mathbf{x}, \cdot)$ es estrictamente creciente para $\beta > \beta^*$ y entonces para todo $c \in \mathbb{R}$ existe un $\hat{\beta}$, tal que para todo $\beta > \hat{\beta}$, $F_\alpha(\mathbf{x}, \beta) > c$, por lo que $\hat{\beta}$ es una cota superior. Con un argumento similar se puede demostrar que para todo c existe una cota inferior $\check{\beta} \in \mathbb{R}$.

Para algún $c \in \mathbb{R}$, $A_c \neq \emptyset$, entonces la función $F_\alpha(\mathbf{x}, \cdot)$, restringida sobre dicho A_c alcanza el mínimo y el mínimo es el mínimo de la función sobre todo \mathbb{R} . Entonces en el problema de minimización (27) se alcanza el mínimo o $\text{argmin}_\beta F_\alpha(\mathbf{x}, \beta)$ es no vacío. Dicho mínimo puede no ser único, pero los elementos de éste están caracterizados porque cumplen con

$$\frac{\partial^- F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = \frac{\Psi(\mathbf{x}, \beta^-) - \alpha}{1 - \alpha} \leq 0 \leq \frac{\partial^+ F_\alpha(\mathbf{x}, \beta)}{\partial \beta} = \frac{\Psi(\mathbf{x}, \beta) - \alpha}{1 - \alpha},$$

por lo que

$$\Psi(\mathbf{x}, \beta^-) = \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) < \beta] \leq \alpha \leq \Psi(\mathbf{x}, \beta) = \mathbb{P}[\mathbb{L}(\mathbf{x}, \mathbf{L}) \leq \beta], \quad (30)$$

para todo $\beta \in \text{argmin}_\beta F_\alpha(\mathbf{x}, \beta)$. Entonces, por definición el mínimo β que cumple (30) es igual a $\text{VaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L})) = q_\alpha^-(\mathbb{L}(\mathbf{x}, \mathbf{L}))$. Es decir, el α -cuantil más pequeño de $\mathbb{L}(\mathbf{x}, \mathbf{L})$ (recordando que $r = 0$).

Ahora, denotando de nuevo $Z_{\mathbf{x}} = \mathbb{L}(\mathbf{x}, \mathbf{L})$ y $q_{\alpha}^{-} = q_{\alpha}^{-}(\mathbb{L}(\mathbf{x}, \mathbf{L}))$, se tiene en particular que,

$$\begin{aligned}
 \min_{\beta} F_{\alpha}(\mathbf{x}, \beta) &= F_{\alpha}(\mathbf{x}, \text{VaR}_{\alpha}(\mathbb{L}(\mathbf{x}, \mathbf{L}))) = F_{\alpha}(\mathbf{x}, q_{\alpha}^{-}) \\
 &= q_{\alpha}^{-} + \frac{1}{1-\alpha} \mathbb{E} [[Z_{\mathbf{x}} - q_{\alpha}^{-}]^{+}] \\
 &= q_{\alpha}^{-} + \frac{1}{1-\alpha} \int_{\{Z_{\mathbf{x}} > q_{\alpha}^{-}\}} [Z_{\mathbf{x}} - q_{\alpha}^{-}] d\mathbb{P} \\
 &= q_{\alpha}^{-} + \frac{1}{1-\alpha} \left(\int_{\{Z_{\mathbf{x}} > q_{\alpha}^{-}\}} Z_{\mathbf{x}} d\mathbb{P} - q_{\alpha}^{-} \int_{\{Z_{\mathbf{x}} > q_{\alpha}^{-}\}} d\mathbb{P} \right) \\
 &= \frac{1}{1-\alpha} \left(\int_{\{Z_{\mathbf{x}} > q_{\alpha}^{-}\}} Z_{\mathbf{x}} d\mathbb{P} + (1-\alpha - \mathbb{P}[Z_{\mathbf{x}} > q_{\alpha}^{-}]) q_{\alpha}^{-} \right) \\
 &= \text{CVaR}_{\alpha}(Z_{\mathbf{x}}) = \text{CVaR}_{\alpha}(\mathbb{L}(\mathbf{x}, \mathbf{L})).
 \end{aligned}$$

La penúltima igualdad es por la Definición 2.4.2, ecuación (4). \square

Utilizando esto anterior se puede entonces formular el siguiente Teorema que permite optimizar el CVaR_{α} de un portafolio y encontrar simultáneamente el VaR_{α} en ese óptimo.

Teorema 4.1.2. *Para $X \subseteq \mathbb{R}^n$ un conjunto de restricciones dado, los dos siguientes problemas son equivalentes:*

$$\min_{\mathbf{x} \in X} \text{CVaR}_{\alpha}(\mathbb{L}(\mathbf{x}, \mathbf{L})) \quad y \quad \min_{(\mathbf{x}, \beta) \in X \times \mathbb{R}} F_{\alpha}(\mathbf{x}, \beta). \quad (31)$$

Más aún,

$$(\mathbf{x}^*, \beta^*) \in \underset{(\mathbf{x}, \beta) \in X \times \mathbb{R}}{\text{argmin}} F_{\alpha}(\mathbf{x}, \beta) \quad \Leftrightarrow \quad \begin{aligned} \mathbf{x}^* &\in \underset{\mathbf{x} \in X}{\text{argmin}} \text{CVaR}_{\alpha}(\mathbb{L}(\mathbf{x}, \mathbf{L})), \\ \beta^* &\in \underset{\beta \in \mathbb{R}}{\text{argmin}} F_{\alpha}(\mathbf{x}^*, \beta). \end{aligned}$$

Demostración. El problema $\min_{\mathbf{x} \in X} \text{CVaR}_{\alpha}(\mathbb{L}(\mathbf{x}, \mathbf{L})) = \min_{\mathbf{x} \in X} (\min_{\beta \in \mathbb{R}} F_{\alpha}(\mathbf{x}, \beta))$ es equivalente al problema $\min_{(\mathbf{x}, \beta) \in X \times \mathbb{R}} F_{\alpha}(\mathbf{x}, \beta)$ porque se puede encontrar el óptimo en el segundo minimizando, para cada $\mathbf{x} \in X$ fijo, con respecto a $\beta \in \mathbb{R}$ y luego minimizando con respecto a \mathbf{x} .

Dado el Teorema 2.4.2, la conclusión “ \Leftarrow ” es inmediata. Para la dirección “ \Rightarrow ” se asegura por el mismo Teorema que el conjunto $\underset{\beta \in \mathbb{R}}{\text{argmin}} F_{\alpha}(\mathbf{x}^*, \beta)$ es no vacío. \square

Corolario 4.1.1. Si $\mathbb{L}(\mathbf{x}, \mathbf{L})$ es convexa con respecto a \mathbf{x} , entonces $\text{CVaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L}))$ también lo es. En este caso $F_\alpha(\mathbf{x}, \beta)$ es convexa.

Demostración. Si $\mathbb{L}(\mathbf{x}, \mathbf{L})$ es convexa con respecto a \mathbf{x} , entonces $[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+$ también lo es y por lo tanto $F_\alpha(\mathbf{x}, \beta) = \beta + \frac{1}{1-\alpha} \mathbb{E} [[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+]$ es convexa en (\mathbf{x}, β) .

Sea $\lambda \in (0, 1)$ y $\mathbf{z} = (\mathbf{x}, \beta)$. Se tiene que $F_\alpha(\mathbf{z})$ es convexa. Sean $\mathbf{z} = (\mathbf{x}, \beta)$ y $\mathbf{z}' = (\mathbf{x}', \beta')$ arbitrarios. Entonces

$$\begin{aligned} F_\alpha(\lambda \mathbf{z} + (1 - \lambda) \mathbf{z}') &\leq \lambda F_\alpha(\mathbf{z}) + (1 - \lambda) F_\alpha(\mathbf{z}') \\ \Rightarrow F_\alpha(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}', \lambda \beta + (1 - \lambda) \beta') &\leq \lambda F_\alpha(\mathbf{x}, \beta) + (1 - \lambda) F_\alpha(\mathbf{x}', \beta') \\ \Rightarrow \min_{\beta \in \mathbb{R}} F_\alpha(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}', \beta) &\leq \lambda F_\alpha(\mathbf{x}, \beta) + (1 - \lambda) F_\alpha(\mathbf{x}', \beta'). \end{aligned} \quad (32)$$

Como (32) se cumple para β que es arbitrario, se tiene que

$$\min_{\beta \in \mathbb{R}} F_\alpha(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}', \beta) \leq \lambda \min_{\beta \in \mathbb{R}} F_\alpha(\mathbf{x}, \beta) + (1 - \lambda) F_\alpha(\mathbf{x}', \beta')$$

Similarmente, como β' también es arbitrario,

$$\min_{\beta \in \mathbb{R}} F_\alpha(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}', \beta) \leq \lambda \min_{\beta \in \mathbb{R}} F_\alpha(\mathbf{x}, \beta) + (1 - \lambda) \min_{\beta \in \mathbb{R}} F_\alpha(\mathbf{x}', \beta),$$

esto es,

$$\text{CVaR}_\alpha(\mathbb{L}(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}', \mathbf{L})) \leq \lambda \text{CVaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L})) + (1 - \lambda) \text{CVaR}_\alpha(\mathbb{L}(\mathbf{x}', \mathbf{L})). \quad \square$$

Ahora, para calcular el mínimo en el problema $\min_{(\mathbf{x}, \beta) \in X \times \mathbb{R}} F_\alpha(\mathbf{x}, \beta)$, en Andersson et al. (2001) se expone cómo aproximar el valor esperado de la definición de $F_\alpha(\mathbf{x}, \beta)$ utilizando N escenarios, (a saber, $\mathbf{L}_1, \dots, \mathbf{L}_N$) que son muestreados de la distribución de probabilidad $f(\mathbf{L})$. Entonces

$$\begin{aligned} \mathbb{E} [[\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+] &= \int [\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+ d\mathbb{P} \\ &= \int [\mathbb{L}(\mathbf{x}, \mathbf{L}) - \beta]^+ f(\mathbf{L}) d\mathbf{L} \approx \frac{1}{N} \sum_{i=1}^N [\mathbb{L}(\mathbf{x}, \mathbf{L}_i) - \beta]^+. \end{aligned}$$

Utilizando esto se aproxima la solución de dicho problema resolviendo

$$\min_{(\mathbf{x}, \beta) \in X \times \mathbb{R}} \tilde{F}_\alpha(\mathbf{x}, \beta), \quad \text{donde} \quad \tilde{F}_\alpha(\mathbf{x}, \beta) \triangleq \beta + \frac{1}{(1 - \alpha) N} \sum_{i=1}^N [\mathbb{L}(\mathbf{x}, \mathbf{L}_i) - \beta]^+. \quad (33)$$

Un caso bastante común a muchas aplicaciones (Rockafellar y Uryasev, 2002) es el siguiente,

$$\mathbb{L}(\mathbf{x}, \mathbf{L}) = x_1 f_1(L_1) + x_2 f_2(L_2) + \dots + x_m f_m(L_m).$$

Entonces si $\mathbb{L}(\mathbf{x}, \mathbf{L})$ es lineal en \mathbf{x} y si X está dado por igualdades y desigualdades lineales se puede llevar (33) a un problema de optimización lineal, utilizando las variables dummy $\mathbf{z} = (z_1, \dots, z_N)$ de la siguiente manera,

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n, \mathbf{z} \in \mathbb{R}^N, \beta \in \mathbb{R}} & \left\{ \beta + \frac{1}{(1-\alpha)N} \sum_{i=1}^N z_i \right\} \\ \text{s.a. } & \mathbf{x} \in X, \\ & z_i \geq \mathbb{L}(\mathbf{x}, \mathbf{L}_i) - \beta, \\ & z_i \geq 0. \end{aligned} \tag{34}$$

4.1.2. Ejemplos

Rockafellar y Uryasev (2000) estudian el caso en que $\mathbf{x} = (x_1, \dots, x_n)$, donde x_i representa la posición tomada en el instrumento i ,

$$\forall i, x_i \geq 0 \quad \text{y} \quad \sum_{i=1}^n x_i = 1.$$

Sea $\mathbf{r} = (r_1, \dots, r_n)$, donde cada r_i son los retornos de L_i . Se miden las pérdidas de los instrumentos L_i basado en sus retornos, a saber, $L_i = -r_i$, para $i = 1, \dots, n$. Sea $\mathbf{L} = (L_1, \dots, L_n)$, entonces la función $\mathbb{L}(\mathbf{x}, \mathbf{L})$ está dada por

$$\mathbb{L}(\mathbf{x}, \mathbf{L}) \triangleq x_1 L_1 + \dots + x_n L_n = \mathbf{x}' \mathbf{L}.$$

