

Medidas monetarias de riesgo: robustez estadística de CVaR y su cálculo explícito para leyes estables

Reynaldo López Tejeda¹
Erick Treviño Aguilar ²
Centro de investigación en Matemáticas, CIMAT A.C.
Guanajuato, Gto.

Firma del Investigador

10 de noviembre de 2012

¹Estudiante de la Maestría en Probabilidad y Estadística CIMAT, reynaldo@cimat.mx.

²Investigador del área de Economía de la Universidad de Guanajuato, erick.trevino@ugto.org.

Dedicatoria

A mi Dios por llenarme de bendiciones, entre ellas esta experiencia.

A mi familia, por siempre estar conmigo y seguirme impulsando en cada etapa.

A Bivi, por ser mi equipo, mi motivación, mi compañera.

A mis amigos por su apoyo, especialmente a Sauri.

Agradecimientos

Gracias al Dr. Erick Treviño por enseñarme tanto y por su empeño en obtener el mejor aprendizaje posible.

Al Dr. Víctor Pérez-Abreu por su paciencia y apoyo durante este proceso.

Agradezco el apoyo económico recibido de CONACyT y de CIMAT.

Índice general

1. Introducción	7
2. Medidas de riesgo monetario convexas	11
2.1. Definiciones y primeros resultados	11
2.2. Teorema de representación en L^∞	16
2.3. Valor en Riesgo y Valor en Riesgo Condicionado	21
2.4. Medidas de riesgo invariantes en ley	22
2.4.1. Distorsiones cóncavas	22
2.4.2. Medidas de riesgo comonótonas	27
3. Distribuciones α-estables	29
3.1. Definición y caracterizaciones	30
3.2. Parametrizaciones	33
3.2.1. Parametrización A	33
3.2.2. Parametrización M	34
3.2.3. Parametrización B	34
3.2.4. Parametrización C	35
3.2.5. Parametrización E	35
3.3. Propiedad de invarianza	36
3.4. Representación integral de la densidad	36
3.4.1. Resultados preparatorios	37
3.4.2. Demostración del Teorema 56	45
4. CVaR para leyes α-estables	47
4.1. Lemas preparativos	48
4.2. Demostración del Teorema de Representación	51
5. Robustez cualitativa	59
5.1. Medidas empíricas	60
5.2. Espacio de observaciones	62

5.3. Métrica de Prokhorov	62
5.3.1. El Teorema de Strassen	63
5.4. Teorema de Robustez	68
5.5. Estructura de espacio de probabilidad para \mathfrak{F}_n	71
6. Robustez comparativa	75
6.0.1. Espacios de Orlicz y consistencia de estimadores	76
6.0.2. Continuidad de \mathcal{R}_ρ	79
6.1. Robustez comparativa	84
6.1.1. Iqr para distorsiones cóncavas	90
7. Conclusiones	93
A. Implementación de CVaR	97

Capítulo 1

Introducción

La regulación bancaria es uno de los temas centrales en política monetaria de un país, puesto que un sistema financiero en buen funcionamiento es clave para el desarrollo económico de una sociedad. Un tema muy importante dentro de la regulación bancaria son los requerimientos de capital. Estos consisten en definir un monto mínimo de reserva, que cada banco debe tener por cada inversión realizada. El propósito es controlar el riesgo derivado de dicha inversión. En este contexto, es necesario examinar los componentes que dan lugar a la definición de dichos requerimientos, en particular la manera en la que se mide el riesgo que se pretende controlar. La comunidad internacional ha analizado algunos aspectos del mercado financiero, con el fin de encontrar una manera estandarizada de medir el riesgo, y que sea adecuada a las necesidades de los usuarios y las instituciones bancarias. Como resultado, se tienen los Acuerdos de Basilea [22], donde se exponen una serie de medidas regulatorias para el sistema financiero en general. En estos acuerdos, se establece la función Valor en Riesgo (VaR) como medida de riesgo estándar para la obtención de capitales de regulación.

Sin embargo, desde la academia se han hecho diferentes señalamientos a VaR, con respecto a su falta de poder descriptivo para algunos aspectos fundamentales, tales como la diversificación del riesgo. Como respuesta a algunos de estos inconvenientes, se ha propuesto el Valor en Riesgo Condicionado (CVaR).

Por lo anterior, es interesante estudiar el comportamiento de VaR y CVaR bajo distintas circunstancias. En este trabajo se pretende estudiar los alcances descriptivos de VaR y CVaR. Lo anterior se realizará considerando dos aspectos fundamentales.

El primer aspecto es la susceptibilidad de ambas medidas, de ser calculadas explícitamente. Esto se estudiará en el caso de distribuciones estables. Estas leyes tienen propiedades de convergencia, colas pesadas y autosimilitud, que son convenientes en finanzas. De hecho, la propiedad de autosimilitud caracteriza a los procesos estables. La característica fractal o autosimilar, dentro del ambiente financiero, es discutida por Mandelbrot en [20], y constituye una ventaja de estas distribuciones.

La segunda cuestión a considerar, es la robustez de las medidas de riesgo. Algunos autores, como Chernobai et al. [3], señalan que CVaR no es conveniente, al no ser robusta. Entonces consideramos un nuevo concepto de robustez, propuesto por Krätschmer, Schied y Zähle en [18].

A lo largo de este trabajo, revisamos varios conceptos relacionados con las medidas monetarias de riesgo convexas. Como hemos discutido, estas funcionales son útiles en el ámbito financiero y se ha desarrollado una extensa investigación en relación a ellas. Una de las características principales de una medida de riesgo, es la naturalidad con la que pueden ser utilizadas para obtener requerimientos de capital, que ayuden a mitigar el riesgo de una inversión. Asimismo, se pueden utilizar para optimizar portafolios, de acuerdo al riesgo de la posición.

En este ambiente, surge el Valor en Riesgo, como una medida que tiene la cualidad de ser susceptible a ser calculada o aproximada con precisión. Esta medida de riesgo (o alguna forma equivalente), ha sido utilizada desde antes de la formalización de la teoría de medidas de riesgo monetario. En el Capítulo 2 se hace un estudio de los conceptos fundamentales de medidas de riesgo monetario. El enfoque es sobre las medidas de riesgo convexas e invariantes en ley. En particular, se presentan las medidas Valor en Riesgo (VaR) y Valor en Riesgo Condicionado (CVaR). Para el caso de medidas definidas en L^∞ , un teorema de representación es demostrado. En un contexto más general, se analizan las medidas comonótonas y las que son producto de distorsiones cóncavas. Para las medidas de riesgo convexas y comonótonas, también se presenta un teorema de representación. Las medidas de riesgo de distorsiones cóncavas serán importantes en el estudio realizado en el Capítulo 5.

Para el Capítulo 3, presentamos un análisis de las distribuciones α -estables. Presentamos un teorema de representación integral, para las funciones de densidad de tales distribuciones. Los argumentos utilizados, se centran en la utilidad del Teorema de Lévy-Khintchine, el Teorema de Inversión de Fourier y la integración compleja, a través del Teorema integral de Cauchy. Cabe mencionar que las diferentes representaciones de las funciones características de las leyes estables, son conocidas y han sido

objeto de estudio. A pesar de ello, no se conocen expresiones simples para las funciones de distribución. Este resultado no es solamente interesante en el sentido teórico, sino que, como se podrá observar, es útil numéricamente. A través de este resultado, se pueden calcular cuantiles y, en consecuencia, el Valor en Riesgo.

Sin embargo, el tener una expresión para el Valor en Riesgo no es suficiente. Mucho se ha debatido sobre los inconvenientes de utilizar esta medida de riesgo. Uno de los problemas es que VaR no es convexa. La convexidad en una medida de riesgo es importante, dado que representa la idea de que diversificar una inversión, reduce el riesgo. Otro punto en contra, es que no pone especial atención a las distribuciones con colas pesadas. En la reciente crisis del 2008, se ha encontrado que desestimar los efectos de los eventos extremos, puede acarrear difíciles consecuencias. Por estos y otros factores, se ha optado por investigar acerca de medidas de riesgo alternativas, para solventar estos puntos débiles. Sin embargo, estas nuevas propuestas no siempre son tan fáciles de manejar como VaR. Mucho menos si se toma en cuenta que se tienen expresiones exactas para las leyes gaussianas, que son las que se recomienda utilizar en los Acuerdos de Basilea.

Por tales motivos, es interesante buscar expresiones sencillas de manejar, para medidas de riesgo alternativas. Una de las más recomendadas es el CVaR, la cual es una medida de riesgo convexa, que da mayor peso al efecto de las colas pesadas.

En el Capítulo 4 se muestra una expresión integral para CVaR, en el caso de las leyes estables. Las variables aleatorias α -estables han sido consideradas adecuadas para modelar los rendimientos de un portafolio. Esta clase de leyes cuentan con miembros que exhiben colas pesadas. Se conoce además que estas leyes cumplen con criterios similares al Teorema del Límite Central, muy adecuado en el análisis asintótico. La propiedad de autosimilitud, estudiada por Mandelbrot en el contexto financiero, es una de las características únicas de estas distribuciones. A pesar de que en tal caso tendríamos que hablar sobre procesos α -estables y no solamente variables aleatorias, es conveniente saber que se cuenta con esa posibilidad. La expresión integral de CVaR del Capítulo 4 es muy conveniente en el aspecto numérico. Tal es su facilidad de manejo, que en el Apéndice se muestra la implementación de esta fórmula, haciendo uso del software estadístico R.

Siguiendo en el debate entre el uso de Var, CVaR u otra medida de riesgo, se ha criticado que la medida CVaR no es robusta, mientras que VaR sí lo es; ver e.g., Cont et al. [4]. La robustez es una propiedad estadística, que se

corresponde con la idea de que pequeñas perturbaciones en la distribución a estimar, produzcan pequeñas perturbaciones en el estimador.

Esta idea es importante, ya que, como se ha mencionado, las medidas de riesgo se pueden pensar como requerimientos de capital. Si el requerimiento es artificialmente alto, la medida provoca que no se aprovechen todas las potencialidades del mercado. Por otro lado si es bajo, puede haber un exceso de riesgo en el sistema, debido a las propiedades de la medida de riesgo en uso. Este análisis es posible para las medidas de riesgo que son invariantes en ley, para las cuales existe un mapeo \mathcal{R}_ρ , tal que $\mathcal{R}_\rho(\mathbb{P} \circ X^{-1}) = \rho(X)$.

El concepto de robustez, así como su caracterización más importante, se debe a Hampel y su artículo publicado en 1971. En este trabajo, hacemos una revisión profunda del celebrado Teorema de Hampel. De hecho, estaremos estudiando una versión actualizada y más general. El Capítulo 5 consiste en este estudio, además de otro resultado importante: el Teorema de Strassen.

De lo anterior, podría parecer muy riesgoso el uso de CVaR en la práctica cotidiana. Empero la robustez es también un concepto sujeto de estudio. En este sentido, en trabajos recientes realizados por Krätschmer, Schied y Zähle, se ha propuesto un nuevo enfoque a la robustez, el cual se analiza en el Capítulo 6. Bajo este nuevo paradigma, las funciones de Young y las correspondientes topologías débiles que generan, son el concepto crucial. Se presenta un resultado análogo al Teorema de Hampel, que relaciona robustez con cierto tipo de continuidad del estimador. Estas ideas generan flexibilidad en cuanto a la manera de clasificar estimadores, de acuerdo a su robustez. Inclusive, se tiene un índice de robustez, que mostraremos es adecuado para clasificar a las medidas de riesgo, con distintos grados de robustez. Bajo este concepto, CVaR tiene cierto grado de robustez. Finalizamos el capítulo con expresiones de dicho índice para medidas de riesgo convexas, y para las que son producto de distorsiones cóncavas.

Capítulo 2

Medidas de riesgo monetario convexas

Las medidas de riesgo monetario son funcionales que nos ayudan a cuantificar el riesgo de una posición financiera. El estudio de dichas medidas ha sido intensivo en los últimos años. En consecuencia, se encuentran expresiones cada vez más sofisticadas, con el fin de mejorar su capacidad de modelación. En este capítulo introduciremos el concepto de medida de riesgo monetario, así como algunas de sus propiedades. El desarrollo de este capítulo sigue la exposición realizada por Föllmer y Schied [11]. Entre los resultados a estudiar, destacan los teoremas de representación, los cuales dejan ver la versatilidad de estas funcionales. Es de particular interés el teorema de representación en L^∞ . Asimismo se estudian las medidas invariantes en distribución y algunas de sus subclases.

2.1. Medidas de riesgo monetario: Definiciones y primeros resultados

Comenzamos estableciendo el contexto en el que se desarrolla este capítulo. Asimismo estudiamos distintas clases de medidas de riesgo, junto con algunos ejemplos.

Vamos a considerar un conjunto Ω de escenarios. Una posición financiera será representada como una función $X : \Omega \rightarrow \mathbb{R}$, donde $X(\omega)$ se entiende como las ganancias o pérdidas de la posición al final del periodo, de ocurrir el escenario ω . Denotamos como \mathcal{X} al espacio lineal de funciones acotadas, incluyendo las constantes.

Definición 1. Un mapeo $\rho : \mathcal{X} \rightarrow \mathbb{R}$ se llama **medida de riesgo monetario**, si satisface las siguientes condiciones para toda $X, Y \in \mathcal{X}$:

1. *Monotonía:* Si $X \leq Y$, entonces $\rho(X) \geq \rho(Y)$.
2. *Invarianza monetaria:* Si $m \in \mathbb{R}$, entonces $\rho(X + m) = \rho(X) - m$.

Se asumirá además que ρ satisface la condición de:

3. *Normalización:* $\rho(0) = 0$.

La definición anterior cobra sentido cuando analizamos las tres propiedades que la conforman, desde el punto de vista financiero. Tengamos en mente que la función X representa las ganancias potenciales de un portafolio. Entonces la propiedad de monotonía indica que el riesgo de un portafolio, es menor que otro cuando invariablemente asegura más ganancias. La segunda propiedad es de gran interés en el contexto de regulación. Si una medida de riesgo es invariante, entonces existe una cantidad de capital finita que neutraliza el riesgo de la inversión. En tal caso, las entidades reguladoras pueden utilizar estas herramientas para controlar el riesgo de determinado portafolio. La tercera propiedad expresa que la ausencia de inversión, será equivalente a la ausencia de riesgo en la posición. Ahora veamos un primer resultado relacionado con medidas de riesgo.

Lema 2. Toda medida de riesgo monetario ρ es Lipschitz continua con respecto a la norma del supremo $\|\cdot\|$:

$$|\rho(X) - \rho(Y)| \leq \|X - Y\|.$$

Demostración. Dado que $X \leq Y + \|X - Y\|$, se tiene que $\rho(Y) - \|X - Y\| \leq \rho(X)$ por monotonía e invarianza. Intercambiando el papel de X y Y se obtiene la desigualdad en el sentido inverso.

El lema anterior nos será de utilidad para relacionar a ρ con un subconjunto de \mathcal{X} y viceversa. En seguida definimos la clase de medidas de riesgo más relevante en este trabajo, las cuales son las medidas convexas.

Definición 3. Una medida de riesgo monetario $\rho : \mathcal{X} \rightarrow \mathbb{R}$ se dice **medida de riesgo convexa** si satisface la siguiente propiedad

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y), \quad \forall 0 \leq \lambda \leq 1.$$

Como caso especial de medidas de riesgo que satisfacen la propiedad de convexidad, tenemos las medidas coherentes.

Definición 4. Una medida de riesgo convexa ρ es coherente si se satisface:

$$\text{Homogeneidad positiva: } \rho(\lambda X) = \lambda \rho(X), \quad \forall \lambda \geq 0.$$

Podemos observar que para una posición fija X , es posible tener diferentes mediciones de riesgo, dependiendo de la definición de ρ . En contraparte, si la medida de riesgo es fija, entonces de manera natural surge una clasificación de los miembros de \mathcal{X} .

Definición 5. A la clase \mathcal{A}_ρ en \mathcal{X} definida como

$$\mathcal{A}_\rho := \{X \in \mathcal{X} | \rho(X) \leq 0\},$$

se le conoce como el conjunto de aceptación de ρ .

Toda medida de riesgo induce un conjunto de aceptación. Es también posible ir en la dirección inversa y comenzar con una clase \mathcal{A} de posiciones que se consideran aceptables, la medida de riesgo que induce se define de la siguiente forma:

$$\rho_{\mathcal{A}}(X) = \inf\{m \in \mathbb{R} | m + X \in \mathcal{A}\}. \quad (2.1)$$

El siguiente resultado nos proporciona condiciones necesarias y suficientes para que un subconjunto $\mathcal{A} \subset \mathcal{X}$ defina una medida de riesgo.