Restringen el vector de decisión a portafolios con retorno esperado mayor a un retorno R dado. Por esto, si $\mathbf{m} = (m_1, \dots, m_n)$ es la media de las pérdidas \mathbf{L} , entonces la media de las pérdidas del portafolio para un \mathbf{x} dado es $\mu(\mathbf{x}) = \mathbf{x}' \mathbf{m}$ y definen

$$X \triangleq \left\{ \mathbf{x} \in \mathbb{R}^n \mid \forall i, x_i \geq 0, \quad \sum_{i=1}^n x_i = 1, \quad \mu(\mathbf{x}) \leq -R \right\}.$$

Los autores resuelven el problema de optimización del portafolio muestreando N valores de \mathbf{L} según la distribución de dicho vector y utilizan la aproximación lineal que se presentó anteriormente. Comparan los resultados de la aproximación con los obtenidos al resolver el problema de programación cuadrática propuesto por Markowitz (1952), en donde se minimiza la varianza del portafolio sujeto a $\mathbf{x} \in X$. Esto es, si V es la varianza de las pérdidas de \mathbf{L} , entonces la varianza del portafolio para un \mathbf{x} dado es $\sigma^2(\mathbf{x}) = \mathbf{x}'V\mathbf{x}$. Luego, el problema es

$$\min_{\mathbf{x} \in X} \sigma^2(\mathbf{x}).$$

Para poder comparar los resultados se asume que para todo \mathbf{x} , $\mathbb{L}(\mathbf{x}, \mathbf{L})$ se distribuye como una Normal (como es el caso si la distribución de \mathbf{L} es una normal multivariada). En este caso, si $\alpha \geq 0,5$ y la restricción $\mu(\mathbf{x}) \leq -R$ es activa, los dos problemas tienen la misma solución.

Para el caso particular en que se tienen 3 instrumentos (S&P 500, un portafolio de bonos de largo plazo del Gobierno de E.E.U.U y un portafolio de acciones) obtienen como resultado que las diferencias entre las soluciones de los dos enfoques es de unos pocos puntos porcentuales, aunque observan que la velocidad de convergencia del modelo aproximado propuesto por los autores depende de la metodología que se utilice para generar los datos aleatorios.

En este artículo, los autores también aplican esta metodología para cubrir¹ un portafolio cuyas pérdidas no son Normales y presentan algunos portafolios con VaR_α negativo (aceptable para esta medida), pero CVaR_α positivo (no aceptable para esta medida).

Similarmente, Andersson et al. (2001) utilizan la anterior metodología para optimizar el riesgo de crédito de un portafolio. Como lo definen los autores en este artículo, el riesgo de crédito es el riesgo que una contraparte no cumpla sus obligaciones completamente en la fecha establecida o una fecha posterior. En este caso utilizan Monte Carlo para generar la distribución del riesgo crediticio y esta metodología muestra ser muy eficiente, pudiendo ser resuelta con programación lineal, por lo que puede manejar cientos de instrumentos y miles de escenarios en tiempos

¹De la palabra en inglés “*hedge*”

de computación no muy altos.

4.2. Optimización dinámica de portafolios

Un aporte a la metodología expuesta en la sección anterior es propuesto por Zhu y Fukushima (2005), los cuales relajan la hipótesis que el inversor conoce la función de densidad de probabilidad f de \mathbf{L} en la madurez del contrato. Los autores asumen que se sabe que pertenece a un conjunto de distribuciones \mathcal{D} , es decir, se sabe que $f \in \mathcal{D}$. En este caso los autores definen para un nivel de confianza α dado, la medida estática coherente de riesgo “Worst Case CVaR” (WCVaR_α), con respecto a \mathcal{D} por

$$\text{WCVaR}_\alpha(L) = \sup_{f(\cdot) \in \mathcal{D}} \text{CVaR}_\alpha(L).$$

Entonces

$$\text{WCVaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L})) = \sup_{f(\cdot) \in \mathcal{D}} \min_{\beta \in \mathbb{R}} F_\alpha(\mathbf{x}, \beta).$$

Por lo tanto, minimizar el WCVaR_α sobre $\mathbf{x} \in X$ es equivalente a resolver el siguiente problema,

$$\min_{\mathbf{x} \in X} \sup_{f(\cdot) \in \mathcal{D}} \min_{\beta \in \mathbb{R}} F_\alpha(\mathbf{x}, \beta).$$

Basándose en esto, Huang et al. (2006) atacan el problema de posibles flujos de caja presentes en el portafolio antes de la fecha T modelando riesgo de precios en T y adicionalmente el riesgo de tiempo de salida de la posición.² Este riesgo está generado porque un inversor puede cerrar las posiciones y salir del mercado antes de la fecha de maduración de su horizonte de inversión. En este caso la distribución de las pérdidas del inversor cambia dependiendo del momento en que salga de las posiciones, ya que la distribución conjunta de los instrumentos financieros riesgosos es de naturaleza variable con respecto al tiempo.

A continuación se extiende de manera sencilla la metodología de la sección anterior para optimizar portafolios de manera similar a como se expuso en ésta, pero utilizando la medida dinámica de riesgo propuesta en la Sección 3.3.3, ecuación

² “Exit time risk”

(20), para el caso en que la medida estática coherente que la define es CVaR_α , para $\alpha \in (0, 1)$. De esta manera se presenta una metodología suficientemente amplia para modelar, entre otros, el riesgo de tiempo de salida de la posición financiera de manera diferente a Huang et al. (2006), ya que se consideran los cambios que puede tener ésta en el tiempo.

Sea $L \in \mathcal{M}_S^T$ y suponga que la medida de probabilidad $\mathbb{Q} \ll \mathbb{P}$ es tal que $\text{CVaR}_\alpha(\cdot) = \mathbb{E}_{\mathbb{Q}}[\cdot]$ es una medida (estática y coherente) del riesgo para el intervalo $[0, t]$, $t = 1, \dots, T$. Entonces, la medida dinámica de riesgo (denotada en este caso por $\text{MaxCVaR}_\alpha = (\text{MaxCVaR}_{\alpha,t})_{t=0, \dots, T}$ que se propone utilizar está dada por

$$\text{MaxCVaR}_{\alpha,t}(L) = \max_{i=t+1, \dots, T} \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t]. \quad (35)$$

Se ha visto anteriormente que esta medida dinámica es fuertemente coherente.

El enfoque que se presenta para optimizar portafolios busca minimizar inicialmente $\text{MaxCVaR}_{\alpha,0}(L)$, para $L \in \mathcal{M}_S^T$, utilizando lo expuesto en la anterior sección. Suponga que se utiliza el subconjunto $\{L^1, \dots, L^m\} \subseteq \mathcal{M}_S^T$ de riesgos financieros para construir lo que se llamará en general un portafolio dinámico para una inversión en el intervalo $[S, T]$ (ver Sección 4.3 para un ejemplo). Sea

$$\mathbf{L} = (L^1, \dots, L^m) = (\mathbf{L}_t)_{t=0,1, \dots} = (L_t^1, \dots, L_t^m)_{t=0,1, \dots}$$

y sea $\mathbf{x} \in X \subseteq \mathbb{R}^n$ un vector de decisión, para algún subconjunto X de \mathbb{R}^n que representa las restricciones al vector de decisión. Suponga que la medida de probabilidad \mathbb{P} y \mathbf{x} son independientes y que para todo t , la función $\mathbb{L}(\mathbf{x}, \mathbf{L}_t)$ representa las pérdidas del portafolio en la fecha t , es continua en \mathbf{x} y \mathcal{F}_t -medible en \mathbf{L}_t . Suponga que

$$\forall \mathbf{x} \in X, \sup_{t \in \mathbb{N}} \mathbb{E}[|\mathbb{L}(\mathbf{x}, \mathbf{L}_t)|] < \infty$$

y $\mathbb{L}(\mathbf{x}, 0) = 0$. Por esto, se cumple que el proceso $\mathbb{L}(\mathbf{x}, \mathbf{L}) \triangleq (\mathbb{L}(\mathbf{x}, \mathbf{L}_t))_{t=0,1, \dots} \in \mathcal{M}_S^T$, para todo $\mathbf{x} \in X$.

Si se conoce la distribución de probabilidad $f_t(\cdot)$ del vector aleatorio \mathbf{L}_t , entonces la variable $Z_{t,\mathbf{x}} = \mathbb{L}(\mathbf{x}, \mathbf{L}_t)$ es una variable aleatoria con distribución dependiente de \mathbf{x} . Para escoger un valor de \mathbf{x} se debe tener en cuenta la evolución de las variables aleatorias $Z_{0,\mathbf{x}}, Z_{1,\mathbf{x}}, Z_{2,\mathbf{x}}, \dots$

Similarmente a la definición hecha en la sección anterior, se define para cada $t = 0, 1, \dots$ la función $F_\alpha^t(\mathbf{x}, \beta) : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ por

$$F_\alpha^t(\mathbf{x}, \beta) \triangleq \beta + \frac{1}{1-\alpha} \mathbb{E} [[\mathbb{L}(\mathbf{x}, \mathbf{L}_t) - \beta]^+].$$

Entonces, por lo visto anteriormente

$$\text{CVaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L}_t)) = \min_{\beta \in \mathbb{R}} F_\alpha^t(\mathbf{x}, \beta).$$

Por esto

$$\begin{aligned} \text{MaxCVaR}_{\alpha,0}(\mathbb{L}(\mathbf{x}, \mathbf{L})) &= \max_{i=1,\dots,T} \{ \mathbb{E}_{\mathbb{Q}}[\mathbb{L}(\mathbf{x}, \mathbf{L}_i)] \} \\ &= \max_{i=1,\dots,T} \{ \text{CVaR}_\alpha(\mathbb{L}(\mathbf{x}, \mathbf{L}_i)) \} = \max_{i=1,\dots,T} \left\{ \min_{\beta \in \mathbb{R}} F_\alpha^i(\mathbf{x}, \beta) \right\}. \end{aligned} \quad (36)$$

Por último, se tiene que minimizar el ψ_0^{\max} para el portafolio, sobre $\mathbf{x} \in X$ es equivalente a resolver el problema mín-máx-mín,

$$\min_{\mathbf{x} \in X} \left\{ \max_{t=1,\dots,T} \left\{ \min_{\beta \in \mathbb{R}} F_\alpha^t(\mathbf{x}, \beta) \right\} \right\}. \quad (37)$$

Proposición 4.2.1. Si $\mathbb{L}(\mathbf{x}, \mathbf{L})$ es convexa con respecto a \mathbf{x} , entonces

$$\max_{t=1,\dots,T} \left\{ \min_{\beta \in \mathbb{R}} F_\alpha^t(\mathbf{x}, \beta) \right\}$$

también lo es.

Demostración. Como se demostró en el Corolario 4.1.1, se tiene que si $\mathbb{L}(\mathbf{x}, \mathbf{L})$ es convexa con respecto a \mathbf{x} , para cada t , $\min_{\beta \in \mathbb{R}} F_\alpha^t(\mathbf{x}, \beta) \triangleq g_t(\mathbf{x})$ es convexa. Entonces $\max_{t=1,\dots,T} g_t(\mathbf{x})$ es convexa.

En efecto, suponga por contradicción que existen $\mathbf{x}_1, \mathbf{x}_2 \in X$ y $\lambda \in (0, 1)$, tales que $\max_{t=1,\dots,T} g_t(\mathbf{x}_3) > \lambda \max_{t=1,\dots,T} g_t(\mathbf{x}_1) + (1-\lambda) \max_{t=1,\dots,T} g_t(\mathbf{x}_2)$, en donde $\mathbf{x}_3 \triangleq \lambda \mathbf{x}_1 + (1-\lambda) \mathbf{x}_2$. Suponga que j es tal que $g_j(\mathbf{x}_3) = \max_{t=1,\dots,T} g_t(\mathbf{x}_3)$. Entonces se tendría la contradicción,

$$\begin{aligned} \lambda \max_{t=1,\dots,T} g_t(\mathbf{x}_1) + (1-\lambda) \max_{t=1,\dots,T} g_t(\mathbf{x}_2) &< \max_{t=1,\dots,T} g_t(\mathbf{x}_3) = g_j(\mathbf{x}_3) \\ &\leq g_j(\mathbf{x}_1) + (1-\lambda)g_j(\mathbf{x}_2) \leq \lambda \max_{t=1,\dots,T} g_t(\mathbf{x}_1) + (1-\lambda) \max_{t=1,\dots,T} g_t(\mathbf{x}_2), \end{aligned}$$

por lo que tales $\mathbf{x}_1, \mathbf{x}_2$ y $\lambda \in (0, 1)$ no existen y $\max_{t=1,\dots,T} g_t(\mathbf{x})$ es convexa. \square

Por esta Proposición se observa que, de manera similar al enfoque de la Sección 4.1, si $\mathbb{L}(\mathbf{x}, \mathbf{L})$ es convexa con respecto a \mathbf{x} , la minimización en $\mathbf{x} \in X$ es de una función convexa.

De manera similar a como se expuso en la sección anterior, para aproximar el mínimo en el problema (37) se puede aproximar el valor esperado de la definición de $F_\alpha^t(\mathbf{x}, \beta)$, para cada t , utilizando N escenarios de L^1, \dots, L^m . Esto es, se generan datos aleatorios de cada \mathbf{L}_t , para $t = 1, 2, \dots, T$ (a saber, $\mathbf{L}_{t,1}, \dots, \mathbf{L}_{t,N}$), que son muestreados de la distribución de probabilidad $f_t(\mathbf{L}_t)$. Entonces

$$\begin{aligned} \mathbb{E} [\mathbb{L}(\mathbf{x}, \mathbf{L}_t) - \beta]^+ &= \int [\mathbb{L}(\mathbf{x}, \mathbf{L}_t) - \beta]^+ d\mathbb{P} \\ &= \int [\mathbb{L}(\mathbf{x}, \mathbf{L}_t) - \beta]^+ f_t(\mathbf{L}_t) d\mathbf{L}_t \approx \frac{1}{N} \sum_{i=1}^N [\mathbb{L}(\mathbf{x}, \mathbf{L}_{t,i}) - \beta]^+. \end{aligned}$$

Utilizando esto se aproxima la solución de dicho problema resolviendo

$$\begin{aligned} &\min_{\mathbf{x} \in X} \max_{t=1, \dots, T} \min_{\beta \in \mathbb{R}} \left\{ \tilde{F}_\alpha^t(\mathbf{x}, \beta) \right\}, \\ \text{donde } \tilde{F}_\alpha^t(\mathbf{x}, \beta) &\triangleq \beta + \frac{1}{(1-\alpha)N} \sum_{i=1}^N [\mathbb{L}(\mathbf{x}, \mathbf{L}_{t,i}) - \beta]^+. \end{aligned} \tag{38}$$

4.3. Aplicación

4.3.1. Definición del problema

Suponga que la fecha inicial es el 1 de diciembre de 2005. Al analizar los precios históricos del “*London Mercantile Exchange*” (LME) del Aluminio y del Cobre, un especulador cree que éstos van a bajar, dado que se encuentran por encima de su promedio de largo plazo (ver Figuras 4 y 5) y éste cree que los “*commodities*” presentan reversión a la media. El especulador posee USD 100.000 y desea comprar USD 1'000.000 en opciones Put.³ Las opciones las comprará sobre el promedio

³(Hull, 2003) Las Opciones son contratos financieros que se tranzan en las bolsas como en contratos privados entre dos partes. Una opción Put le da al tenedor de la misma el derecho de vender el activo subyacente (acción o “commodity” o etc.) del contrato, en una fecha dada, a un precio X dado. Una opción Call le da al tenedor de la misma el derecho de comprar el activo subyacente, en una fecha dada, a un precio X dado. El precio X en ambos casos es llamado precio de ejercicio, “*Strike Price*” o simplemente “*Strike*” y la fecha dada se llama fecha de expiración.

mensual del “cash official” (llamadas opciones TAPO Put) de estos dos “commodities”, ya que cree que los promedios tienen menos volatilidad que el precio “spot”. Las opciones serán sobre los promedios de octubre y/o noviembre. Para poder invertir

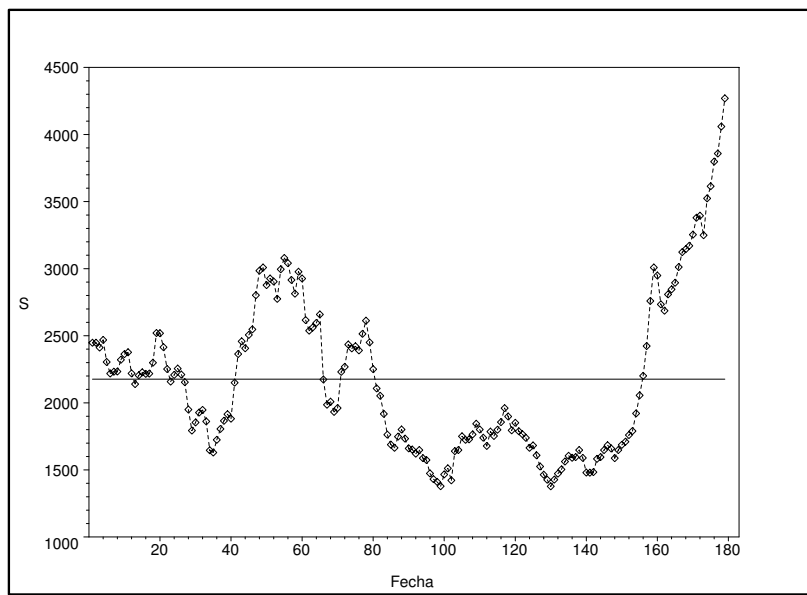


Figura 7: Promedios mensuales de cobre desde el 1 de enero de 91

USD 1'000.000 el especulador se financiará en los USD 900.000 que le hacen falta vendiendo opciones call sobre el promedio mensual del “cash official” (opciones TAPO call). Estas opciones las venderá sobre los promedios de marzo, mayo y/o julio, ya que siente que si la fecha de expiración es más cercana, la incertidumbre sobre los posibles pagos que deba realizar por la venta de estas opciones es menor.

Los precios al 1 de diciembre del 2005 son los presentados en las Tablas 1 y 2.⁴

El especulador desea entonces escoger cuántas toneladas debe comprar de cada una de las opciones de cobre y/o aluminio, sobre el promedio de octubre y/o noviembre comprar y cuántas toneladas de cobre y/o aluminio, sobre el promedio de

Estos contratos se pueden realizar sin entrega real del activo subyacente. En este caso, si el precio real del activo subyacente en la fecha de expiración es S , los ingresos para el comprador de la opción Put (resp. Call) son

$$\text{máx}\{X - S; 0\} \quad (\text{resp.} \quad \text{máx}\{S - X; 0\}).$$

⁴Fuente <http://www.lchclearnet.com/>

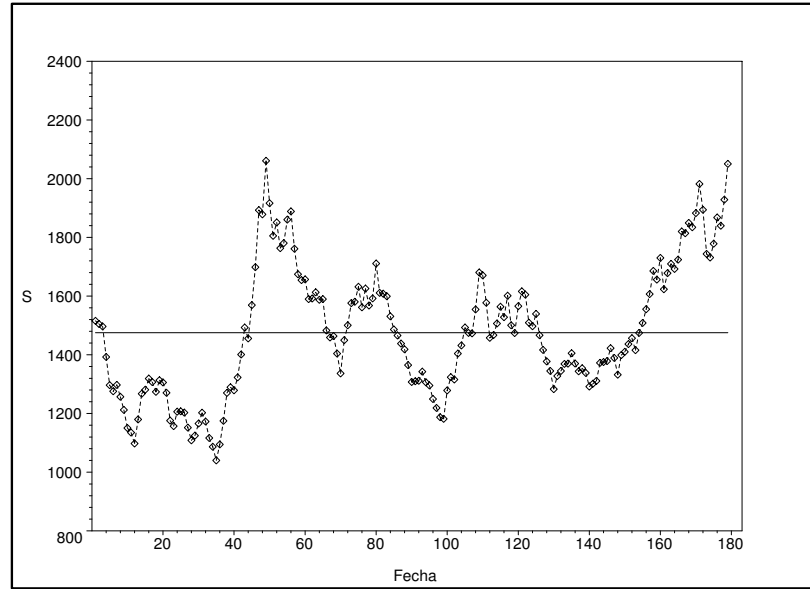


Figura 8: Promedios mensuales de aluminio desde el 1 de enero de 91

marzo, mayo y/o julio debe vender para financiarse, de tal manera que el riesgo de la inversión sea mínimo.

4.3.2. Primer enfoque

El primer enfoque es el presentado en la Sección 4.1.

Las fechas $t = 0, 1, \dots, 11$ serán los meses de la siguiente manera, $t = 0$ es noviembre, $t = 1$ es diciembre, ..., $t = 12$ es noviembre. Se utilizará la siguiente notación.

$S_{co,t}$ = precio promedio del cobre para el mes t ,

$X_{co,t}^{call}$ = precio de ejercicio de una opción call sobre cobre para el mes t ,

$p_{co,t}^{call}$ = prima de una opción call sobre cobre para el mes t ,

$X_{co,t}^{put}$ = precio de ejercicio de una opción put sobre cobre para el mes t ,

$p_{co,t}^{put}$ = prima de una opción put sobre cobre para el mes t .

Al reemplazar co por al se estará hablando del aluminio. Suponga que $r = 0$. Como los pagos de las primas de las opciones se hacen en $t = 0$, son flujos de caja iniciales y no se consideran en la función de pérdidas. Entonces se tiene que las

Cobre				
Fecha de Expiración	Contrato N°	Precio Futuro USD/ton	Strike USD/ton	Prima USD/ton
31-Mar-06	1	4.333,29	4.350,00	254,22
31-May-06	2	4.208,20	4.200,00	321,30
31-Jul-06	3	4.083,90	4.100,00	348,75
Aluminio				
Fecha de Expiración	Contrato N°	Precio Futuro USD/ton	Strike USD/ton	Prima USD/ton
31-Mar-06	4	2.193,43	2.200,00	103,68
31-May-06	5	2.165,14	2.175,00	125,76
31-Jul-06	6	2.125,00	2.125,00	149,14

Tabla 1: Precios Opciones TAPO call

Cobre				
Fecha de Expiración	Contrato N°	Precio Futuro USD/Ton	Strike USD/Ton	Prima USD/Ton
31-Oct-06	7	3.908,16	3.900,00	394,29
30-Nov-06	8	3.849,63	3.850,00	403,35
Aluminio				
Fecha de Expiración	Contrato N°	Precio Futuro USD/Ton	Strike USD/Ton	Prima USD/Ton
31-Oct-06	9	2.091,34	2.100,00	164,03
30-Nov-06	10	2.076,16	2.075,00	162,58

Tabla 2: Precios Opciones TAPO Put

pérdidas por la venta de una opción TAPO call con fecha de expiración t sobre cobre, acumuladas en la fecha T , son iguales a

$$\max\{S_{co,t} - X_{co,t}^{call}; 0\}$$

y se tiene una pérdida de $-p_{co,t}^{call}$ al inicio de la inversión (es decir, un ingreso por la venta de la opción).

Similarmente, las pérdidas por la compra de una opción TAPO put con fecha de expiración t sobre cobre, acumuladas en la fecha T , son iguales a

$$-\max\{X_{co,t}^{put} - S_{co,t}; 0\}$$

y se tiene una pérdida de $p_{co,t}^{put}$ al inicio de la inversión (es decir, un gasto por la compra de la opción).