Proposición 6. Supongamos que $\mathcal{A} \subset \mathcal{X}$ es un conjunto no vacío que satisface las siguientes propiedades

$$\inf\{m \in \mathbb{R} | m \in \mathcal{A}\} > -\infty. \quad (2.2)$$

$$X \in \mathcal{A}, Y \in \mathcal{X}, Y \geq X \implies Y \in \mathcal{A}. \quad (2.3)$$

Entonces el funcional $\rho_{\mathcal{A}}$ tiene las siguientes propiedades:

1. $\rho_{\mathcal{A}}$ es una medida de riesgo monetario.
2. Si \mathcal{A} es un conjunto convexo, entonces $\rho_{\mathcal{A}}$ es una medida de riesgo convexa.
3. Si \mathcal{A} es un cono, entonces $\rho_{\mathcal{A}}$ es homogénea positiva. En particular $\rho_{\mathcal{A}}$ es una medida de riesgo coherente si \mathcal{A} es un cono convexo.
4. \mathcal{A} es un subconjunto de $\mathcal{A}_{\rho_{\mathcal{A}}}$ y $\mathcal{A} = \mathcal{A}_{\rho_{\mathcal{A}}}$ si y sólo si \mathcal{A} es $\|\cdot\|$ -cerrado en \mathcal{X} .

Demostración.

1. De la definición de \mathcal{A} se tiene que $\rho_{\mathcal{A}}$ es monótona e invariante. Ahora fijemos $Y \in \mathcal{A}$. Si $X \in \mathcal{X}$ existe un número finito m tal que $m + X > Y$, entonces

$$\rho_{\mathcal{A}}(X) - m = \rho_{\mathcal{A}}(m + X) \leq \rho_{\mathcal{A}}(Y) \leq 0,$$

de donde $\rho_{\mathcal{A}}$ es finita para todo $X \in \mathcal{X}$ y por lo tanto es una medida de riesgo monetario.

2. Tomemos $X_1, X_2 \in \mathcal{X}$ y $m_1, m_2 \in \mathbb{R}$ tales que $m_i + X_i \in \mathcal{A}$. Sea $\lambda \in [0, 1]$, por convexidad de \mathcal{A} , tenemos que

$$\lambda(m_1 + X_1) + (1 - \lambda)(m_2 + X_2) \in \mathcal{A}.$$

Entonces

$$\begin{aligned} 0 &\geq \rho_{\mathcal{A}}(\lambda(m_1 + X_1) + (1 - \lambda)(m_2 + X_2)) \\ &= \rho_{\mathcal{A}}(\lambda X_1 + (1 - \lambda)X_2) - (\lambda m_1 + (1 - \lambda)m_2). \end{aligned}$$

Reorganizando los términos de la ecuación anterior, observamos la convexidad de $\rho_{\mathcal{A}}$.

3. De manera análoga al inciso 2, se tiene que $\rho_{\mathcal{A}}(\lambda X) \leq \lambda \rho_{\mathcal{A}}(X)$ para $\lambda \geq 0$ si \mathcal{A} es un cono. Ahora sea $m < \rho_{\mathcal{A}}(X)$, entonces $m + X \notin \mathcal{A}$ y entonces $\lambda m + \lambda X \notin \mathcal{A}$ para $\lambda \geq 0$, con lo cual $\lambda m \leq \rho_{\mathcal{A}}(\lambda X)$ y se tiene el resultado.
4. La contención $\mathcal{A} \subset \mathcal{A}_{\rho_{\mathcal{A}}}$ es clara. Ahora suponga que el conjunto \mathcal{A} es $\|\cdot\|$ -cerrado en \mathcal{X} . Para $X \notin \mathcal{A}$ sea m tal que $m > \|X\|$. Existe $\lambda \in (0, 1)$ tal que $\lambda m + (1 - \lambda)X \notin \mathcal{A}$, debido a que el conjunto \mathcal{A} es cerrado en norma. Así,

$$0 \leq \rho_{\mathcal{A}}(\lambda m + (1 - \lambda)X) = \rho_{\mathcal{A}}((1 - \lambda)X) - \lambda m.$$

Además

$$|\rho_{\mathcal{A}}((1 - \lambda)X) - \rho_{\mathcal{A}}(X)| \leq \lambda \|X\|.$$

De lo anterior se concluye que

$$\rho_{\mathcal{A}}(X) \geq \rho_{\mathcal{A}}((1 - \lambda)X) - \lambda \|X\| \geq \lambda(m - \|X\|) > 0.$$

Esto demuestra que X no pertenece a $\mathcal{A}_{\rho_{\mathcal{A}}}$. □

Es interesante analizar las propiedades del nuevo objeto \mathcal{A}_ρ y relacionarlas con las de la medida de riesgo que representa.

Proposición 7. *Sea ρ una medida de riesgo monetario con conjunto de aceptación $\mathcal{A} := \mathcal{A}_\rho$. Se satisfacen los siguientes enunciados:*

1. \mathcal{A} es no vacío, cerrado en \mathcal{X} con respecto a la norma supremo $\|\cdot\|$ y satisface las condiciones (2.2) y (2.3).
2. ρ está caracterizada por \mathcal{A} :

$$\rho(X) = \rho_{\mathcal{A}_\rho}(X) = \inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}\}.$$

3. ρ es convexa si y sólo si \mathcal{A} es convexo.
4. ρ es homogénea positiva si y sólo si \mathcal{A} es un cono. En particular ρ es coherente si y sólo si \mathcal{A} es un cono convexo.

Demostración.

1. Las propiedades (2.2) y (2.3) se tienen de la definición de ρ . El conjunto \mathcal{A} es cerrado en norma, debido a que ρ es Lipschitz continua.
2. De la propiedad de invarianza,

$$\begin{aligned} \inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}_\rho\} &= \inf\{m \in \mathbb{R} \mid \rho(m + X) \leq 0\} \\ &= \inf\{m \in \mathbb{R} \mid \rho(X) \leq m\} \\ &= \rho(X). \end{aligned}$$

Lo anterior implica que $\rho_{\mathcal{A}_\rho} = \rho$, donde

$$\rho_{\mathcal{A}}(X) := \inf\{m \in \mathbb{R} \mid m + X \in \mathcal{A}\}.$$

3. Si ρ es una medida de riesgo convexa entonces, para $X, Y \in \mathcal{A}$, se cumple que $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y) \leq 0$. Es decir, el conjunto \mathcal{A} es convexo. Suponga ahora que el conjunto \mathcal{A} es convexo. Entonces $\rho_{\mathcal{A}}$ es convexa como consecuencia a la Proposición 6, inciso dos. Entonces ρ es convexa debido a que $\rho = \rho_{\mathcal{A}_\rho}$.
4. Se demuestra en forma analoga al inciso anterior. □

Veamos dos ejemplos de medidas de riesgo en específico.

Ejemplo 8. Consideramos un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Ahora, $\tilde{X} \in \mathcal{L}^2 = \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ representa la función de pagos de un activo. Dicho activo tiene un precio $\pi(\tilde{X})$ y varianza $\sigma(\tilde{X})$. Entonces la variable aleatoria $X := \tilde{X}(1+r)^{-1} - \pi(\tilde{X})$ describe las ganancias netas descontadas. El cociente de Sharpe se define como

$$\frac{\tilde{X} - \pi(\tilde{X})(1+r)}{\sigma(\tilde{X})} = \frac{\mathbb{E}[X]}{\sigma(X)}.$$

El conjunto de aceptación se conforma de las posiciones X , tales que su cociente de Sharpe está acotado inferiormente por una constante $c > 0$, y tiene la siguiente forma

$$\mathcal{A}_c := \{X \in \mathcal{L}^2 \mid \mathbb{E}[X] \geq c\sigma(X)\}.$$

La correspondiente funcional ρ_c está dada por

$$\rho_c(X) = \mathbb{E}[-X] + c\sigma(X).$$

Ejemplo 9. Localizados en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, fijamos $n \in \mathbb{N}$. Sean $X \in \mathcal{X}$ y X_1, \dots, X_n copias independientes de X . El funcional

$$\rho(X) = -\mathbb{E}[\min(X_1, \dots, X_n)] \quad (2.4)$$

se conoce como MINVAR y se trata de una medida de riesgo coherente en \mathcal{X} .

2.2. Teorema de representación en L^∞

Una medida de riesgo puede ser definida mediante un conjunto de aceptación sin hacer referencia a una estructura probabilista. En esta sección haremos un estudio sistemático de la estructura de las medidas de riesgo. Para este fin, trabajaremos con el espacio de variables esencialmente acotadas definidas en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. El espacio de variables esencialmente acotadas es un espacio de Banach, lo cual nos permitirá adoptar la teoría de dualidad del análisis funcional. En esta teoría, el Teorema de separación de Hahn-Banach juega un papel crucial. Denotamos por \mathcal{X} al espacio de las funciones acotadas medibles en (Ω, \mathcal{F}) . Se tiene que \mathcal{X} con la norma supremo $\|\cdot\|$ es un espacio de Banach. Sean \mathcal{M}_1 el espacio de las medidas de probabilidad en (Ω, \mathcal{F}) y $\mathcal{M}_{1,f}$ el

conjunto de todas las funciones finitamente aditivas $Q : \mathcal{F} \rightarrow [0, 1]$ tales que $Q(\Omega) = 1$. Consideramos las medidas de riesgo tales que

$$\rho(X) = \rho(Y) \quad \text{si } X = Y \text{ } \mathbb{P}\text{-c.s.} \quad (2.5)$$

De lo anterior, el espacio \mathcal{X} es equivalente a L^∞ . Consideremos un funcional $\alpha : \mathcal{M}_{1,f} \rightarrow \mathbb{R} \cup \{\infty\}$ tal que

$$\inf_{Q \in \mathcal{M}_{1,f}} \alpha(Q) \in \mathbb{R}.$$

Para cada $Q \in \mathcal{M}_{1,f}$, la funcional $X \mapsto \mathbb{E}_Q[-X] - \alpha(Q)$ es convexa, monótona e invariante en \mathcal{X} . Entonces

$$\rho(X) := \sup_{Q \in \mathcal{M}_{1,f}} (\mathbb{E}_Q[-X] - \alpha(Q)) \quad (2.6)$$

define una medida de riesgo convexa en \mathcal{X} tal que

$\rho(0) = -\inf_{Q \in \mathcal{M}_{1,f}} \alpha(Q)$. El funcional α se conoce como la función de penalización para ρ en $\mathcal{M}_{1,f}$. Definimos $\alpha^{\min} := \sup_{X \in \mathcal{A}_\rho} \mathbb{E}_Q[-X]$.

Ahora comenzamos el desarrollo del resultado principal. Vemos que las medidas de probabilidad absolutamente continuas con respecto a \mathbb{P} , caracterizan a la funcional α definida previamente.

Lema 10. *Sea ρ una medida de riesgo convexa que cumple (2.5) y que está representada por una función de penalización α como en (2.6). Entonces $\alpha(Q) = +\infty$ para cualquier $Q \in \mathcal{M}_{1,f}(\Omega, \mathcal{F})$ que no sea absolutamente continua con respecto a \mathbb{P} .*

Demostración. *Si $Q \in \mathcal{M}_{1,f}(\Omega, \mathcal{F})$ no es absolutamente continua con respecto a \mathbb{P} , entonces existe $A \in \mathcal{F}$ tal que $Q(A) > 0$ y $\mathbb{P}(A) = 0$. Sean $X \in \mathcal{A}_\rho$ y $X_n := X - n\mathbf{1}_A$. Entonces $\rho(X_n) = \rho(X)$ \mathbb{P} -c.s. Entonces*

$$\alpha(Q) \geq \alpha^{\min}(Q) \geq \mathbb{E}_Q[-X_n] = \mathbb{E}_Q[-X] + nQ(A) \rightarrow \infty.$$

□

Denotaremos por $\mathcal{M}_1(\mathbb{P})$ al conjunto de las medidas de probabilidad en (Ω, \mathcal{F}) que son absolutamente continuas con respecto a \mathbb{P} .

Definición 11. *Una función $f : X \rightarrow \mathbb{R}$ se dice semicontinua por abajo si el conjunto $\{x \in X \mid f(x) < \alpha\}$ es abierto para todo $\alpha \in \mathbb{R}$.*

En este momento entramos a los resultados y requerimientos de la teoría de espacios de Banach. Comenzamos con los conceptos de separación algunas clases de topologías, los cuales serán útiles al momento de establecer el Teorema de Hahn-Banach.

Definición 12. Sea E un espacio topológico. Se dice que un funcional lineal continuo l separa a E si existen dos conjuntos convexos disjuntos \mathcal{B} y \mathcal{C} tales que

$$l(x) \leq l(y) \quad \forall x \in \mathcal{C}, y \in \mathcal{B}.$$

Definición 13. Sean E un espacio lineal y F una clase de funcionales lineales que lo separan. La F -topología en E , denotada por $\sigma(E, F)$ es la topología en E que se obtiene tomando como base todos los conjuntos de la forma

$$\{y \in E \mid |l_i(y) - l_i(x)| < r, \quad i = 1, \dots, n\},$$

donde $n \in \mathbb{N}$, $x \in E$, $l_i \in F$ y $r > 0$. Si E es localmente convexo, entonces para E' de la forma

$$E' := \{l : E \rightarrow \mathbb{R} \mid l \text{ es lineal y continuo}\},$$

la E' -topología se llama topología débil en E .

Podemos considerar a E como un espacio de funcionales lineales de E' mediante $x(l) := l(x)$ para $l \in E'$ y $x \in E$. Entonces la E -topología $\sigma(E', E)$ es llamada la topología débil* en E' . Al espacio E' se le conoce como espacio dual de E . Por ejemplo L^∞ es el dual de L^1 y viceversa, con la topología débil*. El siguiente resultado es el Teorema de separación de Hahn-Banach, el cual se enunciará solamente. Una demostración puede ser revisada en [6].

Teorema 14 (Teorema de Hahn-Banach). Supongamos que \mathcal{B} y \mathcal{C} son dos subconjuntos no vacíos, convexos y disjuntos de un espacio localmente convexo E . Entonces si \mathcal{B} es compacto y \mathcal{C} es cerrado, existe un funcional lineal l en E tal que

$$\sup_{x \in \mathcal{C}} l(x) < \inf_{y \in \mathcal{B}} l(y).$$

Teorema 15. La integral

$$l(F) = \int F d\mu, \quad F \in \mathcal{X}$$

define una correspondencia uno a uno entre las funcionales lineales de \mathcal{X} y las medidas finitamente aditivas de variación finita sobre \mathcal{F} .

A continuación se demuestra el resultado principal de este apartado. En el caso de una medida de riesgo convexa, dado un conjunto de aceptación, se obtiene una representación para ρ . Más aún, vamos a observar que algunas propiedades topológicas del conjunto de aceptación garantizan dicha representación.

Teorema 16 (Representación en L^∞). *Supongamos que $\rho : L^\infty \rightarrow \mathbb{P}$ es una medida de riesgo convexa. Entonces las condiciones siguientes son equivalentes:*

1. ρ puede ser representado por una función de penalización en $\mathcal{M}_1(\mathbb{P})$.
2. ρ puede ser representado por la restricción de α^{\min} a $\mathcal{M}_1(\mathbb{P})$

$$\rho(X) = \sup_{Q \in \mathcal{M}_{1,f}(\mathbb{P})} (\mathbb{E}_Q[-X] - \alpha^{\min}(Q)), \quad X \in L^\infty.$$

3. ρ es continua por la derecha: Si $X_n \searrow X$ \mathbb{P} -c.s. entonces $\rho(X_n) \nearrow \rho(X)$.
4. ρ tiene la propiedad de Fatou: para cualquier sucesión acotada (X_n) que converge \mathbb{P} -c.s. a alguna X ,

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

5. ρ es semicontinua por abajo para la topología débil* $\sigma(L^\infty, L^1)$.
6. El conjunto de aceptación \mathcal{A}_ρ de ρ es débil* cerrado en L^∞ , es decir, \mathcal{A}_ρ es cerrado con respecto a la topología $\sigma(L^\infty, L^1)$.

Demostración. 1. Es claro que 2 implica 1. Veamos que 1 implica 4.
Del teorema de convergencia dominada

$$\begin{aligned} \rho(X) &= \sup_{Q \in \mathcal{M}_1} (\lim_{n \uparrow \infty} \mathbb{E}_Q[-X_n] - \alpha(Q)) \\ &\leq \liminf_{n \uparrow \infty} \sup_{Q \in \mathcal{M}_1} (\mathbb{E}_Q[-X_n] - \alpha(Q)) \\ &= \liminf_{n \uparrow \infty} \rho(X_n). \end{aligned}$$

2. Para ver que 3 es equivalente a 4, notemos que por monotonía $\rho(X_n) \leq \rho(X)$ si $X_n \searrow X$. Entonces $\rho(X_n) \nearrow \rho(X)$. Ahora consideremos una sucesión (X_n) acotada que converge \mathbb{P} -c.s. a X . Definimos $Y_m := \sup_{n \geq m} X_n$ y se tiene que $Y_m \searrow X$. Entonces la propiedad de Fatou se cumple.

3. Veamos que **3** \Rightarrow **5**. Se debe probar que el conjunto $\mathcal{C} := \{\rho \leq c\}$ es débil* cerrado para $c \in \mathbb{R}$. Definimos $\mathcal{C}_r := \mathcal{C} \cap \{X \in L^\infty \mid \|X\|_\infty \leq r\}$ para $r > 0$. Si (X_n) es una sucesión en \mathcal{C}_r que converge en L^1 a alguna X , entonces existe una subsucesión que converge \mathbb{P} -c.s. y $X \in \mathcal{C}_r$. Entonces \mathcal{C}_r es cerrado en L^1 y en consecuencia \mathcal{C} es débil* cerrado.

4. La implicación **5** \Rightarrow **6** es clara.

5. Finalmente probemos **6** \Rightarrow **2**. Fijamos $X \in L^\infty$ y sea

$$m = \sup_{Q \in \mathcal{M}_1(\mathbb{P})} (\mathbb{E}_Q[-X] - \alpha^{\min}(Q)).$$

Supongamos que $m + X \notin \mathcal{A}_\rho$. Dado que \mathcal{A}_ρ es débil* cerrado, podemos utilizar el Teorema de separación de Hahn-Banach en el espacio localmente compacto $(L^\infty, \sigma(L^\infty, L^1))$ con $\mathcal{C} = \mathcal{A}_\rho$ y $\mathcal{B} = \{m + X\}$. Sea $l := \mathbb{E}[YZ]$ para $Y \in \mathcal{A}_\rho$ y alguna $Z \in L^1$, notemos que

$$\beta := \inf_{Y \in \mathcal{A}_\rho} l(Y) > l(m + X) =: \gamma > -\infty. \quad (2.7)$$

Fijemos $Y \geq 0$ y notemos que $\rho(\lambda Y) \leq \rho(0)$ para $\lambda \geq 0$ por monotonía. Entonces $\lambda Y + \rho(0) \in \mathcal{A}_\rho$ y

$$-\infty < \lambda < l(\lambda Y + \rho(0)) = \lambda l(Y) + l(\rho(0)).$$

De lo anterior $l(Y) \geq 0$ y entonces $Z > 0$ con probabilidad positiva. Definimos

$$\frac{dQ_0}{d\mathbb{P}} := \frac{Z}{\mathbb{E}[Z]},$$

la cual es una medida de probabilidad en $\mathcal{M}_1(\mathbb{P})$. De (2.7) se tiene

$$\alpha^{\min}(Q_0) = \sup_{Y \in \mathcal{A}_\rho} \mathbb{E}_{Q_0}[-Y] = \frac{-\beta}{\mathbb{E}[Z]}$$

Pero

$$\mathbb{E}_{Q_0}[X] + m = \frac{l(m + X)}{\mathbb{E}[Z]} = \frac{\gamma}{\mathbb{E}[Z]} < \frac{\beta}{\mathbb{E}[Z]} = -\alpha^{\min}(Q_0).$$

De lo anterior se obtiene el resultado. □

2.3. Valor en Riesgo y Valor en Riesgo Condicionado

Definición 17. Para una variable aleatoria X y $\varepsilon \in (0, 1)$, un cuantil de probabilidad de la función de distribución de X al nivel ε es un número q tal que se satisface

$$\mathbb{P}[X \leq q] \geq \varepsilon \quad \text{y} \quad \mathbb{P}[X < q] \leq \varepsilon.$$

El conjunto de los cuantiles de probabilidad ε es un intervalo de la forma

$$[q_X^-(\varepsilon), q_X^+(\varepsilon)],$$

con

$$\begin{aligned} q_X^-(t) &= \sup\{x | \mathbb{P}[X < x] < t\} = \inf\{x | \mathbb{P}[X \leq x] \geq t\} \\ q_X^+(t) &= \inf\{x | \mathbb{P}[X \leq x] > t\} = \sup\{x | \mathbb{P}[X < x] \leq t\}. \end{aligned}$$

Definición 18. El **Valor en Riesgo de nivel ε** (VaR_ε) se define como

$$VaR_\varepsilon(X) = -q_X^+(\varepsilon) = q_X^-(1 - \varepsilon) = \inf\{m \in \mathbb{R} \mid \mathbb{P}[m + X < 0] \leq \varepsilon\}.$$

En forma equivalente $VaR_\varepsilon(X)$ satisface la siguiente igualdad

$$\mathbb{P}(X \leq -VaR_\varepsilon(X)) = \varepsilon. \quad (2.8)$$

Es conveniente resaltar que VaR_ε no es, en general, una medida de riesgo convexa. Sin embargo, si nos restringimos al espacio de variables aleatorias con distribución normal, VaR_ε es una medida de riesgo coherente.

Para terminar esta sección definimos una de las medidas de riesgo más importantes: el valor en riesgo condicionado. Cabe mencionar que esta medida de riesgo se estudiará con mayor detalle a lo largo de los siguientes capítulos, así como sus propiedades.

Definición 19. El **Valor en Riesgo Promedio de nivel ε** ($AVaR_\varepsilon$), con $\varepsilon \in (0, 1)$ es una funcional de \mathcal{X} a los números reales dada por

$$AVaR_\varepsilon(X) := \frac{1}{\varepsilon} \int_0^\varepsilon VaR_y(X) dy.$$

Algunas veces se llama también **Valor en Riesgo Condicionado** ($CVaR_\varepsilon$) o **Expected Shortfall** (ES_ε). De hecho una manera equivalente de definir el funcional $CVaR_\varepsilon$ es la siguiente

Definición 20.

$$CVaR_\varepsilon(X) = -\mathbb{E}[X | X \leq -VaR_\varepsilon(X)] \quad (2.9)$$

2.4. Medidas de riesgo invariantes en ley

En esta sección se estudian las medidas que no distinguen entre variables aleatorias con la misma distribución. Se presenta el contexto en donde se realiza el análisis y se tendrá mayor interés en dos subclases de tales medidas: las distorsiones cóncavas y las comonótonas. Asimismo revisamos el concepto de integral de Choquet, el cual resulta ser relevante para caracterizar algunas medidas de riesgo. Finalizamos este capítulo con un teorema de representación para medidas comonótonas.

Definición 21. *Una medida de riesgo monetario ρ en $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ se dice invariante en ley si $\rho(X) = \rho(Y)$ si y sólo si X e Y tienen la misma distribución bajo \mathbb{P} .*

Definición 22. *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Un átomo en dicho espacio es un conjunto $A \in \mathcal{F}$ tal que, para cualquier $B \in \mathcal{F}$ con $B \subset A$, se cumple una de las siguientes propiedades: $\mathbb{P}(A) > 0$ y $\mathbb{P}(B) = 0$, o bien $\mathbb{P}(A) = \mathbb{P}(B)$.*

Otra manera de entender un espacio de probabilidad sin átomos, es que es el dominio de una variable aleatoria continua no trivial. Es decir, el espacio tiene cierta regularidad con respecto a la medida de probabilidad. Por lo general, se considera la existencia de una variable aleatoria con distribución uniforme en $(0, 1)$.

2.4.1. Distorsiones cóncavas

La motivación de este y el subsecuente tema, es revisar algunas características de

$$\rho_\mu(X) = \int CVaR_\varepsilon(X) \mu(d\varepsilon), \quad (2.10)$$

que son las mezclas de $CVaR_\varepsilon$ con respecto a una medida de probabilidad μ en $[0, 1]$. El primer resultado importante nos muestra una manera de representar dichas mezclas, para lo cual necesitamos el siguiente lema.

Lema 23. *La siguiente identidad*

$$\psi'_+(t) = \int_{(t,1]} s^{-1} \mu(ds), \quad 0 < t < 1, \quad (2.11)$$

donde ψ'_+ representa la derivada por la derecha de ψ , define una correspondencia inyectiva entre las medidas de probabilidad μ definidas en $[0, 1]$ y las funciones cóncavas crecientes $\psi : [0, 1] \rightarrow [0, 1]$ con $\psi(0) = 0$ y $\psi(1) = 1$. Más aún, $\psi(0+) = \mu(\{0\})$.

Demostración. *En primer lugar tomamos una medida de probabilidad μ en $[0, 1]$. Definamos al mapeo ψ como en (2.11) y $\psi(1) = 1$. Notamos que ψ es creciente cóncava. De hecho*

$$1 - \psi(0+) = \int_0^1 \psi'(t) dt = \int_{(0,1]} \frac{1}{s} \int_0^1 \mathbf{1}_{\{t < s \leq 1\}} dt \mu(ds) = \mu((0, 1]) \leq 1.$$

Con lo anterior podemos definir $\psi(0) = 0$ y obtener una función creciente y cóncava en $[0, 1]$.

Por otro lado, si tenemos una función ψ como en el lema, se tiene que $\psi'_+(t)$ es decreciente y continua por la derecha, en $(0, 1)$. Entonces puede ser escrita como $\psi'_+(t) = \nu((t, 1])$ para alguna medida ν en $(0, 1]$ localmente finita. Definimos una medida μ en $(0, 1]$ como $\mu(dt) = t\nu(dt)$. Claramente se cumple (2.11) y, del teorema de Fubini,

$$\mu((0, 1]) = \int_0^1 \int_{(0,1]} \mathbf{1}_{\{t < s\}} \nu(ds) dt = 1 - \psi(0+) \leq 1.$$

Finalmente denotamos $\mu(\{0\}) = \psi(0+)$ y obtenemos una medida de probabilidad μ en $[0, 1]$. □

La función ψ tomará un lugar relevante para representar a las mezclas ρ_μ . Ahora tenemos un primer teorema de representación.

Teorema 24. *Para una medida de probabilidad μ en $[0, 1]$, sea ψ la función cóncava definida en el Lema 23. Entonces, si $X \in \mathcal{X}$,*

$$\begin{aligned} \rho_\mu(-X) &= \psi(0+) CVaR_0(-X) + \int_0^1 q_X(t) \psi'(1-t) dt \\ &= \int_{-\infty}^0 (\psi(\mathbb{P}[X > x]) - 1) dx + \int_0^\infty \psi(\mathbb{P}[X > x]) dx. \end{aligned} \quad (2.12)$$

Demostración. De la definición de $CVaR_\varepsilon$, el lema anterior y dado que $VaR_\varepsilon(-X) = q_X^-(1 - \varepsilon)$

$$\int_{(0,1]} CVaR_\varepsilon(-X)\mu(d\varepsilon) = \int_0^1 q_X(t)\psi'(1-t)dt.$$

De lo anterior y tomando en cuenta que $\psi(0+) = \mu(\{0\})$, $CVaR_0(-X) = \text{ess sup } X$, se obtiene (2.12). En el caso de la segunda igualdad, supongamos que $X \geq 0$. Entonces si F_X es la función de distribución de X ,

$$q_X^+(t) = \sup\{x \geq 0 | F_X(x) \leq t\} = \int_0^\infty \mathbf{1}_{\{F_X(x) \leq t\}} dx.$$

Utilizando el teorema de Fubini en (2.12) y el hecho de que $\int_0^y \psi'(t)dt = (\psi(y) - \psi(0+))\mathbf{1}_{\{y>0\}}$,

$$\begin{aligned} \int_0^1 q_X(t)\psi'(1-t)dt &= \int_0^\infty \mathbf{1}_{\{F_X(x) \leq 1-t\}} \psi'(t) dt dx \\ &= -\psi(0+) \text{ess sup } X + \int_0^\infty \psi(1 - F_X(x)) dx. \end{aligned}$$

En el caso general, si $X \in L^\infty$, definimos $C = -\text{ess inf } X$. Notemos que $X + C \geq 0$ y aplicamos ρ_μ a $-X - C$ de la siguiente manera

$$\begin{aligned} C + \rho_\mu(-X) &= \int_0^\infty \psi(\mathbb{P}[X > x - C]) dx \\ &= \int_{-C}^0 \psi(\mathbb{P}[X > x]) dx + \int_0^\infty \psi(\mathbb{P}[X > x]) dx \\ &= C + \int_{-\infty}^0 (\psi(\mathbb{P}[X > x]) - 1) dx + \int_0^\infty \psi(\mathbb{P}[X > x]) dx, \end{aligned}$$

la cual es precisamente la igualdad deseada. \square

Una medida de riesgo puede ser definida también a través de una función con características similares a ψ del Lema 23.

Definición 25. Sea $\psi : [0, 1] \rightarrow [0, 1]$ una función creciente tal que $\psi(0) = 0$ y $\psi(1) = 1$. La función

$$c_\psi(A) = \psi(\mathbb{P}[A]), \quad A \in \mathcal{F},$$

se conoce como la distorsión de la medida de probabilidad \mathbb{P} con respecto a la función de distorsión ψ .

Enseguida tenemos el concepto de función submodular, el cual es fundamental en el resultado de representación al final.

Definición 26. Una función $c : \mathcal{F} \rightarrow [0, 1]$ se llama monótona si

$$c(A) \leq c(B) \quad \text{para } A \subset B$$

y normalizada si

$$c(\emptyset) = 0, \quad c(\Omega) = 1.$$

Una función monótona se dice submodular o bien, 2-alternante si

$$c(A \cup B) + c(A \cap B) \leq c(A) + c(B).$$

Claramente cualquier distorsión c_ψ es normalizada y monótona.

El próximo resultado relaciona la concavidad de alguna función ψ , la distorsión que produce y la submodularidad de tal distorsión.

Proposición 27. Sea c_ψ la distorsión de \mathbb{P} con respecto a la función de distorsión ψ . Si ψ es cóncava, entonces c_ψ es submodular. Además, si el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ es sin átomos, se cumple entonces el recíproco.

Demostración. Supongamos que ψ es cóncava y tomemos $A, B \in \mathcal{F}$ tales que $\mathbb{P}[A] \leq \mathbb{P}[B]$. De la concavidad de c_ψ observemos

$$\frac{c(A) - c(A \cap B)}{\mathbb{P}[A] - \mathbb{P}[A \cap B]} \geq \frac{c(A \cup B) - c(B)}{\mathbb{P}[A \cup B] - \mathbb{P}[B]}.$$

Multiplícando ambos lados por $\mathbb{P}[A \cup B] - \mathbb{P}[B]$, de lo cual c_ψ es submodular.

Para probar el recíproco, tomemos una variable aleatoria U con distribución uniforme en $[0, 1]$. Definamos dos conjuntos $A = \{0 \leq U \leq y\}$ y $B = \{z - y \leq U \leq z\}$, para $0 \leq y \leq z \leq 1$. Ahora denotemos $x_1 = \mathbb{P}[A \cap B]$ y $x_2 = \mathbb{P}[A \cup B]$. Se observa que $y = \mathbb{P}[A] = \frac{x_1 + x_2}{2}$. Como ψ es submodular, $\psi(x) + \psi(z) \leq 2\psi(y)$. La desigualdad anterior se cumple para cualesquiera $0 \leq x \leq z \leq 1$. Entonces ψ es cóncava. \square

Definición 28. Sea $c : \mathcal{F} \rightarrow [0, 1]$ una función monótona y normalizada. La integral de Choquet de una función acotada y medible en (Ω, \mathcal{F}) con respecto a c se define como

$$\int X dc = \int_{-\infty}^0 (c(X > x) - 1) dx + \int_0^{\infty} c(X > x) dx.$$

El Teorema 24, en el contexto de distorsiones cóncavas y la integral de Choquet, tiene las siguientes consecuencias directas.

Corolario 29. *Para una medida de probabilidad μ en $[0, 1]$, sean ψ la distorsión cóncava definida en el Lema 23 y c_ψ la distorsión de \mathbb{P} con respecto de ψ . Entonces si $X \in L^\infty$,*

$$\rho_\mu(X) = \int (-X) dc_\psi.$$

Corolario 30. *Una medida de riesgo convexa ρ es invariante en ley y continua por arriba si y sólo si*

$$\rho(X) = \sup_{\psi} \left(\int (-X) dc_\psi - \gamma^{\min}(\psi) \right),$$

donde el supremo se toma sobre la clase de las funciones de distorsión cóncavas ψ y además

$$\gamma^{\min}(\psi) = \sup_{X \in \mathcal{A}_\rho} \int (-X) dc_\psi.$$

Definición 31. *Sea $\beta \geq 1$ y consideremos la función de distorsión cóncava $\psi_\beta(x) = x^{\frac{1}{\beta}}$. Entonces para $\beta \in \mathbb{N}$, se define la medida de riesgo*

$$\text{MAXVAR}_\beta(X) = \int (-X) dc_{\psi_\beta}. \quad (2.13)$$

Se conoce que la medida de riesgo anteriormente definida tiene la siguiente propiedad. Si Y_1, \dots, Y_β son variables aleatorias independientes idénticamente distribuidas tales que $\max(Y_1, \dots, Y_\beta) \stackrel{d}{=} X$, entonces

$$\text{MAXVAR}_\beta(X) = -\mathbb{E}[Y_1].$$

Definición 32. *Sea $\beta \geq 1$ y consideremos la función de distorsión cóncava $\tilde{\psi}_\beta(x) = (1 - (1 - x)^\beta)^{\frac{1}{\beta}}$. Entonces para $\beta \in \mathbb{N}$, se define la medida de riesgo*

$$\text{MAXMINVAR}_\beta(X) = \int (-X) dc_{\tilde{\psi}_\beta}. \quad (2.14)$$

En este caso, si X_1, \dots, X_β son copias independientes de X y Y_1, \dots, Y_β son v.a.i.i.d. tales que $\max(Y_1, \dots, Y_\beta) \stackrel{d}{=} \min(X_1, \dots, X_\beta)$, se tiene que $\text{MAXMINVAR}_\beta(X) = -\mathbb{E}[Y_1]$.