El vector de decisión $\mathbf{x} = (x_1, \dots, x_{10}) \in \mathbb{R}^{10}$ se interpreta de la siguiente manera, x_i es la cantidad de toneladas compradas del contrato i (ver Tablas 1 y 2 para la definición de los contratos). Sea L_i la pérdida del contrato i , entonces la función de pérdidas acumuladas en la fecha T es igual a

$$\begin{aligned} \sum x_i L_i = & x_1 \max\{S_{co,4} - X_{co,4}^{call}; 0\} + x_2 \max\{S_{co,6} - X_{co,6}^{call}; 0\} \\ & + x_3 \max\{S_{co,8} - X_{co,8}^{call}; 0\} + x_4 \max\{S_{al,4} - X_{al,4}^{call}; 0\} \\ & + x_5 \max\{S_{al,6} - X_{al,6}^{call}; 0\} + x_6 \max\{S_{al,8} - X_{al,8}^{call}; 0\} \\ & - x_7 \max\{X_{co,11}^{put} - S_{co,11}; 0\} - x_8 \max\{X_{co,12}^{put} - S_{co,12}; 0\} \\ & - x_9 \max\{X_{al,11}^{put} - S_{al,11}; 0\} - x_{10} \max\{X_{al,12}^{put} - S_{al,12}; 0\} \end{aligned}$$

y se genera un flujo de caja inicial de

$$\begin{aligned} - x_1 p_{co,4}^{call} - x_2 p_{co,6}^{call} - x_3 p_{co,8}^{call} - x_4 p_{al,4}^{call} - x_5 p_{al,6}^{call} \\ - x_6 p_{al,8}^{call} + x_7 p_{co,11}^{put} + x_8 p_{co,12}^{put} + x_9 p_{al,11}^{put} + x_{10} p_{al,12}^{put}. \end{aligned} \quad (39)$$

Como la venta de las opciones TAPO call se hace para financiar la inversión, se pide que se cumpla que la cantidad de dinero que financiación esté entre los siguientes valores,

$$890000 \leq x_1 p_{co,3}^{call} + x_2 p_{co,5}^{call} + x_3 p_{co,7}^{call} + x_4 p_{al,3}^{call} + x_5 p_{al,5}^{call} + x_6 p_{al,7}^{call} \leq 910000. \quad (40)$$

Entonces, el valor de las opciones compradas no será menor que 990.000, pero no puede ser mayor que 100.000 el dinero que se obtenga de la financiación con la venta de las opciones, esto es

$$\begin{aligned} 990000 \leq x_7 p_{co,11}^{put} + x_8 p_{co,12}^{put} + x_9 p_{al,11}^{put} + x_{10} p_{al,12}^{put} \\ \leq 100000 + x_1 p_{co,4}^{call} + x_2 p_{co,6}^{call} + x_3 p_{co,8}^{call} + x_4 p_{al,4}^{call} + x_5 p_{al,6}^{call} + x_6 p_{al,8}^{call}. \end{aligned} \quad (41)$$

El especulador está comprando las opciones put y vendiendo opciones call, entonces también se debe cumplir que

$$x_i \geq 0, \quad \forall i = 1, \dots, 10. \quad (42)$$

Por lo tanto,

$$X \triangleq \{ \mathbf{x} = (x_1, \dots, x_{10}) \in \mathbb{R}^{10} \mid \text{se cumplen (40), (41) y (42)} \}.$$

Las pérdidas L_i dependen únicamente de $S_{co,t}$ y $S_{al,t}$, los promedios mensuales del cobre y aluminio, ya que las demás variables ya son conocidas y se encuentran en las Tablas 1 y 2. Se resolverá el problema utilizando la aproximación (33), para lo cual se generarán datos aleatorios para el promedio del cobre y del aluminio de la siguiente manera. Considere los retornos logarítmicos de las variables de interés,

$$r_{co,t} \triangleq \ln \left(\frac{S_{co,t}}{S_{co,t-1}} \right) \quad \text{y} \quad r_{al,t} \triangleq \ln \left(\frac{S_{al,t}}{S_{al,t-1}} \right)$$

y suponga que

$$\begin{bmatrix} r_{co,t} & r_{al,t} \end{bmatrix} \Big| \mathcal{F}_{t-1} \sim N(\boldsymbol{\mu}_t, \Sigma_t) = N \left(\begin{bmatrix} \mu_{co,t} & \mu_{al,t} \end{bmatrix}, \begin{bmatrix} \sigma_{co,t}^2 & \sigma_{coal,t} \\ \sigma_{coal,t} & \sigma_{al,t}^2 \end{bmatrix} \right),$$

donde se asume que la media $\boldsymbol{\mu}_t$ y la matriz de covarianzas Σ_t siguen un proceso basado en el GARCH multivariado (ver Jarrow ,1998),

$$\begin{aligned} \sigma_{co,t}^2 &= \gamma_{11} V_{co} + \beta_{11} \sigma_{co,t-1}^2 + \alpha_{11} (r_{co,t-1} - \mu_{co,t})^2, \\ \sigma_{coal,t} &= \omega_{12} + \beta_{12} \sigma_{coal,t-1} + \alpha_{12} (r_{co,t-1} - \mu_{co,t})(r_{al,t-1} - \mu_{al,t}), \\ \sigma_{al,t}^2 &= \gamma_{22} V_{al} + \beta_{22} \sigma_{al,t-1}^2 + \alpha_{22} (r_{al,t-1} - \mu_{al,t})^2, \end{aligned} \quad (43)$$

donde V_{co} y V_{al} son las varianzas de largo plazo de los retornos del cobre y del aluminio, γ_{ii} , α_{ij} y β_{ij} son parámetros a estimar que cumplen que $\gamma_{ii} \geq 0$, $\beta_{ii} \geq 0$, $\alpha_{ii} \geq 0$, $\gamma_{ii} + \beta_{ii} + \alpha_{ii} = 1$, para $i = 1, 2$ y $\omega_{12} \leq (\gamma_{11} V_{co} \gamma_{22} V_{al})^{1/2}$, $\beta_{12} \leq (\beta_{11} \beta_{22})^{1/2}$, $\alpha_{12} \leq (\alpha_{11} \alpha_{22})^{1/2}$.

Adicional al modelo se estiman las medias de la siguiente manera,

$$\begin{aligned} \mu_{co,t} &= \gamma_{11} M_{co} + \beta_{11} \mu_{co,t-1} + \alpha_{11} r_{co,t-1}, \\ \mu_{al,t} &= \gamma_{22} M_{al} + \beta_{22} \mu_{al,t-1} + \alpha_{22} r_{al,t-1}, \end{aligned} \quad (44)$$

donde M_{co} y M_{al} son las medias de largo plazo del cobre y del aluminio. No se asume que $\mu_{co,t}$ y $\mu_{al,t}$ son cero, ya que los retornos esperados son una variable importante para el inversor y también se asume que éstas son variables en el tiempo.

Entonces, el modelo se estima con la información histórica por Máxima Verosimilitud. Luego, como $r_{co,0}$, $r_{al,0}$, $S_{co,0}$ y $S_{al,0}$ son conocidos, se estiman $\tilde{\boldsymbol{\mu}}_1$ y $\tilde{\boldsymbol{\Sigma}}_1$ con el modelo y se generan $\tilde{r}_{co,1}$ y $\tilde{r}_{al,1}$ aleatoriamente, siguiendo la distribución $N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$. Con base en $\tilde{r}_{co,1}$ y $\tilde{r}_{al,1}$ se estiman $\tilde{\boldsymbol{\mu}}_2$ y $\tilde{\boldsymbol{\Sigma}}_2$ en el modelo (43) y (44) para generar $\tilde{r}_{co,2}$ y $\tilde{r}_{al,2}$ según la distribución $N(\tilde{\boldsymbol{\mu}}_2, \tilde{\boldsymbol{\Sigma}}_2)$. Así sucesivamente se genera un camino posible de los precios

$$\begin{aligned}\tilde{S}_{co,1} &= S_{co,0} e^{\tilde{r}_{co,1}}, \quad \tilde{S}_{co,2} = \tilde{S}_{co,1} e^{\tilde{r}_{co,2}}, \quad \dots, \quad \tilde{S}_{co,11} = \tilde{S}_{co,10} e^{\tilde{r}_{co,11}} \text{ y} \\ \tilde{S}_{al,1} &= S_{al,0} e^{\tilde{r}_{al,1}}, \quad \tilde{S}_{al,2} = \tilde{S}_{al,1} e^{\tilde{r}_{al,2}}, \quad \dots, \quad \tilde{S}_{al,11} = \tilde{S}_{al,10} e^{\tilde{r}_{al,11}}.\end{aligned}$$

Se generan entonces de esta manera N caminos posibles de estos precios, con esto se calculan los escenarios $\mathbf{L}_1, \dots, \mathbf{L}_N$ y se resuelve entonces el problema

$$\min_{(\mathbf{x}, \beta) \in X \times \mathbb{R}} \beta + \frac{1}{(1 - \alpha) N} \sum_{i=1}^N [\mathbb{L}(\mathbf{x}, \mathbf{L}_i) - \beta]^+. \quad (45)$$

4.3.3. Segundo enfoque

El segundo enfoque es el presentado en la Sección 4.2.

Se sigue con la misma notación y se toma $L^i = (L_t^i)$ como el proceso de las pérdidas del contrato i . Los flujos de caja iniciales son los mismos que en (39), entonces para la venta de una opción TAPO call con fecha de expiración 6 (mayo) sobre cobre (contrato 2), se tiene que

$$(L_0^1, \dots, L_5^1, L_6^1, \dots, L_{12}^1) = (0, \dots, 0, \max\{S_{co,6} - X_{co,6}^{call}, 0\}, \dots, \max\{S_{co,6} - X_{co,6}^{call}, 0\}).$$

Para los contratos 1, 3, 4, 5, 6 se tiene una expresión similar.

Las pérdidas acumuladas por la compra de una opción TAPO put con fecha de expiración 11 (octubre) sobre cobre (contrato 7), son

$$\begin{aligned}(L_0^7, \dots, L_{10}^7, L_{11}^7, L_{12}^7) \\ = (0, \dots, 0, -\max\{X_{co,11}^{put} - S_{co,11}, 0\}, -\max\{X_{co,11}^{put} - S_{co,11}, 0\}).\end{aligned}$$

Para los contratos 8, 9, 10 se tiene una expresión similar.

El vector de decisión $\mathbf{x} = (x_1, \dots, x_{10}) \in \mathbb{R}^{10}$ se interpreta de la misma manera como se explicó anteriormente y por esto X no cambia. Además,

$$\mathbb{L}(\mathbf{x}, \mathbf{L}) = (\mathbb{L}(\mathbf{x}, \mathbf{L}_t))_{t=0,1,\dots,10}, \quad \text{con } \mathbb{L}(\mathbf{x}, \mathbf{L}_t) = \sum_{i=1}^{10} x_i L_t^i.$$

Observe que las mismas simulaciones que se realizaron en el enfoque anterior se pueden utilizar para aproximar la solución de este problema, por lo que los datos simulados utilizados son los mismos. Esto además permite que se puedan comparar los resultados que se obtengan más fácilmente.

El problema que se resuelve es entonces,

$$\min_{\mathbf{x} \in X} \max_{t=1,\dots,12} \min_{\beta \in \mathbb{R}} \beta + \frac{1}{(1-\alpha)N} \sum_{i=1}^N \left[\sum_{j=1}^{10} x_i L_{t,i}^j - \beta \right]^+, \quad (46)$$

donde $L_{t,i}^j$ es la simulación i de la pérdida L^j (contrato j) para el periodo t .

Observe también que en esta notación, minimizar el CVaR_α del primer enfoque es equivalente a resolver

$$\min_{\mathbf{x} \in X} \min_{\beta \in \mathbb{R}} \beta + \frac{1}{(1-\alpha)N} \sum_{i=1}^N \left[\sum_{j=1}^{10} x_i L_{12,i}^j - \beta \right]^+, \quad (47)$$

así que la minimización en dicho enfoque sobre \mathbf{x} se realiza sobre una función no mayor que en la del segundo enfoque, por lo que en el óptimo se tiene que el CVaR_α del portafolio será menor o igual que el $\text{MaxCVaR}_{\alpha,0}^{\text{máx}}$ del mismo.