Definición 33. Sea $\beta \geq 1$ y consideremos la función de distorsión cóncava $\hat{\psi}_\beta(x) = (1 - (1 - x)^{\frac{1}{\beta}})^\beta$. Entonces para $\beta \in \mathbb{N}$, se define la medida de riesgo

$$\text{MINMAXVAR}_\beta(X) = \int (-X) dc_{\hat{\psi}_\beta}. \quad (2.15)$$

Si Y_1, \dots, Y_β son v.a.i.i.d. tales que $\max(Y_1, \dots, Y_\beta) \stackrel{d}{=} X$, se tiene que $\text{MINMAXVAR}_\beta(X) = -\mathbb{E}[\min(Y_1, \dots, Y_\beta)]$.

2.4.2. Medidas de riesgo comonótonas

Las medidas de riesgo comonótonas son la clave para representar a las mezclas de CVaR de la forma (2.10). Terminamos el capítulo con la noción de medida comonótona y el resultado de representación antes mencionado.

Definición 34. Dos funciones medibles X y Y en (Ω, \mathcal{F}) se dicen comonótonas si

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0, \quad \forall (\omega, \omega') \in \Omega \times \Omega. \quad (2.16)$$

Una medida de riesgo monetario ρ sobre \mathcal{X} se dice comonótona si

$$\rho(X + Y) = \rho(X) + \rho(Y)$$

para todas $X, Y \in \mathcal{X}$ que sean comonótonas.

El siguiente resultado será de gran utilidad para caracterizar a las medidas de riesgo comonótonas. Se omite la demostración, pero se puede encontrar en [11].

Teorema 35. Una medida de riesgo monetario ρ en \mathcal{X} es comonótona si y sólo si existe una función normalizada c en (Ω, \mathcal{F}) tal que

$$\rho(X) = \int (-X) dc \quad X \in \mathcal{X}.$$

En este caso c está dada por $c(A) = \rho(-\mathbf{1}_A)$.

El teorema anterior y el Corolario 29 revelan la siguiente propiedad: las distorsiones de \mathbb{P} con respecto a la función ψ , a la manera del Lema 23, son comonótonas. Estamos listos para el teorema que relaciona las expresiones (2.10) con un conjunto de medidas de riesgo en particular.

Teorema 36. *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad sin átomos. La clase de medidas de riesgo*

$$\rho_\mu(X) = \int CVaR_\varepsilon(X)\mu(d\lambda), \quad \mu \in \mathcal{M}_1([0, 1]), \quad (2.17)$$

coincide con la clase de las medidas de riesgo convexas invariantes en ley en L^∞ que son comonótonas. En particular, cualquier medida de riesgo convexa que es invariante en ley y comonótona es también coherente y continua por arriba.

Demostración. *La comonotonidad de ρ_μ se sigue del Corolario 29 y el Teorema 35. Por otro lado, consideremos a ρ como una medida de riesgo convexa, invariante en ley y comonótona. Del Teorema 35,*

$$\rho(X) = \int (-X)dc \quad X \in \mathcal{X}, \quad c(A) = \rho(-\mathbf{1}_A), \quad A \in \mathcal{F}.$$

Como ρ es invariante en ley y la distribución de $\mathbf{1}_A$ depende únicamente de la probabilidad $\mathbb{P}[A]$, entonces $c(A) = c(B)$ si y sólo si $\mathbb{P}[A] = \mathbb{P}[B]$. Así, existe una función creciente ψ en $[0, 1]$ tal que $\psi(0) = 0$, $\psi(1) = 1$ y $c(A) = \psi(\mathbb{P}[A])$. Es fácil ver que para cualesquiera $A, B \in \mathcal{F}$, las funciones $\mathbf{1}_{A \cup B}$ e $\mathbf{1}_{A \cap B}$ son comonótonas. Por lo tanto

$$\begin{aligned} c(A \cap B) + c(A \cup B) &= \rho(-\mathbf{1}_{A \cap B}) + \rho(-\mathbf{1}_{A \cup B}) = \rho(-\mathbf{1}_{A \cap B} - \mathbf{1}_{A \cup B}) \\ &= \rho(-\mathbf{1}_A - \mathbf{1}_B) \leq c(A) + c(B), \end{aligned}$$

donde la última igualdad se debe a la subaditividad de ρ como medida de riesgo convexa. Se ha probado que c es submodular y de la Proposición 27, ψ es cóncava. Por medio del Corolario 29, sabemos que la integral de Choquet con respecto a c se identifica con la medida de riesgo ρ_μ , con μ relacionada a ψ por el Lema 23. \square

Capítulo 3

Distribuciones α -estables

Un argumento clásico en la modelación de activos financieros, es que los precios son el resultado de un gran número de decisiones individuales, a la vez que son la suma cumulativa de pequeños choques. Esto ha llevado a modelar los precios con distribuciones gaussianas, formalizado con alguna de las formas del Teorema del Límite Central. Tal es la teoría para la modelación de precios, desarrollada en 1900 por la visión pionera de Louis Bachelier. Sin embargo, estudios posteriores con evidencia empírica destacan que los histogramas de precios, poseen colas más pesadas que los correspondientes a la distribución gaussiana. En respuesta, Mandelbrot [19] propone la clase de distribuciones α -estables, que denotaremos \mathfrak{S} . Una motivación importante, es que esta clase paramétrica genera colas pesadas. Para una selección de los parámetros ($\alpha = 2$) se recupera la distribución gaussiana, y es el único caso en que se tienen “colas ligeras”. Existen otras distribuciones paramétricas que producen colas pesadas tales como la distribución t – *student* o la distribución hiperbólica. Sin embargo, los elementos de la clase \mathfrak{S} , satisfacen resultados generalizados del Teorema del Límite Central. Lo anterior en conexión con el hecho de que definen elementos en la clase \mathfrak{S} , de distribuciones infinitamente divisibles; ver Teorema 41, abajo. Una exposición sistemática con una perspectiva contemporánea de la importancia de colas pesadas en la modelación de activos y cuantificación de riesgos puede verse, e.g., en Hass y Pigorsch [12]. Otra motivación importante para considerar la clase \mathfrak{S} en la modelación de activos financieros es cuando consideramos la dinámica de precios. Se ha observado empíricamente en datos de alta frecuencia, que las trayectorias de precios tienen la propiedad de autosimilitud; ver e.g., Evertsz [9] y Mandelbrot [20].

Una característica importante de las distribuciones α -estables, es que las funciones características están parametrizadas por cuatro parámetros $(\alpha, \beta, \gamma, \lambda)$ vía la representación de Lévy-Khintchine; ver Teorema 40. No obstante, no existe una fórmula cerrada para la función de densidad, ni tampoco para la función de distribución. Nuestro objetivo en este capítulo es presentar la clase \mathfrak{S} y una representación integral de su función de densidad. En este sentido, mencionamos que para la medición de riesgos con la medida de riesgo CVaR, las funciones cuantiles son cruciales. Para este fin, es necesario contar con alguna expresión para la función de densidad.

El capítulo esta basado en el trabajo de Zolotarev [29] y esta organizado de la siguiente forma. En la Sección 3.1 introducimos la clase de distribuciones α -estables \mathfrak{S} . Para tal propósito, comenzamos con la clase de distribuciones infinitamente divisibles. En la Sección 3.2, damos una lista de las diferentes parametrizaciones que se conocen en la literatura de las distribuciones estables. En la Sección 3.4 presentamos la representación integral de la función de densidad de la clase \mathfrak{S} ; ver Teorema 56.

3.1. Definición y caracterizaciones

Definición 37 (Distribuciones infinitamente divisibles). *Consideramos una sucesión $\{X_{n,j}\}_{j=1,\dots,k_n}$ de variables aleatorias independientes, así como una sucesión $\{A_n\}_{n \in \mathbb{N}} \subset \mathbb{R}$. Definamos*

$$Z_n := X_{n,1} + \dots + X_{n,k_n} - A_n.$$

Denotemos por F_n la distribución de Z_n . Entonces \mathfrak{S} será la clase de funciones de distribución, que son el límite débil de sucesiones de funciones de distribución de la forma $\{F_n\}_{n=1}^\infty$.

La caracterización de distribuciones en la clase \mathfrak{S} requerirá la aplicación de la función característica.

Definición 38. *Sea μ una medida de probabilidad en \mathbb{R}^n . Su transformada de Fourier es denotada como $\hat{\mu}$ y es una función en \mathbb{R}^n definida por*

$$\hat{\mu}(u) = \int_{-\infty}^{+\infty} e^{i\langle u, x \rangle} \mu(dx).$$

Sea X un vector aleatorio valuado en \mathbb{R}^n . Su función característica está definida por

$$\phi_X(u) := E[e^{i\langle u, X \rangle}].$$

Observamos que si X tiene función de distribución F entonces $\phi_X(u) = \hat{F}$.

Recordamos ahora el teorema de unicidad de la función característica. Para la demostración ver e.g., Jacod y Protter [15], Teorema 14.1.

Teorema 39. *Sean μ_1 y μ_2 dos medidas de probabilidad en \mathbb{R}^n . Si $\widehat{\mu}_1 = \widehat{\mu}_2$, entonces $\mu_1 = \mu_2$.*

El siguiente resultado caracteriza a la clase \mathfrak{G} .

Teorema 40 (Lévy-Khintchine). *Una función de distribución G pertenece a la clase \mathfrak{G} si y sólo si la función característica \mathfrak{g} tiene la forma*

$$\mathfrak{g}(t) = \exp \left\{ ita - bt^2 + \int_{x \neq 0} (e^{itx} - 1 - it \sin(x)) dH(x) \right\} \quad (3.1)$$

para $a, b \geq 0$ y una función H que es no decreciente en $(-\infty, 0) \cup (0, +\infty)$ y satisface

$$\int_{0 < |x| < 1} x^2 dH(x) < \infty.$$

En este caso decimos que la función H es la función espectral de la distribución G .

El siguiente resultado caracteriza en forma alternativa a la clase \mathfrak{G} e identifica la correspondiente función espectral.

Teorema 41 (Gnedenko). *Definimos*

$$a_n(y) := \sum_{j=1}^{k_n} E[X_{n,j} 1_{\{|X_{n,j}| < y\}}] \quad y$$

$$\sigma_n^\epsilon := \sum_{j=1}^{k_n} \text{Var}[X_{n,j} 1_{\{|X_{n,j}| < \epsilon\}}].$$

Sean G una distribución en la clase \mathfrak{G} y A_n definido de la forma

$$A_n := a_n(y) - a - \int_{|u| < y} u dH(u) + \int_{|u| \geq y} u^{-1} dH(u).$$

Entonces las funciones de distribución F_n convergen débilmente a G si y sólo si

1. En todo punto de continuidad x de la función espectral, se satisface

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{k_n} \left(F_{nj}(x) - \frac{1}{2} - \frac{1}{2} \text{sgn}(x) \right) = H(x).$$

2.

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \sigma_n^\epsilon = \lim_{\epsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \sigma_n^\epsilon = 2b.$$

Como caso especial de las distribuciones infinitamente divisibles tenemos las distribuciones estables.

Definición 42 (Distribuciones estables). *Considere una sucesión $\{X_n\}_{j \in \mathbb{N}}$ de variables aleatorias independientes e idénticamente distribuidas. Considere también dos sucesiones de números reales $\{A_n\}_{n \in \mathbb{N}}$ y $\{B_n\}_{n \in \mathbb{N}}$. Definamos*

$$Z_n := \frac{X_1 + \dots + X_n}{B_n} - A_n.$$

Denotemos por F_n la distribución de Z_n . Supongamos que la sucesión de funciones de distribución $\{F_n\}_{n=1}^\infty$ converge a una función de distribución F . En este caso decimos que F es una distribución estable. La clase de distribuciones estables será denotada como \mathfrak{S} .

Recuerde la definición de operación de convolución entre dos funciones de variable real $f, g : \mathbb{R} \rightarrow \mathbb{R}$:

$$f * g(t) := \int_{-\infty}^{+\infty} f(x)g(t-x)dx. \quad (3.2)$$

Teorema 43 (Caracterización de distribuciones estables). *Una función de distribución G pertenece a la clase \mathfrak{S} si y solo si para todo par de números reales b_1 y b_2 existe un número real b y un número a tal que*

$$G(x/b_1) * G(x/b_2) = G((x-a)/b). \quad (3.3)$$

El caso en que la constante a de la ecuación (3.3) se puede tomar igual a cero es importante y distinguimos esta clase de distribuciones.

Definición 44. *La función de distribución G es **estrictamente estable** si se satisface la ecuación (3.3) con $a = 0$, independientemente de las constantes b_1 y b_2 . A esta clase la denotaremos como \mathfrak{M} .*

En el siguiente resultado vemos la conexión de variables infinitamente divisibles (estables), con la operación de convolución. En particular, la notación μ^{*n} se define recursivamente mediante $\mu^{*n} := \mu * \mu^{*(n-1)}$.

Teorema 45. *Una medida de probabilidad μ en \mathbb{R}^d es **infinitamente divisible** si para todo entero positivo n , existe una medida de probabilidad μ_n en \mathbb{R}^d tal que $\mu = \mu_n^{*n}$.*

Sea μ una medida de probabilidad infinitamente divisible en \mathbb{R}^d . La medida es **estable** si para todo $a > 0$, existen $b > 0$ y $c \in \mathbb{R}^d$ tales que

$$\hat{\mu}(z)^a = \hat{\mu}(bz)e^{i\langle c, z \rangle}.$$

La medida es **estrictamente estable** si para todo $a > 0$, existe $b > 0$ tal que

$$\hat{\mu}(z)^a = \hat{\mu}(bz).$$

Para distribuciones en la clase \mathfrak{S} podemos identificar explícitamente la función espectral.

Teorema 46. *Sea H la función espectral de una distribución G en la clase \mathfrak{S} . Existen constantes no negativas C_1, C_2 y $\alpha \in (0, 2)$ tal que*

$$H(x) := \begin{cases} -C_1 x^\alpha & x > 0, \\ C_2 (-x)^\alpha & x < 0. \end{cases} \quad (3.4)$$

Demostración. ver Zolotarev [29], Teorema C.1, pp.7. □

3.2. Parametrizaciones

La función característica de las leyes infinitamente divisibles cuenta con la representación de Lévy-Khintchine (3.1). En particular, las leyes α -estables tienen varias representaciones de tal tipo, muchas de ellas exploradas por Zolotarev [29].

En este apartado presentamos las diferentes parametrizaciones de la función característica de las distribuciones α -estables.

3.2.1. Parametrización A

Teorema 47. *Una distribución no-degenerada es α -estable con vector de parámetros*

$$(\alpha, \beta, \gamma, \lambda) \in (0, 2] \times [-1, 1] \times (-\infty, +\infty) \times (0, +\infty),$$

si y sólo si el logaritmo de su función característica \mathfrak{g} , puede ser representado como

$$\log \mathfrak{g}(t) = \lambda(it\gamma - |t|^\alpha + it\omega_A(t, \alpha, \beta)).$$

Donde

$$\omega_A(t, \alpha, \beta) := \begin{cases} |t|^{\alpha-1} \beta \tan(\pi\alpha/2) & \alpha \neq 1, \\ -\beta(2/\pi) \log |t| & \alpha = 1. \end{cases} \quad (3.5)$$

La distribución con respecto a esta parametrización satisface

$$Y(\alpha, \beta, \gamma, \lambda) \sim \lambda^{\frac{1}{\alpha}} Y(\alpha, \beta, 0, 1) + \lambda(\gamma + b_0), \text{ para alguna } b_0.$$

Esto justifica interpretar a λ como un parámetro de escalamiento y a γ un parámetro de no centralidad.

3.2.2. Parametrización M

Teorema 48. *Una distribución no-degenerada es α -estable con vector de parámetros*

$$(\alpha, \beta, \gamma, \lambda) \in (0, 2] \times [-1, 1] \times (-\infty, +\infty) \times (0, +\infty),$$

si y sólo si el logaritmo de su función característica \mathbf{g} , puede ser representado como

$$\log \mathbf{g}(t) = \lambda(it\gamma - |t|^\alpha + it\omega_M(t, \alpha, \beta)).$$

Donde

$$\omega_M(t, \alpha, \beta) := \begin{cases} (|t|^{\alpha-1} - 1)\beta \tan(\pi\alpha/2) & \alpha \neq 1, \\ -\beta(2/\pi) \log |t| & \alpha = 1. \end{cases} \quad (3.6)$$

3.2.3. Parametrización B

Recordemos la definición de la función signo:

$$\text{sgn}(x) = 1_{[0, +\infty)} - 1_{(-\infty, 0)}(x).$$

Utilizaremos la siguiente notación

$$K(\alpha) := \alpha - 1 - \text{sgn}(1 - \alpha). \quad (3.7)$$

Teorema 49. *Una distribución no-degenerada es α -estable con vector de parámetros*

$$(\alpha, \beta, \gamma, \lambda) \in (0, 2] \times [-1, 1] \times (-\infty, +\infty) \times (0, +\infty),$$

si y sólo si el logaritmo de su función característica \mathbf{g} , puede ser representado como

$$\log \mathbf{g}(t) = \lambda(it\gamma - |t|^\alpha \omega_B(t, \alpha, \beta)).$$

Donde

$$\omega_B(t, \alpha, \beta) := \begin{cases} \exp(-i\frac{\pi}{2}\beta K(\alpha) \text{sgn}(t)) & \alpha \neq 1, \\ \frac{\pi}{2} + i\beta \log(|t|) \text{sgn}(t) & \alpha = 1. \end{cases} \quad (3.8)$$

3.2.4. Parametrización C

La parametrización de esta sección se satisface para las distribuciones en la clase \mathfrak{M} . Utilizaremos la notación

$$\theta_\alpha := \min\left(1, \frac{2}{\alpha} - 1\right).$$

Teorema 50. *Una distribución no-degenerada, es estrictamente α -estable con vector de parámetros*

$$(\alpha, \theta, \lambda) \in (0, 2] \times (-\infty, \theta_\alpha] \times (0, +\infty),$$

si y sólo si el logaritmo de su función característica \mathfrak{g} , puede ser representado como

$$\log \mathfrak{g}(t) = -\lambda |t|^\alpha \omega_C(t, \alpha, \theta).$$

Donde

$$\omega_C(t, \alpha, \theta) := \exp(-i\theta\alpha\frac{\pi}{2}\text{sgn}(t)). \quad (3.9)$$

3.2.5. Parametrización E

La parametrización E, descrita en el siguiente teorema, será crucial para la inferencia estadística de distribuciones estrictamente estables por el método de momentos.