4.3.4. Primeros resultados

Para hacer las simulaciones de los precios se estimó el modelo (43) y (44) por máxima verosimilitud, utilizando los promedios históricos de los dos “*commodities*” desde el 1 de enero de 1991.⁵ Dado que es un modelo GARCH no típico, ya que se asume que el promedio de los retornos no es igual a cero, se hace esto en un programa específicamente diseñado en EXCEL.⁶ No obstante, se asume que los promedios de

⁵Fuente Bloomberg

⁶Ver los cálculos en el archivo GARCH aluminio+cobre+correl.xls

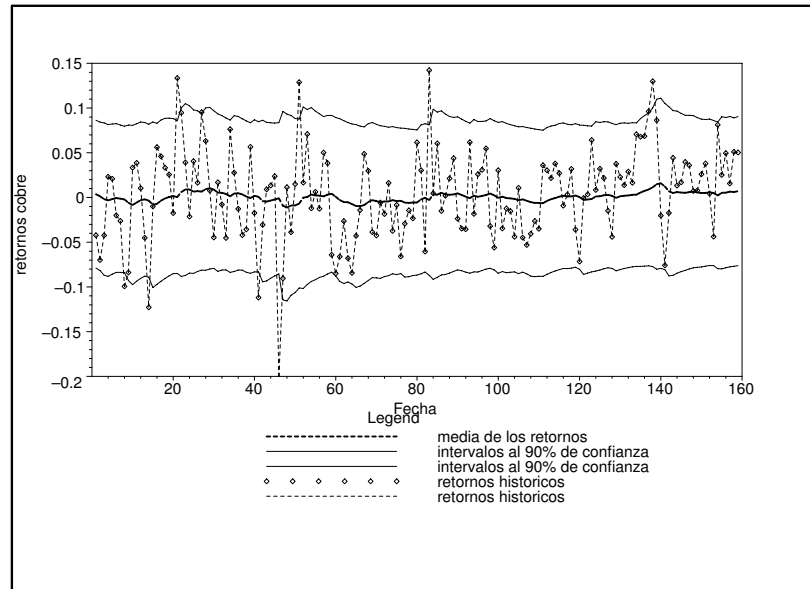


Figura 9: Modelo GARCH para el cobre

largo plazo de los retornos del cobre y del aluminio si son iguales a cero, es decir, $M_{co} = 0$ y $M_{al} = 0$. El modelo obtenido es (ver Figuras 6 y 7),

$$\begin{aligned}\mu_{co,t} &= 0,8261 \mu_{co,t-1} + 0,0417 r_{co,t-1}, \\ \sigma_{co,t}^2 &= 0,0004 + 0,8261 \sigma_{co,t-1}^2 + 0,0417 (r_{co,t-1} - \mu_{co,t})^2, \\ \mu_{al,t} &= 0,8398 \mu_{al,t-1} + 0,0452 r_{al,t-1}, \\ \sigma_{al,t}^2 &= 0,0002 + 0,8398 \sigma_{al,t-1}^2 + 0,0452 (r_{al,t-1} - \mu_{al,t})^2, \\ \sigma_{coal,t} &= 0,0002 + 0,8329 \sigma_{coal,t-1} + 0,0314 (r_{co,t-1} - \mu_{co,t})(r_{al,t-1} - \mu_{al,t}).\end{aligned}$$

Los datos al cierre del 31 de noviembre del 2005 son,

$$\begin{aligned}S_{co,0} &= \text{USD/ton } 4,269,34, & \mu_{co,0} &= 0,00800, & \sigma_{co,0}^2 &= 0,00258, & \sigma_{coal,0} &= 0,00136, \\ S_{al,0} &= \text{USD/ton } 2,050,59, & \mu_{al,0} &= 0,00531, & \sigma_{al,0}^2 &= 0,00174.\end{aligned}$$

Se hacen 10.000 simulaciones, es decir, $N = 10.000$.⁷

Utilizando las rutinas de optimización de MATLAB⁸ se resuelven los problemas (45) y (46).

⁷Ver archivo Simulaciones.xls

⁸Ver archivos adjuntos

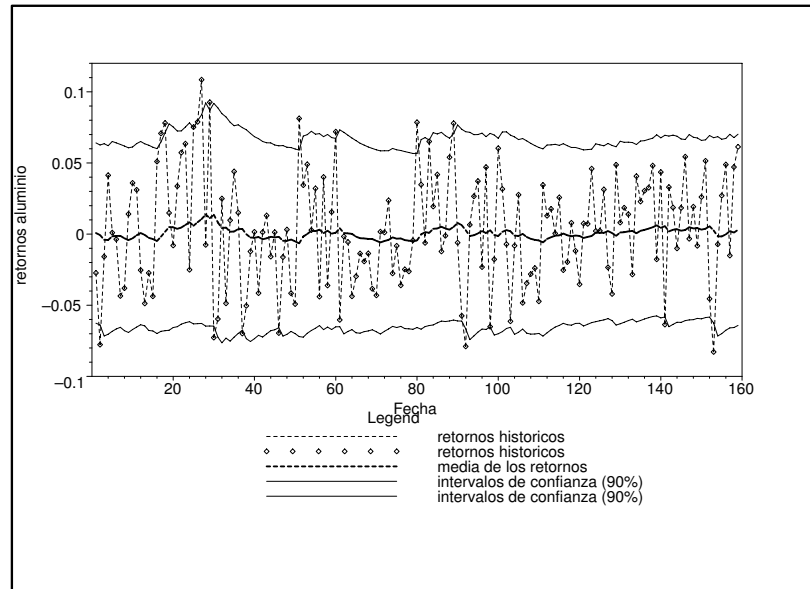


Figura 10: Modelo GARCH para el aluminio

4.3.4.1. *Resultados primer enfoque*

Para el Primer Enfoque se obtuvo que se deben vender opciones por 8584,10 toneladas de Aluminio con expiración en marzo únicamente (contrato 4) y comprar 245,94 toneladas de Cobre con expiración en noviembre (contrato 8), 3181,01 toneladas de Aluminio con expiración en octubre (contrato 9), 2269,76 toneladas de Aluminio con expiración en noviembre (contrato 10) y el β es igual a USD 823.761,02. Si se disminuye el β óptimo en 1 tonelada se aumenta la función objetivo, cuyo valor es USD 1'705.338,77, por lo tanto, el CVaR_{90} y VaR_{90} óptimos son aproximadamente

$$\text{CVaR}_{90}(\mathbb{L}(\mathbf{x}_1^*, \mathbf{L})) \approx \text{USD } 1'705,339 \text{ y } \text{VaR}_{90}(\mathbb{L}(\mathbf{x}_1^*, \mathbf{L})) \approx \text{USD } 823,761.$$

4.3.4.2. *Resultados parciales segundo enfoque*

Para el Primer Enfoque se obtuvo que se deben vender opciones por 8584,10 toneladas de Aluminio con expiración en marzo únicamente (contrato 4) y comprar 879,73 toneladas de Cobre con expiración en noviembre (contrato 7), 879,62 toneladas de Cobre con expiración en noviembre (contrato 8), 882,81 toneladas de Aluminio

con expiración en octubre (contrato 9), 882,83 toneladas de Aluminio con expiración en noviembre (contrato 10). Esta solución se obtuvo ingresando al programa de optimización el escenario inicial $\mathbf{x}_0 = (695, 695, 695, 695, 695, 695, 890, 890, 890, 890)$.

Observe que del mes 1 (diciembre) al 3 (febrero) no hay flujos de caja. Hay posibles flujos en 4 (marzo) pero en 5 (abril) los flujos de caja acumulados son los mismos que en 4. Siguiendo con esta lógica solo se estudian los flujos de caja acumulados en los meses donde hay cambios, es decir, los meses 4, 6, 8, 11 y 12. En el \mathbf{x} óptimo (\mathbf{x}_2^*) se tiene que

$$\begin{aligned}\tilde{F}_{90}^4(\mathbf{x}_2^*, \beta_1^*) &= 1'735.758,63, & \text{con } \beta_1^* &= 881.614,27, \\ \tilde{F}_{90}^6(\mathbf{x}_2^*, \beta_2^*) &= 1'735.758,63, & \text{con } \beta_2^* &= 881.614,27, \\ \tilde{F}_{90}^8(\mathbf{x}_2^*, \beta_3^*) &= 1'735.758,63, & \text{con } \beta_3^* &= 881.614,27, \\ \tilde{F}_{90}^{11}(\mathbf{x}_2^*, \beta_4^*) &= 1'720.914,98, & \text{con } \beta_4^* &= 856.826,22, \\ \tilde{F}_{90}^{12}(\mathbf{x}_2^*, \beta_5^*) &= 1'709.844,02, & \text{con } \beta_5^* &= 840.719,81.\end{aligned}$$

El máximo de éstos es para $t = 4$, por lo tanto

$$\text{MaxCVaR}_{\alpha,0}(\mathbb{L}(\mathbf{x}_2^*, \mathbf{L})) = \max_{i=1,\dots,T} \{\mathbb{E}_{\mathbb{Q}_i}[\mathbb{L}(\mathbf{x}_2^*, \mathbf{L}_t) \mid \mathcal{F}_0]\} \approx \text{USD } 1'735,758.$$

Si cualquier β_i^* es disminuido en 1 tonelada, el valor de la función objetivo respectiva aumenta. Entonces, observe que β_i^* es el VaR_{90} para las pérdidas acumuladas del portafolio en el mes i , entonces el VaR_{90} más alto en el proceso de la inversión, denotado $\text{VaR}_{90}^{\text{máx}}$, es aproximadamente USD 881.614.

En este enfoque puede haber más de una solución \mathbf{x}^* óptima (ver siguiente Sección) y no es adecuado desde el punto de vista de selección de portafolios tomar una solución óptima cualquiera, por razones que serán claras a continuación.

4.3.4.3. *Comparación de los resultados*

Se analiza ahora el resultado del Primer Enfoque desde el punto de vista del Segundo Enfoque. Reemplazando la solución \mathbf{x}_1^* se obtiene lo siguiente,

$$\begin{aligned}\tilde{F}_{90}^4(\mathbf{x}_1^*, \beta_1^*) &= 1'735.758,63, & \text{con } \beta_1^* &= 881.614,27, \\ \tilde{F}_{90}^6(\mathbf{x}_1^*, \beta_2^*) &= 1'735.758,63, & \text{con } \beta_2^* &= 881.614,27, \\ \tilde{F}_{90}^8(\mathbf{x}_1^*, \beta_3^*) &= 1'735.758,63, & \text{con } \beta_3^* &= 881.614,27, \\ \tilde{F}_{90}^{11}(\mathbf{x}_1^*, \beta_4^*) &= 1'715.857,22, & \text{con } \beta_4^* &= 840.625,61, \\ \tilde{F}_{90}^{12}(\mathbf{x}_1^*, \beta_5^*) &= 1'705.339,77, & \text{con } \beta_5^* &= 823.761,02.\end{aligned}$$

Por lo tanto,

$$\text{MaxCVaR}_{\alpha,0}(\mathbb{L}(\mathbf{x}_1^*, \mathbf{L})) = \text{MaxCVaR}_{\alpha,0}(\mathbb{L}(\mathbf{x}_2^*, \mathbf{L})) \approx \text{USD } 1'735,758.$$

Entonces, el resultado obtenido para el Primer Enfoque es óptimo también para el Segundo Enfoque, pero además se tiene que logra disminuir el riesgo para los meses 11 y 12 (octubre y noviembre), ya que el CVaR_{90} para estos meses es menor que el obtenido en el óptimo \mathbf{x}_2^* . Lo mismo sucede con el VaR_{90} asociado a estos meses. Esto ocurre porque al minimizar $\text{MaxCVaR}_{\alpha,0}$ solo se minimiza el peor caso entre las diferentes fechas, pero no se busca minimizar el riesgo en las demás fechas. Por lo tanto, es necesario en el Segundo Enfoque, encontrar una manera de optimizar el resultado entre los \mathbf{x}^* óptimos del problema

$$\min_{\mathbf{x} \in X} \left\{ \max_{t=1, \dots, T} \left\{ \min_{\beta \in \mathbb{R}} F_{\alpha}^t(\mathbf{x}, \beta) \right\} \right\}. \quad (48)$$

A continuación se propone una metodología para hacer esto.

4.3.5. Segundo enfoque completado

4.3.5.1. *Descripción*

Se propone utilizar un caso particular de la medida (21) en este contexto para optimizar la solución \mathbf{x}_2^* , presentada. Su definición es

$$\overline{\text{CVaR}}_{\alpha,t}(L) = \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}}[L_i | \mathcal{F}_t]. \quad (49)$$

Se busca entonces en este caso resolver el problema

$$\begin{aligned} & \min_{\mathbf{x}} \overline{\text{CVaR}}_{\alpha,0}(\mathbb{L}(\mathbf{x}, \mathbf{L})), \\ \text{s.a. } & \text{MaxCVaR}_{\alpha,0}(\mathbb{L}(\mathbf{x}, \mathbf{L})) \leq \min_{\mathbf{x} \in X} \text{MaxCVaR}_{\alpha,0}(\mathbb{L}(\mathbf{x}, \mathbf{L})) + \epsilon, \end{aligned}$$

para un $\epsilon > 0$ dado. Este ϵ es la cantidad en la que el inversor está dispuesto a aumentar el valor del riesgo máximo de la inversión, para buscar que el riesgo ponderado disminuya en las demás fechas.