Teorema 51. *Una distribución no-degenerada, es estrictamente α -estable con vector de parámetros*

$$(\nu, \theta, \tau) \in \left(\frac{1}{4}, +\infty\right) \times (-\infty, \theta_\alpha] \times (-\infty, +\infty),$$

si y sólo si el logaritmo de su función característica \mathfrak{g} , puede ser representado como

$$\log \mathfrak{g}(t) = -\exp(\nu^{-\frac{1}{2}}(\log(|t|) + \tau - i\frac{\pi}{2}\theta\text{sgn}(t)) + \mathfrak{E}(\nu^{-\frac{1}{2}} - 1)).$$

En la que \mathfrak{E} es la constante de Euler, definida por

$$\mathfrak{E} := \lim_{n \rightarrow \infty} \left\{ \log(n) - \sum_{j=1}^n \frac{1}{j} \right\}.$$

3.3. Propiedad de invarianza

En el siguiente resultado, recordamos una propiedad de invarianza de la distribución α -estable, que será importante en la Sección 3.4; ver e.g., Zolotarev [29], pp. 60-61.

Proposición 52. *Sea X una variable aleatoria con distribución α -estable de parámetros $(\alpha, \beta, \gamma, \lambda)$. La variable aleatoria $-X$ tiene distribución α -estable de parámetros $(\alpha, -\beta, -\gamma, \lambda)$.*

3.4. Representación integral de la función de densidad de una ley α -estable

En esta sección utilizaremos la parametrización de la función característica en la forma B, que presentamos en la subsección 3.2.3. Fijaremos valores de la siguiente forma

Hipótesis 53. *Las distribuciones estándares en la forma B, es decir, $\gamma = 0$ y $\lambda = 1$. Los demás parámetros satisfacen $1 < \alpha < 2$, $-1 < \beta < 1$. Consideraremos, sin pérdida de generalidad, valores no negativos $x > 0$.*

Utilizaremos la siguiente

Notación 54.

1. $s(\alpha) = \operatorname{sgn}(1 - \alpha)$

2. $\tau = \beta \frac{K(\alpha)}{\alpha}$

3. $\tau^* = \tau \operatorname{sgn}(x)$

- 4.

$$U_\alpha(\phi, \tau) := \left(\frac{\sin(\frac{\pi\alpha(\phi+\tau)}{2})}{\cos(\frac{\pi\phi}{2})} \right)^{\frac{\alpha}{1-\alpha}} \frac{\cos(\frac{\pi\phi(\alpha-1)+\pi\alpha\tau}{2})}{\cos(\frac{\pi\phi}{2})}.$$

Observación 55. *Para $\alpha > 1$ tenemos*

$$s(\alpha) = -1, K(\alpha) = \alpha \text{ y } \tau = \beta.$$

Además, como $x > 0$, tenemos que $\tau^ = \tau = \beta$.*

Debido a que los valores de los parámetros γ y λ son fijos, utilizaremos la notación $g(x, \alpha, \beta)$, para la función de densidad de la distribución α -estable. Ella será identificada mediante la función característica:

$$g(x, \alpha, \beta) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \mathbf{g}(t, \alpha, \beta) dt \quad (3.10)$$

$$= \frac{1}{\pi} \int_0^{+\infty} e^{itx} \mathbf{g}(t, \alpha, -\beta) dt, \quad (3.11)$$

debido al teorema de inversión de Lévy. Partiendo de esta caracterización, nuestro objetivo es demostrar el siguiente teorema de representación integral de la distribución estable.

Teorema 56. *Sea $\alpha > 1$ y β tal que $|\beta| < 1$. Para $x \neq 0$ la función de densidad de la distribución estándar α -estable tiene la representación integral*

$$g(x, \alpha, \beta) = \frac{\alpha |x|^{\frac{1}{\alpha-1}}}{2|1-\alpha|} \int_{-\tau^*}^1 U_\alpha(\phi, \tau^*) \exp\{-|x|^{\frac{\alpha}{\alpha-1}} U_\alpha(\phi, \tau^*)\} d\phi. \quad (3.12)$$

El Teorema 56 es un caso especial del Teorema 2.2.3 de Zolotarev [29], el cual abarca otros valores de los parámetros (α, β) .

Es suficiente con demostrar el Teorema 56 para el caso $x > 0$, ya que la función de densidad g satisface:

$$g(-x, \alpha, \beta) = g(x, \alpha, -\beta), \quad (3.13)$$

debido a la Proposición 52.

3.4.1. Resultados preparatorios

La demostración del Teorema 56 requerirá varios resultados preparatorios. La integración compleja será la técnica principal que utilizaremos para demostrar la representación integral. En la siguiente observación hacemos un rápido recordatorio de la definición de la integral compleja.

Observación 57. *Recordatorio de la integral compleja. Una función $f : \mathbb{C} \rightarrow \mathbb{C}$ tiene la forma*

$$f(a + ib) = u(a + ib) + iv(a + ib).$$

Sea I un intervalo de la forma $[a, b]$. La integral compleja a lo largo de una curva γ , con parametrización $G : I \rightarrow \mathbb{C}$, se define componente por

componente, respetando la multiplicación compleja. Supongamos una parametrización de la forma

$$G(t) = A(t) + iB(t).$$

La derivada es entonces

$$DG(t) = A'(t) + iB'(t).$$

Utilizaremos en la siguiente construcción la siguiente convención: El símbolo \circ denota composición de funciones, \times denota multiplicación compleja y \cdot denota multiplicación real. La integral compleja se construye de la siguiente forma

$$\begin{aligned} \int_{\gamma} f(z)dz &:= \int_I f \circ G(t) \times DG(t)dt \\ &= \int_I [u \circ G(t) + iv \circ G(t)] \times [A'(t) + iB'(t)] dt \\ &= \int_I [u \circ G(t) \cdot A'(t) - v \circ G(t) \cdot B'(t)] dt \\ &\quad + i \int_I [u \circ G(t) \cdot B'(t) + v \circ G(t) \cdot A'(t)] dt. \end{aligned} \quad (3.14)$$

Sea C el semirayo en el plano complejo definido por la condición

$$\text{Arg}(z) = -\frac{3}{4}\pi.$$

Denotaremos mediante \mathfrak{g}^+ la extensión analítica de \mathfrak{g} , del intervalo $(0, \infty)$, al dominio $\mathbb{C} - C$. El logaritmo de la función característica \mathfrak{g}^+ será importante y utilizaremos la siguiente notación

$$\Psi(z, \alpha, \beta) := \log(\mathfrak{g}^+(z, \alpha, \beta)). \quad (3.15)$$

Recuerde que τ depende de α y β . Utilizaremos la siguiente versión de Ψ correspondiendo a la parametrización B

$$\Psi(z, \alpha, \beta) = -z^\alpha \exp \left\{ -i \frac{\pi}{2} \tau \alpha \right\}. \quad (3.16)$$

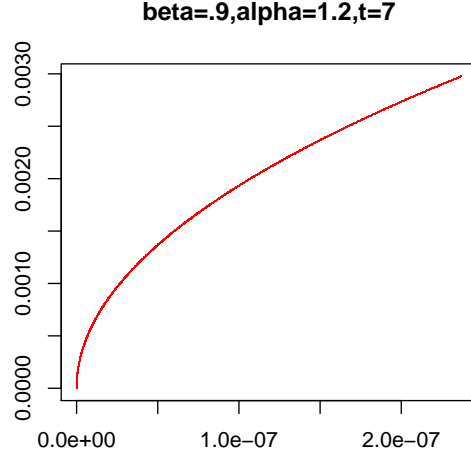
Sea N la función definida de la siguiente forma

$$N(z) := izx + \Psi(z, \alpha, -\beta). \quad (3.17)$$

Sea Γ la curva definida como

$$\Gamma := \left\{ z \mid \Im(N(z)) = 0, \text{arg}(z) \in \left[-\tau \frac{\pi}{2}, \frac{\pi}{2} \right] \right\}. \quad (3.18)$$

Una ilustración de la curva Γ se puede ver en la figura 3.1.

Figura 3.1: Curva Γ .

Proposición 58. *La función N satisface a lo largo de la curva Γ la siguiente identidad integral*

$$\Re \int_{\Gamma} \exp\{N(z)\} dz = \frac{\alpha s(\alpha)}{1-\alpha} \int_{-\tau^* \frac{\pi}{2}}^{\frac{\pi}{2}} x^{\frac{1}{\alpha-1}} U_{\alpha} \left(\frac{2\phi}{\pi}, \tau \right) \exp \left\{ -x^{\frac{\alpha}{\alpha-1}} U_{\alpha} \left(\frac{2\phi}{\pi}, \tau \right) \right\} d\phi. \quad (3.19)$$

Demostración. *Para estimar la integral compleja (3.19), es necesario parametrizar la curva Γ . Para este fin utilizaremos las coordenadas polares:*

$$z = re^{i\theta} = r(\cos(\theta) + i \sin(\theta)).$$

Sea I el intervalo $[-\tau^* \frac{\pi}{2}, \frac{\pi}{2}]$. Sea $G : I \rightarrow \mathbb{C}$ la función definida en coordenadas polares por

$$r(\theta) := \left(\frac{\sin(\alpha\theta - \frac{\pi}{2}\tau\alpha)}{x \cos(\theta)} \right)^{\frac{1}{1-\alpha}}. \quad (3.20)$$

Más abajo demostraremos que G es una parametrización de la curva Γ . En

concordancia con la notación de la Observación 57, denotemos

$$\begin{aligned} A &:= \Re(G), \\ B &:= \Im(G), \\ u &:= \Re(N), \\ v &:= \Im(N). \end{aligned}$$

Observemos que para cada complejo $z \in \Gamma$ se satisface

$$v(z) = 0,$$

debido a la definición de Γ . Como consecuencia, tenemos

$$\int_{\Gamma} \exp\{N(z)\} dz = \int_{\Gamma} \exp\{u(z)\} dz.$$

Comparando con la fórmula (3.14) de la definición de integral compleja, obtenemos

$$\Re \int_{\Gamma} \exp\{N(z)\} dz = \int_I \exp\{u \circ G(\theta)\} A'(\theta) d\theta. \quad (3.21)$$

La demostración ahora consiste en escribir explícitamente el integrando $u \circ G(\theta)$, así como mostrar la siguiente relación, para el integrador a lo largo de la curva Γ :

$$xA'(\theta) = \frac{\alpha}{1-\alpha} u(r(\theta), \theta). \quad (3.22)$$

1. En coordenadas polares, los puntos de la curva Γ satisfacen

$$xr \cos(\theta) - r^\alpha \sin\left(\alpha\theta - \frac{\pi}{2}\theta\alpha\right) = 0,$$

debido a la ecuación (3.23), abajo. Esta relación implica la parametrización de la curva (3.20).

2. En coordenadas polares la función N tiene la forma

$$ixr(\cos(\theta) + i\sin(\theta)) - r^\alpha (\cos(\alpha\theta) + i\sin(\alpha\theta)) \left(\cos\left(-\frac{\pi}{2}\theta\alpha\right) + i\sin\left(-\frac{\pi}{2}\theta\alpha\right) \right).$$

Al realizar los productos involucrados en la anterior expresión, vemos que la parte imaginaria tiene la forma

$$xr \cos(\theta) - r^\alpha \cos(\alpha\theta) \sin\left(-\frac{\pi}{2}\theta\alpha\right) - r^\alpha \sin(\alpha\theta) \cos\left(-\frac{\pi}{2}\theta\alpha\right).$$

La anterior expresión se puede reducir por suma de ángulos a

$$\Im(N) = xr \cos(\theta) - r^\alpha \sin\left(\alpha\theta - \frac{\pi}{2}\theta\alpha\right). \quad (3.23)$$

En forma análoga se calcula la parte real de la función N y en particular tenemos la forma de la función u en coordenadas polares:

$$u(r, \theta) = \Re(N(r, \theta)) = xr \sin(\theta) + r^\alpha \cos\left(\alpha\theta + \frac{\pi}{2}\tau\alpha\right). \quad (3.24)$$

3. Una propiedad fundamental del componente u de la función N , es su relación con la función U_α a lo largo de la curva Γ :

$$u(\theta, r) = x^{\frac{\alpha}{\alpha-1}} U_\alpha\left(\frac{2\theta}{\pi}, \tau\right).$$

La relación se obtiene al sustituir la parametrización (3.20) en la definición de la función u en coordenadas polares (3.24).

4. Ahora mostramos la relación (3.22). Debemos calcular

$$\frac{d}{d\theta} (r(\theta) \cos(\theta)) = r'(\theta) \cos(\theta) - r(\theta) \sin(\theta). \quad (3.25)$$

Para simplificar los cálculos introducimos la notación

$$T := \alpha\theta + \frac{\pi}{2}\tau^*\alpha.$$

La derivada de $r(\theta)$ es

$$r'(\theta) = \frac{1}{1-\alpha} \left(\frac{\sin(T)}{x \cos(\theta)} \right)^{\frac{\alpha}{1-\alpha}} \frac{\cos(T)\alpha x \cos(\theta) + x \sin(\theta) \sin(T)}{(x \cos(\theta))^2}. \quad (3.26)$$

Al sustituir (3.26) en (3.25) y reducir,

$$\frac{d}{d\theta} (r(\theta) \cos(\theta)) = \frac{\alpha r^\alpha(\theta)}{x(1-\alpha)} \cos(T) + \frac{r^\alpha(\theta)}{1-\alpha} \frac{\sin(\theta) \sin(T)}{x \cos(\theta)} - r(\theta) \sin(\theta).$$

La función r satisface $r^\alpha \sin(T) = xr \cos(\theta)$ y entonces tenemos la igualdad

$$\frac{r^\alpha \sin(\theta) \sin(T)}{x \cos(\theta)} = r \sin(\theta).$$

Como consecuencia tenemos

$$\begin{aligned} \frac{d}{d\theta} (r(\theta) \cos(\theta)) &= \frac{\alpha r^\alpha(\theta)}{x(1-\alpha)} \cos(T) + \frac{\alpha}{1-\alpha} r(\theta) \sin(\theta) \\ &= \frac{\alpha}{x(1-\alpha)} [r^\alpha(\theta) \cos(T) + xr(\theta) \sin(\theta)]. \end{aligned}$$

□

El siguiente es el Lema 2.2.2 de Zolotarev [29] especializado al caso $\alpha > 1$.

Lema 59. *Sea $\epsilon \in (0, \pi/4)$ un número fijo. Definamos la curva*

$$C_r := \left\{ z \mid \text{mod}(z) = r, \epsilon - \frac{\pi}{2\alpha}(1 - \beta K(\alpha)) \leq \arg(z) \leq \frac{\pi}{2\alpha}(1 + \beta K(\alpha)) \right\}.$$

Entonces se satisface la siguiente propiedad asintótica

$$\lim_{r \rightarrow +\infty} \int_{C_r} e^{izx} \mathbf{g}^+(z, \alpha, -\beta) dz = 0.$$

Demostración. *Definamos*

$$Q(C_r) := \int_{C_r} e^{izx} \mathbf{g}^+(z, \alpha, -\beta) dz.$$

La función N se definió en la ecuación (3.17) como

$$N(z) := izx + \Psi(z, \alpha, -\beta).$$

Observe que $Q(C_r) = \int_{C_r} e^{N(z)} dz$. Tenemos

$$\begin{aligned} Q(C_r) &\leq l(C_r) \cdot \max_{z \in C_r} \|\exp\{N(z)\}\|_{\mathbb{C}} \\ &\leq l(C_r) \cdot \exp \left\{ \max_{z \in C_r} \Re N(z) \right\}, \end{aligned}$$

debido a la definición de integral compleja; ver e.g., Stein y Shakarchi [26], Proposición 1.3.1, parte iii).

Observe que $l(C_r) \leq \pi r$. Vamos a verificar la siguiente cota superior para la parte real de la función N :

$$U(z) := \Re(N(z)) \leq -2 \log r. \quad (3.27)$$

Como consecuencia tendremos el enunciado del lema:

$$Q(C_r) \leq \pi r \exp\{-2 \log r\} \leq \frac{\pi}{r}.$$

Ahora establecemos la estimación (3.27).

1. En coordenadas polares, la función U es de la forma:

$$U(r, \theta) = -xr \sin(\theta) - r^\alpha \cos\left(\alpha\theta + \frac{\pi}{2}\beta\alpha\right). \quad (3.28)$$

Para $\theta \in \left[\epsilon - \frac{\pi}{2\alpha}(1 - \beta K(\alpha)), \frac{\pi}{2\alpha}(1 + \beta K(\alpha))\right]$ existe una constante $c > 0$ tal que

$$\cos\left(\alpha\theta + \frac{\pi}{2}\beta\alpha\right) \geq c,$$

debido a que la función coseno es no negativa en el intervalo $[-\frac{\pi}{2}, \frac{\pi}{2}]$ y se anula exclusivamente en los extremos. Como consecuencia tenemos la desigualdad

$$U(r, \theta) \leq -xr \sin(\theta) - cr^\alpha.$$

2. Se satisface

$$\lim_{r \rightarrow \infty} \frac{r^\alpha}{\log(r)} = +\infty,$$

como consecuencia a una aplicación de la regla de L'Hopital. En particular se satisface la siguiente propiedad: para toda constante $K > 0$ existe $r_0 > 0$ tal que

$$r^\alpha \geq K \log(r), r > r_0.$$

3. Para $\alpha > \alpha' > 1$ y $b, M > 0$, claramente se satisface

$$\lim_{r \rightarrow \infty} \frac{br^\alpha - Mr}{r^{\alpha'}} = +\infty,$$

como consecuencia a una aplicación de la regla de L'Hopital. En particular se satisface la siguiente propiedad: existe $r_0 > 0$ tal que

$$br^\alpha - Mr \geq r^{\alpha'}.$$

□

El siguiente es el Lema 2.2.3 de Zolotarev [29]. Recuerde que denotamos mediante C , el semirayo en el plano complejo negativo caracterizado por la condición

$$\arg(z) = -\frac{3}{4}\pi.$$

Lema 60. Sea Γ una curva en el plano complejo que satisface las siguientes condiciones:

1. Γ comienza en cero.
2. El contorno no interseca el corte C .
3. La curva converge a infinito en norma y los puntos z satisfacen

$$\epsilon - \frac{\pi}{2\alpha}(1 - \beta K(\alpha)) \leq \arg(z) \leq \frac{\pi}{2\alpha}(1 + \beta K(\alpha)).$$

Entonces

$$\int_0^{+\infty} e^{izx} \mathfrak{g}^+(z, \alpha, -\beta) dz = \int_{\Gamma} e^{izx} \mathfrak{g}^+(z, \alpha, -\beta) dz. \quad (3.29)$$

Demostración. Siguiendo la demostración de Zolotarev [29], pp.70, definimos las siguientes curvas

1. T_r es la porción de la curva Γ comenzando en la intersección z_r con la circunferencia $\text{Mod}(z) = r$ y continuando hasta el origen. Estamos considerando una parametrización en el sentido de las manecillas del reloj.
2. L_r es la porción de la circunferencia $\text{Mod}(z) = r$, comenzando en el punto $z = z_r$ y continuando hasta $z = r + 0i$.
3. Para $R > r$, $-L_R$ es la porción de la circunferencia $\text{Mod}(z) = R$, comenzando en el punto $z = R + 0i$ y continuando hasta el punto de intersección $z = z_R$.
4. $I_{R,r}$ es el intervalo $[r, R]$ considerado como una curva del plano complejo.

Definimos un conjunto $M_{r,R}$ definido como la siguiente unión de puntos

$$I_{r,R} \cup -L_R \cup T_r \cup L_r.$$

Este conjunto define una curva en el plano complejo que claramente se puede parametrizar. Tenemos en este caso

$$\int_{M_{R,r}} N(z) dz = 0,$$

debido al Teorema de Cauchy, ya que la función N es analítica en el dominio $\mathbb{C} - C$.

Se satisface

$$\begin{aligned} \lim_{r \rightarrow 0, R \rightarrow \infty} \int_{M_{R,r}} N(z) dz &= \lim_{r \rightarrow 0, R \rightarrow \infty} \int_{-L_R \cup [r, R]} N(z) dz \\ &= - \int_{\Gamma} N(z) dz + \int_0^{+\infty} N(z) dz, \end{aligned}$$

debido al Lema 59. □

3.4.2. Demostración del Teorema 56

Tenemos lo siguiente

$$\begin{aligned} g(x, \alpha, \beta) &= \frac{1}{\pi} \int_0^{+\infty} e^{itx} \mathfrak{g}(t, \alpha, -\beta) dt, \\ &= \frac{1}{\pi} \int_0^{+\infty} e^{itx} \mathfrak{g}^+(t, \alpha, -\beta) dt, \\ &= \frac{1}{\pi} \Re \int_0^{+\infty} e^{izx} \mathfrak{g}^+(z, \alpha, -\beta) dz, \\ &= \frac{1}{\pi} \Re \int_0^{+\infty} \exp\{N(z)\} dz. \end{aligned}$$

Como consecuencia tenemos que

$$g(x, \alpha, \beta) = \frac{1}{\pi} \Re \int_{\Gamma} e^{izx} \mathfrak{g}^+(z, \alpha, -\beta) dz,$$

debido a (3.29). Mas aun,

$$g(x, \alpha, \beta) = \frac{1}{\pi} \frac{\alpha s(\alpha)}{1 - \alpha} \int_{-\tau^* \frac{\pi}{2}}^{\frac{\pi}{2}} x^{\frac{1}{\alpha-1}} U_{\alpha} \left(\frac{2\phi}{\pi}, \tau \right) \exp \left\{ -x^{\frac{\alpha}{\alpha-1}} U_{\alpha} \left(\frac{2\phi}{\pi}, \tau \right) \right\} d\phi,$$

debido a la ecuación (3.19) de la Proposición 58. La demostración del Teorema concluye con el cambio de variable

$$\phi = \phi' \frac{\pi}{2}.$$

□

Capítulo 4

CVaR para leyes α -estables

Debido a la gran popularidad de la medida de riesgo VaR, existe una amplia variedad de implementaciones para ella. Desde los modelos clásicos de distribuciones gaussianas que se referencian en el acuerdo de convergencia monetaria de Basilea [22], hasta modelos más sofisticados, basados en procesos de volatilidad estocástica, ver e.g., Engle y Manganelli [8]. La estimación de riesgo cuantificada por CVaR es una línea de investigación en progreso. En este capítulo supondremos que las ganancias y pérdidas se distribuyen de acuerdo a una ley α -estable y exponemos una representación integral para el cálculo del CVaR debida a Stoyanov et al. [27].

Nuestro objetivo es demostrar las formulas de representación (4.4) y (4.5) más abajo, del calculo del CVaR para distribuciones estables. En este objetivo, el Teorema 56, visto en el Capítulo 3 será crucial.

Introduciremos varias notaciones. Existen intersecciones con la Notación 54. No obstante, utilizaremos la notación de Stoyanov et al. [27].

1. $\bar{\theta}_0 = \frac{1}{\alpha} \arctan \left(\bar{\beta} \tan \frac{\pi \cdot \alpha}{2} \right),$

2. $\bar{\beta} = -\text{sgn}(VaR_\varepsilon(X)) \beta,$

- 3.

$$\theta_0 := \theta_0(\beta) := \frac{1}{\alpha} \arctan \left(\beta \tan \frac{\pi \cdot \alpha}{2} \right), \quad (4.1)$$

4.

$$g(\theta) := \frac{\sin(\alpha(\bar{\theta}_0 + \theta) - 2\theta)}{\sin \alpha(\bar{\theta}_0 + \theta)} - \frac{\alpha \cos^2 \theta}{\sin^2 \alpha(\bar{\theta}_0 + \theta)}, \quad (4.2)$$

$$v(\theta) := (\cos \alpha \bar{\theta}_0)^{\frac{1}{\alpha-1}} \left(\frac{\cos \theta}{\sin \alpha(\bar{\theta}_0 + \theta)} \right)^{\frac{\alpha}{\alpha-1}} \frac{\cos(\alpha \bar{\theta}_0 + (\alpha-1)\theta)}{\cos \theta}. \quad (4.3)$$

A lo largo de este capítulo, denotaremos por $S_\alpha(\lambda, \beta, \gamma)$ al conjunto de leyes α -estables con parametrización M, vista en 3.2. Se presenta entonces el resultado principal de este capítulo.

Teorema 61. *Sea $X \in S_\alpha(1, \beta, 0)$ con $\alpha > 1$. Si $VaR_\varepsilon(X) \neq 0$, entonces el VaR condicional de X con nivel de significancia ε admite la siguiente representación integral:*

$$CVaR_\varepsilon(X) = \frac{\alpha}{1-\alpha} \frac{|VaR_\varepsilon(X)|}{\pi \cdot \varepsilon} \int_{-\bar{\theta}_0}^{\pi/2} g(\theta) \exp\left(-|VaR_\varepsilon(X)|^{\frac{\alpha}{\alpha-1}} v(\theta)\right) d\theta, \quad (4.4)$$

Además, si $VaR_\varepsilon(X) = 0$, entonces

$$CVaR_\varepsilon(X) = \frac{2}{(\pi - 2\theta_0)} \Gamma\left(\frac{\alpha-1}{\alpha}\right) \frac{\cos \theta_0}{(\cos \alpha \theta_0)^{\frac{1}{\alpha}}}. \quad (4.5)$$

4.1. Lemas preparativos

En esta sección fijamos una variable aleatoria X con función de distribución α -estable $X \in S_\alpha(1, \beta, 0)$. La función de densidad la denotaremos mediante $f_X(u; \alpha, \beta)$. El primer resultado preparativo para la demostración del Teorema 61 es el siguiente lema.

Lema 62.

$$CVaR_\varepsilon(X) = \frac{1}{\varepsilon} \int_0^\infty u \cdot f_X(u; \alpha, -\beta) du - \frac{1}{\varepsilon} \int_0^{|VaR_\varepsilon(X)|} u \cdot f_X(u; \alpha, -\text{sgn}(VaR_\varepsilon(X))\beta) du \quad (4.6)$$

Demostración. *De la definición de VaR y de esperanza condicional se obtiene lo siguiente*

$$\begin{aligned} CVaR_\varepsilon(X) &= \frac{\mathbb{E}\left[-X \mathbf{1}_{\{-X \geq VaR_\varepsilon(X)\}}\right]}{\mathbb{P}(X \leq -VaR_\varepsilon(X))} \\ &= \frac{1}{\varepsilon} \int_{VaR_\varepsilon(X)}^\infty u \cdot f_X(-u; \alpha, \beta) du. \end{aligned} \quad (4.7)$$

Es claro que cuando $VaR_\varepsilon(X) \geq 0$, la expresión (4.7) se puede escribir como

$$CVaR_\varepsilon(X) = \frac{1}{\varepsilon} \int_0^\infty u \cdot f_X(u; \alpha, -\beta) du - \frac{1}{\varepsilon} \int_0^{VaR_\varepsilon(X)} u \cdot f_X(u; \alpha, -\beta) du,$$

debido a la propiedad de invarianza para leyes estables

$$f_X(-u; \alpha, \beta) = f_X(u; \alpha, -\beta),$$

de la Proposición 52.

Ahora bien, si $VaR_\varepsilon(X) < 0$,

$$\begin{aligned} CVaR_\varepsilon(X) &= \\ &= \frac{1}{\varepsilon} \int_0^\infty u \cdot f_X(u; \alpha, -\beta) du + \frac{1}{\varepsilon} \int_{VaR_\varepsilon(X)}^0 u \cdot f_X(-u; \alpha, \beta) du \\ &= \frac{1}{\varepsilon} \int_0^\infty u \cdot f_X(u; \alpha, -\beta) du - \frac{1}{\varepsilon} \int_0^{-VaR_\varepsilon(X)} u \cdot f_X(u; \alpha, \beta) du, \end{aligned}$$

con lo cual se obtiene el resultado. \square

Lema 63. La función definida por

$$L(\theta; \alpha, \beta) = \frac{\alpha}{\alpha - 1} \cos \theta \left(\frac{\cos(\alpha\theta_0 + (\alpha - 1)\theta)}{\cos \alpha\theta_0 \cos \theta} \right)^{\frac{1}{\alpha}}, \quad (4.8)$$

es una antiderivada de

$$(v(\theta; \alpha, \beta))^{\frac{1-\alpha}{\alpha}} = (\cos \alpha\theta_0)^{-\frac{1}{\alpha}} \frac{\sin \alpha(\theta_0 + \theta)}{\cos \theta} \left(\frac{\cos(\alpha\theta_0 + (\alpha - 1)\theta)}{\cos \theta} \right)^{\frac{1-\alpha}{\alpha}}$$

y tiene las siguientes propiedades:

$$\begin{aligned} \lim_{\theta \rightarrow -\theta_0^+} L(\theta; \alpha, \beta) &= \frac{\alpha \cos \theta_0}{(1 - \alpha)(\cos \alpha\theta_0)^{\frac{1}{\alpha}}}, \\ \lim_{\theta \rightarrow \pi/2^-} L(\theta; \alpha, \beta) &= 0, \end{aligned}$$

donde $v(\theta; \alpha, \beta)$ y θ_0 fueron definidos en (4.2), (4.3) respectivamente.

Demostración. De la definición de $v(\theta; \alpha, \beta)$ y denotando $A = (\cos \alpha\theta_0)^{-\frac{1}{\alpha}}$,

$$\begin{aligned} v(\theta; \alpha, \beta)^{\frac{1-\alpha}{\alpha}} &= A \frac{\sin \alpha(\theta_0 + \theta)}{\cos \theta} \left(\frac{\cos(\alpha\theta_0 + (\alpha-1)\theta)}{\cos \theta} \right)^{\frac{1-\alpha}{\alpha}} \\ &= A \frac{\sin(\alpha(\theta_0 + \theta) + \theta - \theta)}{\cos \theta} \left(\frac{\cos(\alpha\theta_0 + (\alpha-1)\theta)}{\cos \theta} \right)^{\frac{1-\alpha}{\alpha}} \\ &= A \sin(\alpha(\theta_0 + \theta) - \theta) \left(\frac{\cos(\alpha\theta_0 + (\alpha-1)\theta)}{\cos \theta} \right)^{\frac{1-\alpha}{\alpha}} \\ &\quad + A \sin \theta \left(\frac{\cos(\alpha\theta_0 + (\alpha-1)\theta)}{\cos \theta} \right)^{\frac{1}{\alpha}}. \end{aligned}$$

Por otro lado,

$$L(\theta; \alpha, \beta) = A \frac{\alpha}{1-\alpha} [\cos(\alpha\theta_0 + (\alpha-1)\theta)]^{\frac{1}{\alpha}} (\cos \theta)^{\frac{\alpha-1}{\alpha}}, \quad (4.9)$$

de lo cual

$$\begin{aligned} L'(\theta; \alpha, \beta) &= A \frac{\alpha}{1-\alpha} \frac{\alpha-1}{\alpha} [\cos(\alpha\theta_0 + (\alpha-1)\theta)]^{\frac{1-\alpha}{\alpha}} [-\sin(\alpha\theta_0 + (\alpha-1)\theta)] (\cos \theta)^{\frac{\alpha-1}{\alpha}} \\ &\quad + A \frac{\alpha}{1-\alpha} [\cos(\alpha\theta_0 + (\alpha-1)\theta)]^{\frac{1}{\alpha}} \frac{\alpha-1}{\alpha} (\cos \theta)^{\frac{-1}{\alpha}} (-\sin \theta) \\ &= v(\theta; \alpha, \beta)^{\frac{1-\alpha}{\alpha}}. \end{aligned}$$

Nótese que las expresiones relacionadas con los límites se obtienen de manera sencilla de la ecuación (4.9). \square

Recuerde que definimos la función θ_0 en la ecuación (4.1).

Lema 64. La función $\theta_0 = \theta_0(\beta)$ es una función decreciente de β . Más aún, se cumple

$$-\frac{\pi}{2} < -\frac{\pi(2-\alpha)}{2\alpha} \leq \theta_0(\beta) \leq \frac{\pi(2-\alpha)}{2\alpha} < \frac{\pi}{2}, \quad (4.10)$$

donde $\beta \in [-1, 1]$. Además

$$0 \leq \sin \alpha(\theta_0 + \theta) \leq 1, \quad \text{para } \theta \in [-\theta_0, \pi/2] \quad (4.11)$$

y $\sin \alpha(\theta_0 + \theta) = 0$ si $\theta = \pi/2$ y $\beta = -1$.