Utilizando de nuevo lo expuesto en la Sección 4.1, se obtiene que el problema a resolver es

$$\begin{aligned} & \min_{\mathbf{x} \in X} \left\{ \sum_{i=t+1}^T \lambda_i^t \left(\min_{\beta \in \mathbb{R}} F_{\alpha}^t(\mathbf{x}, \beta) \right) \right\}, \\ \text{s.a. } & \max_{t=1, \dots, T} \left\{ \min_{\beta \in \mathbb{R}} F_{\alpha}^t(\mathbf{x}, \beta) \right\} \leq \min_{\beta \in \mathbb{R}} F_{\alpha}^t(\mathbf{x}_2^*, \beta) + \epsilon \end{aligned}$$

y se aproxima nuevamente con

$$\begin{aligned} & \min_{\mathbf{x} \in X} \left\{ \sum_{i=t+1}^T \lambda_i^t \left(\min_{\beta \in \mathbb{R}} \tilde{F}_{\alpha}^t(\mathbf{x}, \beta) \right) \right\}, \\ \text{s.a. } & \max_{t=1, \dots, T} \left\{ \min_{\beta \in \mathbb{R}} \tilde{F}_{\alpha}^t(\mathbf{x}, \beta) \right\} \leq \min_{\beta \in \mathbb{R}} \tilde{F}_{\alpha}^t(\mathbf{x}_2^*, \beta) + \epsilon \end{aligned}$$

4.3.5.2. Aplicación y Resultados

Dado que los únicos meses en los que puede haber cambios en los flujos aleatorios acumulados son el 4, 6, 8, 11 y 12, se toma $\lambda_1^0 = \lambda_2^0 = \lambda_3^0 = \lambda_5^0 = \lambda_7^0 = \lambda_9^0 = \lambda_{10}^0 = 0$ y $\lambda_4^0 = \lambda_6^0 = \lambda_8^0 = \lambda_{11}^0 = \lambda_{12}^0 = \frac{1}{5}$. Entonces

$$\begin{aligned} & \min_{\mathbf{x} \in X} \left\{ \sum_{t=1}^T \lambda_i^t \left(\min_{\beta \in \mathbb{R}} \beta + \frac{1}{(1-\alpha)N} \sum_{i=1}^N \left[\sum_{j=1}^{10} x_i L_{t,i}^j - \beta \right]^+ \right) \right\}, \\ \text{s.a. } & \max_{t=1, \dots, T} \left\{ \min_{\beta \in \mathbb{R}} \beta + \frac{1}{(1-\alpha)N} \sum_{i=1}^N \left[\sum_{j=1}^{10} x_i L_{t,i}^j - \beta \right]^+ \right\} \leq \min_{\beta \in \mathbb{R}} \tilde{F}_{\alpha}^t(\mathbf{x}_2^*, \beta) + \epsilon \end{aligned}$$

Se resuelve el problema para $\epsilon = 0, 1, 1000, 10000, 100000, 1000000$ y los resultados se presentan en la Tabla 3. Si se disminuye en 1 los valores de los β_i^* , las funciones objetivo aumentan, por lo que éstos son los VaR_{90} para el respectivo periodo, i .

ϵ	0,1	1.000	10.000	100.000	1'000.000
$\tilde{F}_{90}^4(\mathbf{x}^*, \beta_1^*)$	1'735.754	1'701.050	1'599.741	1'195.752	600.739
$\tilde{F}_{90}^6(\mathbf{x}^*, \beta_2^*)$	1'735.758	1'736.759	1'652.146	1'195.752	600.750
$\tilde{F}_{90}^8(\mathbf{x}^*, \beta_3^*)$	1'735.758	1'736.759	1'745.759	1'835.759	2'137.538
$\tilde{F}_{90}^{11}(\mathbf{x}^*, \beta_4^*)$	1'706.673	1'708.767	1'721.642	1'824.501	2'134.110
$\tilde{F}_{90}^{12}(\mathbf{x}^*, \beta_5^*)$	1'706.673	1'708.765	1'721.642	1'824.501	2'134.110
β_1^*	881.612	863.985	812.528	607.338	305.124
β_2^*	881.612	892.179	854.677	607.338	305.123
β_3^*	881.612	892.179	909.299	1'059.607	1'383.090
β_4^*	826.952	834.272	858.376	1'032.949	1'374.225
β_5^*	826.952	834.272	858.376	1'032.949	1'374.225
$\text{CVaR}_{90}(\mathbb{L}(\mathbf{x}^*, \mathbf{L}))$	1'706.673	1'708.765	1'721.642	1'824.501	2'134.110
$\text{VaR}_{90}(\mathbb{L}(\mathbf{x}^*, \mathbf{L}))$	826.952	834.272	858.376	1'032.949	1'374.225
$\text{MaxCVaR}_{90,0}(\mathbb{L}(\mathbf{x}^*, \mathbf{L}))$	1'735.758	1'736.759	1'745.759	1'835.759	2'137.538
$\overline{\text{CVaR}}_{90,0}(\mathbb{L}(\mathbf{x}^*, \mathbf{L}))$	1'724.124	1'718.420	1'688.186	1'575.253	1'521.450
$\text{VaR}_{90}^{\text{máx}}(\mathbb{L}(\mathbf{x}^*, \mathbf{L}))$	881.612	892.179	909.299	1'059.625	1'383.090

Tabla 3: Resultados Segundo Enfoque Completado

El resultado para $\epsilon = 0,1$ mejora el primer resultado mostrado utilizando el Segundo Enfoque, ya que mantiene el valor de $\text{MaxCVaR}_{90,0}$, pero disminuye el valor de todas las demás medidas de riesgo presentadas y las medidas de riesgo estáticas en los dos últimos periodos. El CVaR_{90} es naturalmente mayor que el obtenido en el primer enfoque, al igual que el VaR_{90} , pero el aumento es pequeño comparado con la disminución en el CVaR_{90} y VaR_{90} para el penúltimo periodo.

Para los diferentes valores de ϵ se observa que si el valor de ϵ aumenta de 0,1 a 1.000 los cambios no son tan grandes en las diferentes medidas, pero al aumentar a 100.000 y 1'000.000 se tiene una mejoría en el valor de la medida $\overline{\text{CVaR}}_{90,0}$, pero a expensas de un aumento significativo en el valor de las demás medidas.

Observe que las medidas CVaR_{90} y VaR_{90} no necesariamente aumentarían al mejorar el valor de $\overline{\text{CVaR}}_{90,0}$, ya que ésta es una medida del valor promedio del riesgo en los diferentes meses y podría haber disminuido al tener disminuciones en el riesgo del último periodo. Esto no sucede para la medida $\text{MaxCVaR}_{90,0}$, ya que si disminuye $\overline{\text{CVaR}}_{90,0}$, es decir, el promedio de los riesgos, debe aumentar el riesgo máximo en las diferentes fechas (porque de lo contrario, al mejorar $\overline{\text{CVaR}}_{90,0}$ se estaría mejorando

MaxCVaR_{90,0}, que ya ha sido minimizada) y por lo tanto MaxCVaR_{90,0} es la única medida entre las medidas propuestas capaz de detectar posibles aumentos en los flujos de caja en una fecha dada, que puedan ser demasiado altos para lograr sostener la posición abierta durante toda la vida de la misma.

Por otro lado, observe que solucionar el problema con $\epsilon = 1'000.000$ es igual que resolver el problema sin restricciones (o con $\epsilon = \infty$), ya que el aumento en MaxCVaR_{90,0} es apenas de 401.780. Se observa entonces que los resultados de la optimización del portafolio minimizando $\overline{\text{CVaR}}_{90,0}$ sin restricciones no es adecuada en general, porque se está aumentando significativamente el riesgo que se está corriendo en los últimos tres periodos de análisis (fechas 8, 11 y 12) y de no observar las demás medidas de riesgo se estaría ignorando dicho riesgo.

4.3.6. Pérdidas esperadas, realizadas y comparación de las medidas

Los diferentes valores de ϵ y el resultado del primer enfoque tienen como resultado también diferentes portafolios en los cuales se puede invertir. La decisión de en cuál de estos portafolios es mejor invertir es decisión del inversor según sus preferencias y aversión al riesgo, para lo cual se puede también analizar los retornos esperados. En la Tabla 4 se muestran los valores esperados de las pérdidas simuladas (Esp) y los valores realizados (Real) de las pérdidas, calculados de los datos que en realidad se dieron. En las Figuras 8 y 9 se presentan los promedios del cobre y aluminio realizados.⁹

Las pérdidas esperadas muestran que el Segundo Enfoque con $\epsilon = 0,1$ logra más rápidamente tener ganancias esperadas altas que los demás enfoques. Además tiene las ganancias esperadas más altas al final de toda la inversión. Dado que, como se observó anteriormente, además logra esto sin ser mucho más riesgoso que el resultado del Primer Enfoque y teniendo un riesgo más balanceado en el sentido en que no tiene meses en los que el riesgo sea demasiado alto comparado con los demás meses, este enfoque es bastante atractivo para cualquier inversor desde el punto de vista de este análisis.

Los resultados obtenidos con el Primer Enfoque también son bastante buenos,

⁹Fuente: Bloomberg

		Mes 4	Mes 6	Mes 8	Mes 11	Mes 12
Enfoque 1	Esp	223.901	223.901	223.901	-155.437	-422.886
	Real	1.966.880	1.966.880	1.966.880	1.966.880	1.966.880
Enfoque 2 $\epsilon = 0,1$	Esp	223.900	223.901	223.901	-477.079	-477.079
	Real	1.966.875	1.966.887	1.966.887	1.966.887	1.966.887
Enfoque 2 $\epsilon = 1.000$	Esp	219.424	226.324	226.324	-472.186	-472.254
	Real	1.927.550	2.025.473	2.025.473	2.025.473	2.025.473
Enfoque 2 $\epsilon = 10.000$	Esp	206.355	216.457	240.214	-387.743	-387.743
	Real	1.812.751	1.956.099	2.069.679	2.069.679	2.069.679
Enfoque 2 $\epsilon = 100.000$	Esp	154.244	154.244	304.801	-159.394	-159.394
	Real	1.354.970	1.354.970	2.074.779	2.074.779	2.074.779
Enfoque 2 $\epsilon = \infty$	Esp	77.491	77.493	393.941	206.439	206.439
	Real	680.729	680.760	2.193.686	2.193.686	2.193.686

Tabla 4: Pérdidas esperadas (Esp) y realizadas (Real)

ya que tiene ganancias esperadas altas para el último mes y un riesgo bajo a lo largo de la inversión. No obstante, que la solución óptima no tenga meses en los que el riesgo es muy alto depende de esta aplicación en particular y no de la medida, ya que CVaR_α en general solamente observa los resultados del último periodo.

Por otro lado, cuando ϵ aumenta, las ganancias esperadas decrecen y como existen los problemas de riesgo ya nombrados, los resultados obtenidos con estos enfoques no serán atractivos en general para ningún inversor desde el punto de vista de este análisis.

Observando las pérdidas realizadas, la aplicación muestra cómo el VaR_α , tal como se da en la discusión teórica, puede subestimar el riesgo, ya que solamente es el valor para el cual las pérdidas no serán mayores, pero no mide la severidad de dichas pérdidas, a diferencia de CVaR_α y $\psi_0^{\text{máx}}$. Esto es bastante dramático en este caso, ya que CVaR_α y $\text{MaxCVaR}_{90,0}$ son en casi todos los casos mayores que el doble del VaR_α y las pérdidas realizadas también. Entonces este caso muestra como CVaR_α y $\text{MaxCVaR}_{90,0}$ pueden en realidad ser unas medidas más cercanas del valor en riesgo real, comparadas con VaR_α .

Se observa también que en esta aplicación en particular resulta que el Primer Enfoque y el Segundo Enfoque con $\epsilon = 0, 1$, que son los que menor riesgo presentaron

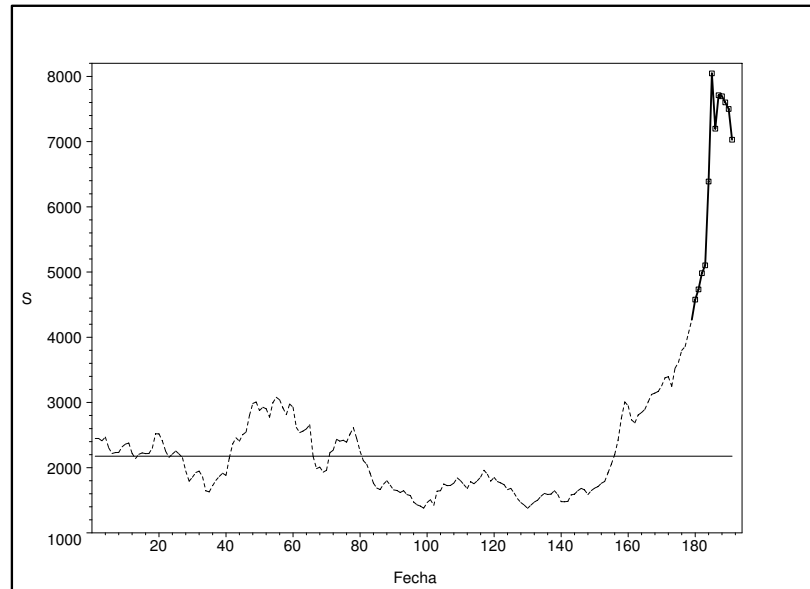


Figura 11: Promedios mensuales de cobre realizados

en el análisis previo, también fueron los que menores pérdidas realizadas tuvieron.