Demostración. Notemos que

$$\tan \frac{\pi\alpha}{2} = \tan \left(\frac{\pi(\alpha - 2)}{2} \right) = -\tan \left(\frac{\pi(2 - \alpha)}{2} \right).$$

Entonces:

$$\theta_0(\beta) = -\frac{1}{\alpha} \arctan \left[\beta \tan \left(\frac{\pi(2 - \alpha)}{2} \right) \right].$$

De lo anterior,

$$\begin{aligned} \theta'_0(\beta) &= -\frac{1}{\alpha} \frac{\left[\beta \tan \left(\frac{\pi(2 - \alpha)}{2} \right) \right]'}{\beta^2 \tan^2 \left(\frac{\pi(2 - \alpha)}{2} \right) + 1} \\ &= -\frac{c_1}{c_2}, \quad \text{con } c_1, c_2 > 0, \end{aligned}$$

y entonces $\theta_0(\beta)$ es decreciente en el intervalo $[-1, 1]$, obteniendo

$$\theta_0(1) = -\frac{\pi(2 - \alpha)}{2\alpha} \leq \theta_0(\beta) \leq \frac{\pi(2 - \alpha)}{2\alpha} = \theta_0(-1).$$

Las desigualdades (4.10) son consecuencia de que $\alpha \in (1, 2)$. Para el caso en que $\theta \in [-\theta_0, \pi/2]$, con $\theta_0(\beta) < \pi/2$,

$$0 \leq \alpha(\theta_0 + \theta) \leq \alpha(\theta_0 + \pi/2) < \pi.$$

Por lo tanto (4.11) se cumple. Concluimos la igualdad $\sin \alpha(\theta_0 + \theta) = 0$ debido a la definición de θ_0 . \square

4.2. Demostración del Teorema de Representación 61

Consideremos en primer lugar el caso $VaR_\varepsilon(X) \neq 0$. De la expresión para la densidad de X dada en el Teorema 56 y la expresión (4.6) para el CVaR, podemos escribir a $CVaR_\varepsilon(X)$ como

$$CVaR_\varepsilon(X) = \frac{1}{\varepsilon} I_1 - \frac{1}{\varepsilon} I_2, \quad (4.12)$$

donde

$$I_1 = \int_0^\infty u \int_{\theta_0}^{\pi/2} \frac{\alpha u^{\frac{1}{\alpha-1}}}{\pi(\alpha-1)} v(\theta, \alpha, -\beta) \exp \left(-u^{\frac{\alpha}{\alpha-1}} v(\theta, \alpha, -\beta) \right) d\theta du,$$

$$I_2 = \int_0^{|VaR_\varepsilon(X)|} u \int_{-\bar{\theta}_0}^{\pi/2} \frac{\alpha u^{\frac{1}{\alpha-1}}}{\pi(\alpha-1)} v(\theta, \alpha, \bar{\beta}) \exp\left(-u^{\frac{1}{\alpha-1}} v(\theta, \alpha, \bar{\beta})\right) d\theta du.$$

Observemos que se escribe θ_0 en I_1 , ya que $\theta_0(-\beta) = -\theta_0$. Dado que $v(\theta, \alpha, \bar{\beta})$ es positiva, por el teorema de Tonelli,

$$I_2 = \frac{\alpha}{\pi(\alpha-1)} \int_{-\bar{\theta}_0}^{\pi/2} \int_0^{|VaR_\varepsilon(X)|} u^{\frac{1}{\alpha-1}} v(\theta, \alpha, \bar{\beta}) \exp\left(-u^{\frac{1}{\alpha-1}} v(\theta, \alpha, \bar{\beta})\right) d\theta du.$$

Ahora, con el cambio de variable $t = u^{\frac{1}{\alpha-1}} v(\theta, \alpha, \bar{\beta})$,

$$I_2 = \frac{\alpha}{\pi(\alpha-1)} \int_{-\bar{\theta}_0}^{\pi/2} \int_0^{a(\theta)} t^{\frac{\alpha-1}{\alpha}} u^{-\frac{1}{\alpha-1}} v(\theta; \alpha, \bar{\beta})^{-1} e^{-t} dt d\theta,$$

con $a(\theta) = |VaR_\varepsilon(X)|^{\frac{\alpha-1}{\alpha}} v(\theta; \alpha, \bar{\beta})$. Notemos que $tu^{-\frac{1}{\alpha-1}} v(\theta; \alpha, \bar{\beta})^{-1} = u$ y que $t^{\frac{\alpha-1}{\alpha}} = uv(\theta; \alpha, \bar{\beta})^{\frac{\alpha-1}{\alpha}}$, de lo cual

$$I_2 = \frac{1}{\pi} \int_{-\bar{\theta}_0}^{\pi/2} v(\theta; \alpha, \bar{\beta})^{\frac{1-\alpha}{\alpha}} \int_0^{a(\theta)} t^{\frac{\alpha-1}{\alpha}} e^{-t} dt d\theta.$$

Gracias al Lema 63,

$$I_2 = \frac{1}{\pi} \int_{-\bar{\theta}_0}^{\pi/2} \int_0^{a(\theta)} t^{\frac{\alpha-1}{\alpha}} e^{-t} dt dL(\theta; \alpha, \bar{\beta}).$$

Se procederá mediante integración por partes, para lo cual observemos la expresión

$$L(\theta; \alpha, \bar{\beta}) \left[\int_0^{a(\theta)} t^{\frac{\alpha-1}{\alpha}} e^{-t} dt \right] \Big|_{-\bar{\theta}_0}^{\pi/2}.$$

Del Lema 63, podemos notar que cuando θ tiende a $\pi/2$ por la izquierda, tanto la función $L(\theta; \alpha, \bar{\beta})$ como $a(\theta)$ tienden a cero. En consecuencia, solamente resta analizar el caso cuando θ tiende a $-\bar{\theta}_0$ por la derecha. De la definición de $a(\theta)$ y las propiedades de $v(\theta; \alpha, \bar{\beta})$ enunciadas en el Lema 63, segunda parte, tenemos

$$\lim_{\theta \rightarrow \bar{\theta}_0^+} a(\theta) = \lim_{\theta \rightarrow \bar{\theta}_0^+} v(\theta; \alpha, \bar{\beta}) = \infty.$$

Más aun

$$\lim_{\theta \rightarrow \bar{\theta}_0^+} L(\theta; \alpha, \bar{\beta}) = \frac{\alpha \cos \bar{\theta}_0}{(1-\alpha) \left(\cos \alpha \bar{\theta}_0\right)^{\frac{1}{\alpha}}}.$$

De lo anterior y el hecho de que $\Gamma(z + 1) = z\Gamma(z)$,

$$L(\theta; \alpha, \bar{\beta}) \left[\int_0^{a(\theta)} t^{\frac{\alpha-1}{\alpha}} e^{-t} dt \right] \Big|_{-\bar{\theta}_0}^{\pi/2} = \frac{\cos \bar{\theta}_0}{(\cos \alpha \bar{\theta}_0)^{\frac{1}{\alpha}}} \Gamma\left(\frac{\alpha-1}{\alpha}\right).$$

Además

$$\frac{d}{d\theta} \left(\int_0^{a(\theta)} t^{\frac{\alpha-1}{\alpha}} e^{-t} dt \right) = (a(\theta))^{\frac{\alpha-1}{\alpha}} e^{-a(\theta)} da(\theta).$$

Por lo tanto, como resultado de integrar por partes

$$I_2 = \frac{\cos \bar{\theta}_0}{\pi (\cos \alpha \bar{\theta}_0)^{\frac{1}{\alpha}}} \Gamma\left(\frac{\alpha-1}{\alpha}\right) + I_3, \quad (4.13)$$

donde I_3 está dada por

$$\begin{aligned} I_3 &:= -\frac{1}{\pi} \int_{-\bar{\theta}_0}^{\pi/2} L(\theta; \alpha, \bar{\beta}) (a(\theta))^{\frac{\alpha-1}{\alpha}} e^{-a(\theta)} da(\theta) \\ &= \frac{1}{\pi} \int_{-\bar{\theta}_0}^{\pi/2} L(\theta; \alpha, \bar{\beta}) (a(\theta))^{\frac{\alpha-1}{\alpha}} de^{-a(\theta)}. \end{aligned}$$

De nuevo se utilizará la integración por partes, así que centraremos nuestra atención en la expresión

$$\frac{1}{\pi} L(\theta; \alpha, \bar{\beta}) (a(\theta))^{\frac{\alpha-1}{\alpha}} e^{-a(\theta)} \Big|_{-\bar{\theta}_0}^{\pi/2}.$$

Como en el caso anterior, cuando θ tiende a $\pi/2$ por la izquierda, la expresión completa tiende a cero. Por otro lado de las propiedades de $v(\theta; \alpha, \bar{\beta})$,

$$\lim_{\theta \rightarrow -\bar{\theta}_0^+} a(\theta) e^{-a(\theta)} = 0. \quad (4.14)$$

Más aun

$$\begin{aligned} \lim_{\theta \rightarrow -\bar{\theta}_0^+} L(\theta; \alpha, \bar{\beta}) (a(\theta))^{-\frac{1}{\alpha}} &= |VaR_\varepsilon(X)|^{\frac{1}{1-\alpha}} \lim_{\theta \rightarrow -\bar{\theta}_0^+} L(\theta; \alpha, \bar{\beta}) v(\theta; \alpha, \bar{\beta})^{-\frac{1}{\alpha}} \\ &= \frac{\alpha}{1-\alpha} (\cos \alpha \bar{\theta}_0)^{-\frac{1}{\alpha}} |VaR_\varepsilon(X)|^{\frac{1}{1-\alpha}} \lim_{\theta \rightarrow -\bar{\theta}_0^+} \left[\frac{v(\theta; \alpha, \bar{\beta})}{\cos(\alpha \bar{\theta}_0 + (\alpha-1)\theta)} \right]^{-\frac{1}{\alpha}} (\cos \theta)^{\frac{\alpha-1}{\alpha}}. \end{aligned}$$

De la definición de $v(\theta; \alpha, \bar{\beta})$ tenemos

$$\left[\frac{v(\theta; \alpha, \bar{\beta})}{\cos(\alpha\bar{\theta}_0 + (\alpha-1)\theta)} \right]^{-\frac{1}{\alpha}} = \left[(\cos \alpha\bar{\theta}_0)^{\frac{1}{\alpha-1}} \right]^{-\frac{1}{\alpha}} \left[\frac{\sin(\alpha(\bar{\theta}_0 + \theta))}{\cos \theta} \right]^{\frac{1}{\alpha-1}} (\cos \theta)^{\frac{1}{\alpha}},$$

por lo cual y dado que $\bar{\theta}_0(\beta) \in (-\pi/2, \pi/2)$, con $\bar{\theta}_0$ está fijo en el límite,

$$\begin{aligned} \lim_{\theta \rightarrow -\bar{\theta}_0^+} L(\theta; \alpha, \bar{\beta}) (a(\theta))^{-\frac{1}{\alpha}} &= \\ &= \frac{\alpha}{1-\alpha} |VaR_\varepsilon(X)|^{\frac{1}{1-\alpha}} (\cos \alpha\bar{\theta}_0)^{\frac{1}{1-\alpha}} \lim_{\theta \rightarrow -\bar{\theta}_0^+} \cos \theta \left[\frac{\sin(\alpha(\bar{\theta}_0 + \theta))}{\cos \theta} \right]^{\frac{1}{\alpha-1}} (\cos \theta)^{\frac{1}{\alpha}} \\ &= 0. \end{aligned} \tag{4.15}$$

Por lo tanto, de (4.14) y (4.15) se tiene

$$\begin{aligned} \lim_{\theta \rightarrow -\bar{\theta}_0^+} L(\theta; \alpha, \bar{\beta}) (a(\theta))^{\frac{\alpha-1}{\alpha}} e^{-a(\theta)} &= \lim_{\theta \rightarrow -\bar{\theta}_0^+} L(\theta; \alpha, \bar{\beta}) (a(\theta))^{-\frac{1}{\alpha}} a(\theta) e^{-a(\theta)} \\ &= 0. \end{aligned}$$

Como consecuencia, nuestra expresión se hace cero en ambos casos. Al aplicar la integración por partes en I_3 nos resulta en lo siguiente

$$I_3 = -\frac{1}{\pi} \int_{-\bar{\theta}_0}^{\pi/2} e^{-a(\theta)} d \left[L(\theta; \alpha, \bar{\beta}) (a(\theta))^{\frac{\alpha-1}{\alpha}} \right].$$

Ahora observemos,

$$\begin{aligned} &L(\theta; \alpha, \bar{\beta}) (a(\theta))^{\frac{\alpha-1}{\alpha}} \\ &= \frac{\alpha}{1-\alpha} (\cos \alpha\bar{\theta}_0)^{-\frac{1}{\alpha}} \left[\cos(\alpha\bar{\theta}_0 + (\alpha-1)\theta) \right]^{\frac{1}{\alpha}} (\cos \theta)^{\frac{\alpha-1}{\alpha}} |VaR_\varepsilon(X)| v(\theta; \alpha, \bar{\beta})^{\frac{\alpha-1}{\alpha}} \\ &= \frac{\alpha}{1-\alpha} |VaR_\varepsilon(X)| \left[\frac{\cos(\alpha\bar{\theta}_0 + (\alpha-1)\theta)}{\cos \alpha\bar{\theta}_0} \right]^{\frac{1}{\alpha}} (\cos \theta)^{\frac{\alpha-1}{\alpha}} v(\theta; \alpha, \bar{\beta})^{\frac{\alpha-1}{\alpha}} \\ &= \frac{\alpha}{1-\alpha} |VaR_\varepsilon(X)| \cos(\alpha\bar{\theta}_0 + (\alpha-1)\theta) \frac{\cos \theta}{\sin(\alpha(\bar{\theta}_0 + \theta))} \\ &= \frac{\alpha}{1-\alpha} |VaR_\varepsilon(X)| \left[\cot(\alpha(\bar{\theta}_0 + \theta)) \cos^2 \theta + \sin \theta \cos \theta \right]. \end{aligned}$$

Además

$$\begin{aligned} \left[\cot \left(\alpha(\bar{\theta}_0 + \theta) \right) \cos^2 \theta \right]' &= -2 \cos \theta \sin \theta \cot \left(\alpha(\bar{\theta}_0 + \theta) \right) - \frac{\alpha \cos^2 \theta}{\sin^2 \left(\alpha(\bar{\theta}_0 + \theta) \right)} \\ &= -\sin 2\theta \cot \left(\alpha(\bar{\theta}_0 + \theta) \right) - \frac{\alpha \cos^2 \theta}{\sin^2 \left(\alpha(\bar{\theta}_0 + \theta) \right)} \end{aligned}$$

y

$$(\sin \theta \cos \theta)' = \cos^2 \theta - \sin^2 \theta = \frac{\cos 2\theta \sin \left(\alpha(\bar{\theta}_0 + \theta) \right)}{\sin \left(\alpha(\bar{\theta}_0 + \theta) \right)}.$$

Por lo tanto

$$\begin{aligned} \frac{d}{d\theta} \left[L(\theta; \alpha, \bar{\beta}) (a(\theta))^{\frac{\alpha-1}{\alpha}} \right] &= \\ &= \frac{\alpha}{1-\alpha} |VaR_\varepsilon(X)| \left[\frac{\sin \left(\alpha(\bar{\theta}_0 + \theta) - 2\theta \right)}{\sin \left(\alpha(\bar{\theta}_0 + \theta) \right)} - \frac{\alpha \cos^2 \theta}{\sin^2 \left(\alpha(\bar{\theta}_0 + \theta) \right)} \right] \\ &=: H(\theta). \end{aligned}$$

Gracias a la expresión que hemos hallado para I_3 y a (4.13),

$$I_2 = \frac{\cos \bar{\theta}_0}{\pi \left(\cos \alpha \bar{\theta}_0 \right)^{\frac{1}{\alpha}}} \Gamma \left(\frac{\alpha-1}{\alpha} \right) - \frac{1}{\pi} \int_{-\bar{\theta}_0}^{\pi/2} h(\theta) e^{-a(\theta)} d\theta.$$

Para I_1 , de manera análoga a I_2 se obtiene

$$\begin{aligned} I_1 &= \int_0^\infty u \int_{\theta_0}^{\pi/2} \frac{\alpha u^{\frac{1}{\alpha-1}}}{\pi(\alpha-1)} v(\theta, \alpha, -\beta) \exp \left(-u^{\frac{\alpha}{\alpha-1}} v(\theta, \alpha, -\beta) \right) d\theta du \\ &= \frac{\alpha}{\pi(\alpha-1)} \int_{\theta_0}^{\pi/2} \int_0^\infty u^{\frac{\alpha}{\alpha-1}} v(\theta, \alpha, -\beta) \exp \left(-u^{\frac{\alpha}{\alpha-1}} v(\theta, \alpha, -\beta) \right) dud\theta \\ &= \frac{1}{\pi} \int_{\theta_0}^{\pi/2} v(\theta, \alpha, -\beta)^{\frac{1-\alpha}{\alpha}} \int_0^\infty t^{\frac{\alpha-1}{\alpha}} e^{-t} dt d\theta \\ &= \Gamma \left(\frac{2\alpha-1}{\alpha} \right) \frac{1}{\pi} \int_{\theta_0}^{\pi/2} v(\theta, \alpha, -\beta)^{\frac{1-\alpha}{\alpha}} d\theta \\ &= \frac{1}{\pi} \Gamma \left(\frac{\alpha-1}{\alpha} \right) \frac{\cos \theta_0}{\left(\cos \alpha \theta_0 \right)^{\frac{1}{\alpha}}}. \end{aligned}$$

De la expresión anterior y la ecuación (4.12) se obtiene el resultado para $VaR_\varepsilon(X) \neq 0$. En el otro caso la integral I_2 se anula. En Nolan [21] se establece que

$$\varepsilon = \mathbb{P}(X < 0) = \frac{1}{\pi} \frac{\pi - 2\theta_0}{2},$$

de lo cual

$$CVaR_\varepsilon(X) = \frac{2}{\pi - 2\theta_0} \Gamma\left(\frac{\alpha - 1}{\alpha}\right) \frac{\cos \theta_0}{(\cos \alpha \theta_0)^{\frac{1}{\alpha}}},$$

como se quería probar. □

En este capítulo se ha obtenido una expresión integral cerrada para el CVaR, en el caso de contar con leyes α -estables. Una de las ventajas de la expresión obtenida, es que es fácilmente implementable computacionalmente hablando. De hecho, en el Apéndice se puede encontrar un programa en el que se realiza tal implementación. Cabe mencionar que el tiempo necesario para obtener todos los valores de CVaR, para las leyes $S_\alpha(1, \beta, 0)$, es mínimo. Una gráfica del conjunto de valores de CVaR, se muestra en la figura 4.1, obtenida por medio de nuestra implementación. El programa que la genera se incluye en el Apéndice A.