Por otro lado, se observa de los resultados que CVaR_α no mide el hecho que el riesgo en las fechas intermedias, sobre todo para la fecha 8 (julio), el riesgo es mayor que en las fechas finales. Esto demuestra de nuevo que esta medida puede no asegurar la factibilidad de la inversión durante toda la vida de ésta y entonces puede llevar a resultados en los que puede ser necesario cerrar la posición antes de tiempo, a diferencia de $\text{MaxCVaR}_{90,0}$.

Similarmente, se ha observado que la medida $\overline{\text{CVaR}}_{90,0}$ (que es coherente) también puede ignorar riesgos intermedios que puedan hacer necesario cerrar la posición antes de lo planeado y no es una medida que refleje el riesgo real de la inversión. Para $\epsilon = 1'000.000$ se tiene que en el óptimo $\overline{\text{CVaR}}_{90,0}$ es pequeño porque el CVaR_α es pequeño en las fechas iniciales. Por esto, $\overline{\text{CVaR}}_{90,0} = 1'521.450$, pero $\text{MaxCVaR}_{90,0} = 2'137.538$ y las pérdidas realizadas son $2'193.686$, lo que sucede porque $\overline{\text{CVaR}}_{90,0}$ ignora los riesgos muy altos que se tienen en las últimas fechas. Entonces se observa que ser *coherente* puede no ser suficiente para que la medida sea una medida adecuada y por lo tanto es necesario utilizar medidas *fuertemente coherentes*.

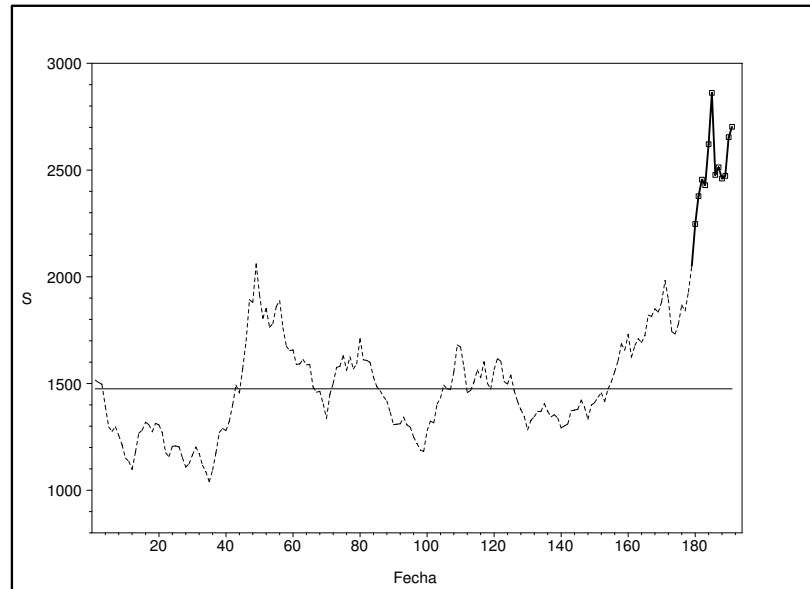


Figura 12: Promedios mensuales de aluminio realizados

De este análisis también se obtiene que es importante tener en cuenta cómo cambia el riesgo en las diferentes fechas al observar los valores de $\tilde{F}_\alpha^i(\mathbf{x}^*, \beta_i^*)$, $i = 1, \dots, T$. Éste es un mapa de los riesgos en que se incurre en las diferentes fechas y entonces un análisis completo de riesgo se ha visto que puede hacerse observando tanto el promedio de estos riesgos ($\overline{\text{CVaR}}_{90,0}$), como el valor máximo de ellos ($\text{MaxCVaR}_{90,0}$). Esto es porque al utilizar como herramienta de análisis y de optimización $\text{MaxCVaR}_{90,0}$ se puede lograr asegurar la viabilidad de la inversión hasta la fecha final y al utilizar además $\overline{\text{CVaR}}_{90,0}$ se puede minimizar el riesgo en las fechas donde el riesgo no sea máximo.

Por último se tiene también que el observar la evolución de las pérdidas (o ganancias) esperadas (como se ve en la Tabla 4) brinda al inversor más herramientas para saber cuál de los posibles portafolios escoger, comparado con observar solamente las pérdidas (o ganancias) esperadas acumuladas del último periodo. En la aplicación en particular se observa que el utilizar $\overline{\text{CVaR}}_{90,0}$ y $\text{MaxCVaR}_{90,0}$ en el análisis puede tener resultados muy favorables desde el punto de vista de pérdidas (o ganancias) esperadas, tanto como riesgo. No obstante, utilizar alguna de las medidas de riesgo como herramienta sin la otra puede no ser adecuado.

Capítulo V

CONCLUSIONES

En el presente trabajo se expone de manera general la teoría de medidas coherentes estáticas de riesgo y luego ésta es utilizada para extender la teoría de medidas de riesgo y optimización de portafolios a un contexto dinámico. Se recopilan las ideas de diferentes autores sobre temas generales de medidas dinámicas de riesgo, se definen las medidas dinámicas *coherentes* y se propone una extensión del concepto de medida coherente del contexto estático al dinámico que es más fuerte que las propuestas por dichos autores.

Esta extensión fuerte, que es argumentada desde el punto de vista teórico, se ve justificada con la aplicación presentada en el Capítulo 4, ya que se obtiene que una medida que sea simplemente coherente (Cheridito, et al. (2004)), o débilmente coherente (propuesto por la mayoría de los autores estudiados), puede ignorar el riesgo en el que incurre un inversionista de un portafolio dado, en fechas intermedias. Estos riesgos generan en el ejemplo flujos de caja grandes que pueden hacer que el inversor tenga que cerrar las posiciones y salir del mercado antes de la fecha final o en la fecha final lo pueden llevar a la quiebra.

La medida MaxCVaR_α , que es la única medida dinámica fuertemente coherente, muestra ser en realidad una mejor medida que VaR_α , CVaR_α y $\overline{\text{CVaR}}_\alpha$, ya que es la única que no ignora riesgos potencialmente muy altos de las fechas intermedias.

La aplicación muestra que el enfoque para optimización de portafolio en el contexto estático, expuesto en la Sección 4.1 puede lograr buenos resultados, ya que se obtiene un portafolio con retornos esperados altos. Además, el valor de la medida de riesgo CVaR_α en este caso es más cercano al riesgo real que el riesgo que mide

VaR_α , ya que ésta última no mide la severidad de las pérdidas en casos extremos. No obstante, CVaR_α no es una medida que en general logre asegurar que no se tengan que cerrar las posiciones antes de la fecha final porque ignora riesgos intermedios.

Se observa que el enfoque para optimizar portafolios analizando el CVaR_α de las pérdidas aleatorias acumuladas, tanto como los retornos esperados en las diferentes fechas brinda al inversor muchas herramientas para poder hacer una selección adecuada de portafolio. El enfoque, basado en dos etapas de optimización de las medidas MaxCVaR_α y $\overline{\text{CVaR}}_\alpha$ muestra ser una alternativa fácil de entender e implementar y con buenos resultados para obtener, desde fechas intermedias, retornos esperados altos, con un riesgo adecuadamente medido por MaxCVaR_α durante toda la tenencia del portafolio. No obstante, no se pretende hacer en este trabajo un análisis de los tiempos necesarios para la optimización y los recursos computacionales requeridos.

Por otro lado, una interpretación que se hace del valor de las medidas estáticas de riesgo no es adecuado en el contexto dinámico. Esto es, para una medida estática coherente ρ , si $\rho(L) > 0$, este valor es interpretado como el capital necesario para que la posición sea aceptable, como se vio en la Sección 2.2. Si $\psi_0(L) > 0$, $\psi_0(L)$ no debe ser interpretado como la cantidad que se debe añadir a la posición en la fecha inicial para que ésta sea aceptable. Esto es porque, aunque el valor $\psi(L)$ en realidad es un valor que al ser añadido a la posición en la fecha inicial logra $\psi_0(L - \psi_0(L)) = \psi_0(L) - \psi_0(L) = 0$, esto ignora que el riesgo puede ir cambiando en el tiempo, como se observa en la aplicación. Por ejemplo, para $\epsilon = 1'000.000$ en el Segundo Enfoque se tiene que el $\text{MaxCVaR}_{90,0}$ es 2'137.538, pero el riesgo hasta la fecha 8 es mucho menor, como lo muestran los CVaR_{90} de los primeros 7 periodos. Por lo tanto, pedir que se añada a la posición el valor $\text{MaxCVaR}_{90,0}$ desde la fecha inicial es exagerado en las primeras fechas. Entonces, el valor que se debe añadir a la posición para que ésta sea aceptable pueden mejor ser unos valores que se añaden a (o retiran de) la posición en las diferentes fechas, que van a depender de la medida dinámica que se utilice y la evolución del riesgo de la posición.

Para terminar, como posibles extensiones de este trabajo se tiene que en futuras aplicaciones de esta optimización de portafolios en dos etapas se puede incluir al problema restricciones de retornos esperados, en donde se puede por ejemplo poner

restricciones para asegurar que desde una fecha dada los retornos esperados sean positivos o mayores que un retorno dado.

Se puede también restringir los valores del vector de decisión \mathbf{x} para buscar que los resultados sean aplicables en la vida real. Por ejemplo, los contratos sobre opciones de Cobre y Aluminio en la bolsa *London Mercantile Exchange* son solamente de 25 toneladas, así que la solución obtenida debe ser en múltiplos de 25.

También se tiene que es sencillo aplicar el segundo enfoque a otras medidas dinámicas de riesgo. Por ejemplo, el valor de α puede cambiar en las diferentes fechas. Entonces para $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_T)$, diferentes niveles de confianza α_i en el periodo i , si $\text{CVaR}_{\alpha_i}(\cdot) = \mathbb{E}_{\mathbb{Q}_i}[\cdot]$, se pueden definir las medidas $\text{MaxCVaR}_{\boldsymbol{\alpha}}$ y $\overline{\text{CVaR}}_{\boldsymbol{\alpha}}$ por

$$\text{MaxCVaR}_{\boldsymbol{\alpha},t}(L) = \max_{i=t+1,\dots,T} \mathbb{E}_{\mathbb{Q}_i}[L_i | \mathcal{F}_t] \quad \text{y} \quad \overline{\text{CVaR}}_{\boldsymbol{\alpha},t} = \sum_{i=t+1}^T \lambda_i^t \mathbb{E}_{\mathbb{Q}_i}[L_i | \mathcal{F}_t].$$

Esto puede ser interesante por ejemplo para un inversor que quiera tener más confianza sobre las posibles pérdidas en fechas cercanas a la final. En este caso $\alpha_T \geq \alpha_{T-1} \geq \dots \geq \alpha_1$. También pueden ser variables los niveles de confianza, por ejemplo, si CVaR_{α_1} es mayor que un M , entonces α_T es igual a un nivel de confianza dado y, de lo contrario, α_T es igual a un nivel menor.

Se puede también investigar las propiedades de todas estas posibles medidas de riesgo o estudiar una caracterización de las medidas fuertemente coherentes. Además, estudiar si las medidas cumplen con el axioma de *Monotonidad temporal* es interesante para la optimización de portafolios, porque si lo cumplen, al minimizar el valor de ψ_0 del portafolio se ayuda a que los valores de ψ_t , con $t > 0$, sean bajos en valor esperado, porque se cumpliría que $\psi_0(L) \geq \mathbb{E}[\psi_t(L)]$.

Por último, el valor de \mathbf{x} no tiene que ser constante en la vida del portafolio. Se podría estudiar la manera de optimizar el problema donde las funciones $F_{\alpha}^t(\mathbf{x}_t, \beta)$, dependen de \mathbf{x}_t , para $t = 1, \dots, T$, con restricciones adicionales a los valores de \mathbf{x}_t , para $t > 1$. Por ejemplo, si $\pi_t(\mathbf{x}_t, \mathbf{L}_t)$ es el valor del portafolio (que debe ser \mathcal{F}_t -medible) en t , para las pérdidas \mathbf{L}_t y el vector \mathbf{x}_t , se puede restringir los valores de \mathbf{x}_t de tal manera que $\mathbf{x}_t \in X$ y se cumpla que no se necesite más ni menos dinero para pasar del portafolio compuesto por \mathbf{x}_t a \mathbf{x}_{t+1} (es decir, que sea un portafolio autofinanciado), esto es, para todo $t > 1$, $\pi_{t+1}(\mathbf{x}_t, \mathbf{L}_{t+1}) = \pi_{t+1}(\mathbf{x}_{t+1}, \mathbf{L}_{t+1})$ c.s.

Apéndice A

ESPERANZA CONDICIONAL

Ver Shreve, 1997.

Definición A.0.1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y \mathcal{G} una sub- σ -álgebra de \mathcal{F} . Sea X una variable aleatoria sobre $(\Omega, \mathcal{F}, \mathbb{P})$. Entonces $\mathbb{E}[X|\mathcal{G}]$ está definida como cualquier variable aleatoria Y que satisfaga lo siguiente,

1. Y es \mathcal{G} -medible.
2. Para todo $A \in \mathcal{G}$, se cumple la “propiedad de promedio parcial”,¹

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P} \quad \text{o} \quad \mathbb{E}[\mathbb{I}_A Y] = \mathbb{E}[\mathbb{I}_A X].$$

Para la existencia y unicidad de dicha variable se utiliza el teorema de Lebesgue-Radon-Nikodym como se muestra a continuación. Primero se enuncia el Teorema (ver Rudin, 1987).

Teorema A.0.1. Si μ y λ son medidas no-negativas, σ -finitas sobre la colección \mathcal{G} y λ es absolutamente continua con respecto a μ , $\lambda \ll \mu$, entonces existe una función Y no negativa y \mathcal{G} -medible, tal que

$$\forall A \in \mathcal{G}, \quad \lambda(A) = \int_A Y d\mu = \mathbb{E}_\mu[Y \mathbb{I}_A].$$

Y se denota por $\frac{d\lambda}{d\mu}$ y se llama la derivada de Radon-Nikodym de λ con respecto a μ .

¹ “Partial averaging property”

Ahora, suponga $\mathbb{E}[|X|] < \infty$ (es decir, $X \in L^1$). Sea $\mathbb{P}_{\mathcal{G}}$ la medida \mathbb{P} restringida a \mathcal{G} . Es decir, como $\text{id} : (\Omega, \mathcal{F}) \rightarrow (\Omega, \mathcal{G})$ es medible, $\mathbb{P}_{\mathcal{G}}(B) \triangleq \mathbb{P}(\text{id}^{-1}(B)) = \mathbb{P}(B)$, para $B \in \mathcal{G}$. Se define la nueva medida \mathbb{Q}_X por

$$\mathbb{Q}_X(A) \triangleq \int_A X d\mathbb{P}_X = \int_A X d\mathbb{P} = \mathbb{E}[X \mathbb{I}_A], \quad \text{para } A \in \mathcal{G}.$$

Claramente $\mathbb{Q}_X \ll \mathbb{P}_{\mathcal{G}}$ y si X es no negativa, entonces \mathbb{Q}_X también lo es. Suponga que X es no negativa. Entonces, el teorema de Radon-Nikodym dice que existe una variable aleatoria $\frac{d\mathbb{Q}_X}{d\mathbb{P}_{\mathcal{G}}}$, la derivada de Radon-Nikodym, que es única en el sentido que cualesquiera dos variables que cumplan esto son iguales casi siempre bajo la medida $\mathbb{P}_{\mathcal{G}}$ y es tal que $\frac{d\mathbb{Q}_X}{d\mathbb{P}_{\mathcal{G}}}$ es \mathcal{G} -medible (lo cual es una condición más fuerte que ser \mathcal{F} -medible) y para cualquier $A \in \mathcal{G}$,

$$\int_A X d\mathbb{P} = \int_A X d\mathbb{P}_{\mathcal{G}} = \int_A \frac{d\mathbb{Q}_X}{d\mathbb{P}_{\mathcal{G}}} d\mathbb{P}_{\mathcal{G}} = \int_A \frac{d\mathbb{Q}_X}{d\mathbb{P}} d\mathbb{P}.$$

Por esto anterior se define $\mathbb{E}[X|\mathcal{G}]$ como alguna variable aleatoria fija de la clase de equivalencia de la derivada de Radon-Nikodym, $\frac{d\mathbb{Q}_X}{d\mathbb{P}_{\mathcal{G}}}$.

Ahora, suponga que Z es \mathcal{G} -medible, entonces

$$\mathbb{E}[Z X|\mathcal{G}] = Z \mathbb{E}[X|\mathcal{G}].$$

En efecto, sea $Y = Z \mathbb{E}[X|\mathcal{G}]$, luego Y es \mathcal{G} -medible por ser el producto de dos variables \mathcal{G} -medibles. Sea $A \in \mathcal{G}$ un conjunto cualquiera, se verificará que

$$\int_A Z \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_A Z X d\mathbb{P},$$

por lo que $\mathbb{E}[Z X|\mathcal{G}] = Z \mathbb{E}[X|\mathcal{G}]$ $\mathbb{P}_{\mathcal{G}}$ -c.s. Suponga que $Z = \mathbb{I}_B$ para algún $B \in \mathcal{G}$. Se tiene que

$$\mathbb{E}[\mathbb{I}_A \mathbb{I}_B \mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[\mathbb{I}_{A \cap B} \mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[\mathbb{I}_{A \cap B} X] = \mathbb{E}[\mathbb{I}_A \mathbb{I}_B X],$$

por la “propiedad de promedio parcial”, ya que $A \cap B \in \mathcal{G}$. Entonces se tiene el resultado para el caso en que Z es una función indicadora. Por esto, se cumple para funciones simples, es decir, si $Z = \sum_{i=1}^n c_i \mathbb{I}_{A_i}$, para algún n y algunos $A_i \in \mathcal{G}$, dado que se tiene que $\mathbb{E}[a_1 X_1 + a_2 X_2|\mathcal{G}] = a_1 \mathbb{E}[X_1|\mathcal{G}] + a_2 \mathbb{E}[X_2|\mathcal{G}]$. Como se tiene para

funciones simples se puede verificar para funciones medibles no negativas utilizando que para cada Z medible no negativa existen funciones simples s_n tales que $0 \leq s_1 \leq s_2 \leq \dots \leq Z$ y $s_n(\omega) \rightarrow Z(\omega)$ si $n \rightarrow \infty$, para todo $\omega \in \Omega$ y utilizando el teorema de convergencia monótona de Lebesgue. Para funciones medibles reales se verifica directamente utilizando que $Z = Z^+ - Z^-$, donde $Z^+ = \max\{Z, 0\}$ y $Z^- = \max\{-Z, 0\}$, que son funciones medibles no negativas.

Si \mathcal{H} es una sub- σ -álgebra de \mathcal{G} , entonces $\mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{H}] = \mathbb{E}[X | \mathcal{H}]$. En efecto, sea $Z = \mathbb{E}[X | \mathcal{G}]$ y $A \in \mathcal{H}$. Se tiene que $\mathbb{E}[X | \mathcal{H}]$ es \mathcal{H} -medible. Se verificará que

$$\mathbb{E}[\mathbb{I}_A \mathbb{E}[X | \mathcal{H}]] = \mathbb{E}[\mathbb{I}_A Z],$$

por lo que $\mathbb{E}[X | \mathcal{H}] = \mathbb{E}[Z | \mathcal{H}] = \mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{H}]$ $\mathbb{P}_{\mathcal{H}}$ -c.s. $A \in \mathcal{H}$ implica $A \in \mathcal{G}$, luego por la “propiedad de promedio parcial” se tiene que $\mathbb{E}[\mathbb{I}_A \mathbb{E}[X | \mathcal{G}]] = \mathbb{E}[\mathbb{I}_A X]$. Similarmente, $\mathbb{E}[\mathbb{I}_A Z] = \mathbb{E}[\mathbb{I}_A \mathbb{E}[X | \mathcal{H}]] = \mathbb{E}[\mathbb{I}_A X]$, por lo que el resultado se tiene.

REFERENCIAS

1. <http://www.lchclearnet.com>
2. Hull J.C. (2003). *Options, Futures & other Derivatives*, Prentice Hall.
3. McNeil A. J., Frey, R., Embrechts P. (2004). *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press.
4. Jarrow R. (1998). *Volatility, New Estimation Techniques for Pricing Derivatives*, Risk Books.
5. Shreve S. (1997). *Stochastic Calculus and Finance*, Carnegie Mellon University.
6. Andersson, F., Mausser, H., Rosen, D. y S. Uryasev. (2001) Credit Risk Optimization with Conditional Value-At-Risk Criterion, *Mathematical Programming*, Series B 89, 273-291
7. Artzner, P., Delbaen, F., Eber J-M. y Heath D. (1999). Coherent measures of risk, *Mathematical Finance*, **Vol. 9, No 3**, 203–228.
8. Artzner, P., Delbaen, F., Eber J-M., Heath D. y Ku Hyejin (2003). Multiperiod Risk and Coherent Multiperiod Risk Measurement, *working paper, ETH Zurich*.
9. Artzner, P., Delbaen, F., Eber J-M., Heath D. y Ku Hyejin (2004). Coherent Multiperiod Risk Adjusted Values and Bellman's Principle, Manuscript, Institut de Recherche Mathématique Avancée, Strasbourg.
10. Balbás A., Garrido J., Mayoral s. (2002) Coherent risk measures in a dynamic framework, Working Paper, Madrid Carlos III University.
11. Cheridito, P., Delbaen, F., Kupper, M. (2004) Dynamic monetary risk measures for bounded discrete-time processes. Working Paper. ETH-Zürich
12. Delbaen, F (2000). Coherent Risk Measures on General Probability Spaces, *Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann*, in: K. Sandmann and P. J. Schönbucher (eds.), Springer-Verlag (2002), 1-37.

13. Delbaen, F. (2002). Coherent risk measures, Monograph, Scuola Normale Superiore, Pisa.
14. Detlefsen K., Scandolo G. (2005) Conditional and dynamic convex risk measures, *Humboldt University*.
15. Hardy, M., Wirch, J. L. (1999). A Synthesis of Risk Measures for Capital Adequacy, *Insurance: Mathematics and Economics*, **25**, 337–347.
16. Hardy, M., Wirch, J. L. (2004). The iterated CTE: A dynamic risk measure, *North American Actuarial Journal*, 62–75.
17. Huang, D., Zhu, S-S., Fabozzi, F. J., Fukushima, M. (2006). Robust CVaR Approach to Portfolio Selection with Uncertain Exit Time, Technical Report 2006-1, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, <http://www.amp.i.kyoto-u.ac.jp/tecrep/>.
18. Markowitz, H.M. (1952). Portfolio Selection, *Journal of Finance*, 7 (1), 77–91.
19. Riedel F. (2004). Dynamic Coherent Risk Measures. *Processes and Applications*, **112**, 185–200.
20. RiskmetricsTM - Technical Document, 4th ed., Morgan Guaranty Trust Co. (1996).
21. Rockafellar, R.T., Uryasev, S. (2002) Conditional Value-at-Risk for General Loss Distributions, *Journal of Banking and Finance*, 26/7, 1443-1471
22. Rockafellar, R.T., Uryasev, S. (2000) Optimization of Conditional Value-At-Risk, *The Journal of Risk*, Vol. 2, No. 3, 21-41
23. Uryasev, S. (2000) Conditional Value-at-Risk: Optimization Algorithms and Applications, *Financial Engineering News*, No. 14, 1-5
24. Wang, T. (2000). A Class of Dynamic Risk Measures. Working paper, University of British Columbia.
25. Zhu, S.S., Fukushima, M. (2005) Worst-Case Conditional Value-at-Risk with Application to Robust Portfolio Management, Technical Report 2005-6, Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, <http://www.amp.i.kyoto-u.ac.jp/tecrep/>.