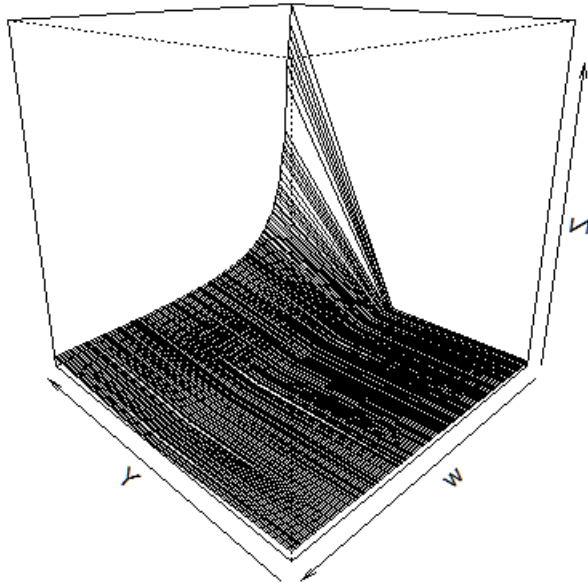


Figura 4.1: Contorno del rango de $CVaR_{0,1}(X)$.

Capítulo 5

Robustez cualitativa

De acuerdo a Huber [14], pp.1, la robustez en un procedimiento estadístico hace referencia a cierto grado de no-sensibilidad de la inferencia, ante pequeñas desviaciones en la distribución subyacente. Las desviaciones pueden ser de dos tipos: pequeñas perturbaciones a un número grande de datos muestrales, o bien grandes perturbaciones a un número reducido de observaciones. Nuestro objetivo en este capítulo, es hacer una exposición de la formalización matemática de esta idea intuitiva. Para este fin, será necesario precisar el contexto en que trabajaremos. Nuestros objetos básicos serán medidas de probabilidad en espacios métricos, y las *medidas empíricas* construidas a partir de una sucesión de observaciones independientes e idénticamente distribuidas, tendrán un lugar destacado. Las desviaciones serán entonces entre medidas de probabilidad en *espacios de observación*. Tales medidas serán cuantificadas por una métrica, prominentemente la *métrica de Prokhorov*.

Formalizamos un procedimiento estadístico a través de un estimador, el cual es una funcional definida en una clase de medidas de probabilidad, en el espacio de observaciones. La funcional estará valuada en un espacio métrico completo separable (espacio Polaco). La robustez del estimador es, entonces, una forma de continuidad con respecto a la métrica de Prokhorov. La intuición la obtenemos al considerar estimadores del tipo M (M-Estimators):

$$T(F) := \int \psi(x) dF(x).$$

Supongamos que F_n es una sucesión de distribuciones empíricas que convergen a la distribución F . Consideramos que F “está generando” las observaciones que producen las distribuciones empíricas. También suponemos que las observaciones sufren algún tipo de contaminación, que

se va reduciendo con n , y que produce las distribuciones empíricas contaminadas \tilde{F}_n . La siguiente convergencia es deseable:

$$\lim_{n \rightarrow \infty} T(\tilde{F}_n) = T(F).$$

Esta propiedad se verifica si ψ es acotada, continua y caracteriza la continuidad débil de la funcional T ; ver Huber [14], Lema 2.1, pp. 21. La distancia de Prokhorov ingresa al análisis, debido a que metriza la topología débil en el espacio de medidas de probabilidad borelianas, de un espacio polaco; ver Huber [14], Teorema 3.8, pp. 28.

Verificar la robustez de un estimador puede ser difícil. Además, contar con condiciones suficientes ágilmente verificables es importante. En esta dirección, el resultado más conocido es el célebre Teorema de Hampel [13], que reduce la robustez de un estimador a la continuidad puntual. Nosotros seguimos una exposición reciente de Krätschmer, Schied y Zähle [18]. En dicho trabajo, se define la robustez con respecto a otras métricas, en espacios de medidas de probabilidad diferentes a la métrica de Prokhorov. El contenido del capítulo está organizado de la siguiente forma. En la Sección 5.1 fijamos un espacio polaco y fijamos la notación para la clase de medidas empíricas definidas en el espacio. En la Sección 5.2 definimos un espacio métrico, en el cual introducimos una estructura probabilista que formaliza la noción de observaciones independientes, generadas por una distribución F . En la Sección 5.3, introducimos la distancia de Prokhorov y hacemos una exposición extensa del Teorema de Strassen, que caracteriza a la métrica de Prokhorov. En la Sección 5.4 presentamos la versión de Krätschmer et al. del Teorema de Hampel. En la Sección 5.5, dotamos de una estructura de espacio de probabilidad, a la clase de medidas empíricas formadas con n observaciones.

5.1. Medidas empíricas

Sea \mathfrak{X} un espacio polaco. Es decir, que es un espacio con una métrica $\mathfrak{p}_{\mathfrak{X}}(\cdot, \cdot)$, bajo la cual es completo y separable. Como ejemplo podemos pensar en \mathbb{R}^d . La métrica $\mathfrak{p}_{\mathfrak{X}}(\cdot, \cdot)$ en \mathfrak{X} induce una topología y la correspondiente σ -álgebra de borelianos la denotaremos $\mathcal{B}(\mathfrak{X})$. Para $x \in \mathfrak{X}$, la medida de probabilidad concentrada en este punto la denotaremos como δ_x . La medida de $A \in \mathcal{B}(\mathfrak{X})$ bajo δ_x , esta definida de la siguiente forma

$$\delta_x(A) := \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases} \quad (5.1)$$

Definición 65.

1. Denotaremos mediante $\mathfrak{F}_n(\mathfrak{X})$ el conjunto de medidas empíricas en \mathfrak{X}^n de la forma

$$\frac{1}{n} \sum_{i=1}^n \delta_{x_i}, \text{ para } (x_1, \dots, x_n) \in \mathfrak{X}^n.$$

2. El conjunto de todas las medidas empíricas en \mathfrak{X} lo denotaremos $\mathfrak{F}(\mathfrak{X})$. Es decir

$$\mathfrak{F}(\mathfrak{X}) = \bigcup_{n \in \mathbb{N}} \mathfrak{F}_n(\mathfrak{X}).$$

Definición 66. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Considere dada una sucesión fija de variables aleatorias independientes $\{X_i\}_{i \in \mathbb{N}}$ con

$$X_i : \Omega \rightarrow \mathfrak{X}.$$

Suponga que la sucesión es idénticamente distribuida, y bajo la medida de probabilidad \mathbb{P} , la ley de X_1 es F .

Para $n \in \mathbb{N}$, las primeras n observaciones X_1, \dots, X_n , inducen una medida aleatoria, valuada en el espacio de medidas de probabilidad \mathfrak{F}_n . Se trata de un mapeo $i_n : \Omega \rightarrow \mathfrak{F}_n$ definido como la distribución empírica de las observaciones en el escenario realizado. Es decir, para $\omega \in \Omega$ tenemos

$$i_n(\omega) := \frac{1}{n} \sum_{i=1}^n \delta_{X_i(\omega)}.$$

Intuitivamente pensamos que la medida empírica de observaciones independientes, aproximan a la función de distribución que “produce” las observaciones. Esta intuición se formaliza en el siguiente resultado.

Teorema 67. La medida empírica converge débilmente a su medida teórica:

$$\lim_{n \rightarrow \infty} i_n \stackrel{\text{ley}}{=} F, \quad \mathbb{P} - \text{c.s.}$$

Demostración. Ver Parthasarathy, [23], Teorema II.7.1. □

5.2. Espacio de observaciones

En esta sección introducimos un modelo probabilístico para formalizar la idea de sucesiones independientes de observaciones, que siguen la distribución F . Para este fin, consideraremos el producto cartesiano

$$\mathfrak{D} := \mathfrak{X}^{\mathbb{N}},$$

con la σ -álgebra del producto cartesiano $(\mathcal{B}(\mathfrak{X}))^{\mathbb{N}}$, la cual es la mínima σ -álgebra bajo la cual las proyecciones canónicas $\pi_i : \mathfrak{D} \rightarrow \mathfrak{X}$ son medibles; ver e.g., Parthasarathy, [23], Definición I.1.2.

Se puede inducir en forma natural una métrica en el espacio \mathfrak{D} , a partir de la métrica $\mathfrak{p}_{\mathfrak{X}}$, de la siguiente forma:

$$\mathfrak{p}_{\mathfrak{D}}(x, y) := \sum_{i=1}^{\infty} \frac{1}{2^i} \mathfrak{p}_{\mathfrak{X}}(x_i, y_i).$$

La σ -álgebra de conjuntos borelianos inducidos por esta métrica, coinciden con la σ -álgebra del producto cartesiano; ver e.g., [23], Teorema I.1.10. En el espacio medible $(\mathfrak{D}, \mathcal{B}(\mathfrak{D}))$ consideramos la medida de probabilidad

$$\mathbb{Q}_F := F^{\mathbb{N}}, \quad (5.2)$$

la cual es determinada por el Teorema de extensión de Kolmogorov, ver e.g. [23], Teorema V.5.1.

5.3. Métrica de Prokhorov

Recuerde que \mathfrak{X} es un espacio métrico completo separable, con métrica $\mathfrak{p}_{\mathfrak{X}}$. Denotemos mediante $\mathcal{B}(\mathfrak{X})$, a la σ -álgebra de los conjuntos Borelianos generada por la topología definida por la métrica $\mathfrak{p}_{\mathfrak{X}}$. Denotemos mediante $\mathcal{M}_1(\mathfrak{X})$, al conjunto de medidas de probabilidad definidas en $\mathcal{B}(\mathfrak{X})$.

Definición 68. 1. Sea $\epsilon > 0$. Para $A \in \mathcal{B}(\mathfrak{X})$ definimos

$$A^\epsilon := \{x \in \mathfrak{X} \mid \mathfrak{p}_{\mathfrak{X}}(a, x) \leq \epsilon \text{ para algun } a \in A\}.$$

2. Para $\mu, \nu \in \mathcal{M}_1(\mathfrak{X})$, su distancia de Prokhorov se define de la siguiente forma

$$\mathfrak{d}_{\text{prokh}}(\mu, \nu) := \inf\{\epsilon > 0 \mid \mu(A) \leq \nu(A^\epsilon) + \epsilon, A \in \mathcal{B}(\mathfrak{X})\}.$$

En la Subsección 5.3.1 más abajo, veremos una caracterización de la distancia de Prokhorov.

5.3.1. El Teorema de Strassen

El Teorema de Strassen [28], Teorema 11 y su Corolario, caracterizan la distancia de Prokhorov en términos de una medida de probabilidad bivariada, con marginales dadas. El argumento crucial se dirige a una aplicación del Teorema de Separación de Hahn-Banach, el Teorema 14. La dificultad, como en muchas de las aplicaciones del teorema de separación en dimensiones infinitas, es demostrar que el conjunto convexo al cual se aplicará el teorema es cerrado, en la topología seleccionada. Nosotros haremos una esquematización de los detalles, para llegar a la aplicación del Teorema de Hahn-Banach siguiendo la exposición de Föllmer y Schied [11], Sección 2.6.

Topología débil

Dejaremos fija una función continua

$$\psi : \mathfrak{X} \rightarrow [1, +\infty).$$

Denotamos mediante $C_\psi(\mathfrak{X})$ a las funciones continuas que son dominadas por ψ , en el siguiente sentido:

$$C_\psi(\mathfrak{X}) := \{f \in C(\mathfrak{X}) \mid \exists c : |f(x)| \leq c \cdot \psi(x) \text{ para } x \in \mathfrak{X}\}.$$

La función ψ determina una clase de medidas de probabilidad, con respecto de las cuales su esperanza es finita,

$$\mathcal{M}_1^\psi(\mathfrak{X}) := \left\{ \mu \in \mathcal{M}_1(\mathfrak{X}) \mid \int \psi(x) \mu(dx) < \infty \right\}.$$

La topología débil en $\mathcal{M}_1^\psi(\mathfrak{X})$ es la mínima topología bajo la cual los siguientes mapeos son continuos:

$$\mu \rightarrow \int f d\mu, \text{ para } \mu \in \mathcal{M}_1^\psi(\mathfrak{X}),$$

para toda $f \in C_\psi(\mathfrak{X})$.

Consideramos el espacio $\mathcal{M}_1^{\bar{\psi}}(\mathfrak{X} \times \mathfrak{X})$ con la topología débil generada por la función

$$\bar{\psi}(x, y) := \psi(x) + \psi(y).$$

Sea $L\mathcal{M}_1^\psi(\mathfrak{X})$ el espacio lineal generado por las medidas en $\mathcal{M}_1^\psi(\mathfrak{X})$ es decir

$$L\mathcal{M}_1^\psi(\mathfrak{X}) := \{a\mu - b\nu \mid \mu, \nu \in \mathcal{M}_1^\psi(\mathfrak{X}); a, b \in \mathbb{R}\}.$$

En forma análoga a la topología débil definida en $\mathcal{M}_1^\psi(\mathfrak{X})$, consideramos la topología débil generada por los mapeos

$$\rho \rightarrow a \int f d\mu - b \int f d\nu, \text{ para } \rho = a\mu - b\nu \in LM_1^\psi(\mathfrak{X}).$$

Clases de medidas de probabilidad con marginales dadas

El siguiente es el Teorema 7 de Strassen en [28]. En la demostración esquematizamos los pasos más importantes, destacando las ideas principales.

Teorema 69. *Sea $\Lambda \subset \mathcal{M}_1^{\bar{\psi}}(\mathfrak{X} \times \mathfrak{X})$ un conjunto convexo y cerrado en la topología $\bar{\psi}$ -débil. Sean μ y ν dos medidas de probabilidad en $\mathcal{M}_1^\psi(\mathfrak{X})$. Entonces existe $\lambda \in \Lambda$ con distribuciones marginales μ y ν si y solo si*

$$\int f_1 d\mu + \int f_2 d\nu \leq \sup_{\lambda \in \Lambda} \int (f_1(x) + f_2(y)) \lambda(dx, dy), \text{ para todo } f_1, f_2 \in C_\psi(\mathfrak{X}). \quad (5.3)$$

Demostración. (bosquejo)

1. Considere copias idénticas L_1 y L_2 de $LM_1^\psi(\mathfrak{X})$. Definimos un mapeo

$$I : \mathcal{M}_1^{\bar{\psi}}(\mathfrak{X} \times \mathfrak{X}) \rightarrow L_1 \times L_2,$$

de la siguiente forma

$$I(\lambda) := (\pi_1 \lambda, \pi_2 \lambda),$$

donde π_i representa la proyección a las marginales:

$$\pi_i : \mathcal{M}_1^{\bar{\psi}}(\mathfrak{X} \times \mathfrak{X}) \rightarrow L_i.$$

La imagen directa de Λ bajo I , definida por

$$H_\Lambda := \{I(\lambda) \mid \lambda \in \Lambda\},$$

es claramente convexa. Más aun, satisface que es un conjunto cerrado en la topología débil de $LM_1^\psi(\mathfrak{X})$. Este es el paso más delicado en la demostración y utiliza el hecho de que \mathfrak{X} es un espacio polaco y el Teorema de Prokhorov, en la caracterización de conjuntos de medidas relativamente compactos.

2. Existe una funcional lineal $l : L_1 \times L_2 \rightarrow \mathbb{R}$ tal que

$$l(\mu, \nu) > \sup_{(\mu', \nu') \in H_\Lambda} l(\mu', \nu'),$$

si y sólo si $(\mu, \nu) \notin H_\Lambda$, debido al Teorema 14, de separación de Hahn-Banach. Es posible aplicar el Teorema de separación debido al paso anterior, ya que el conjunto $C := H_\Lambda$ es convexo y cerrado, y el conjunto $B := \{(\mu, \nu)\}$ es compacto.

3. La funcional lineal l necesariamente es de la forma

$$l(\rho_1, \rho_2) = \int f_1 d\rho_1 + \int f_2 d\rho_2,$$

para algún par de funciones $f_1, f_2 \in C_\psi(\mathfrak{X})$.

□

Variantes y consecuencias del Teorema de Strassen

Utilizaremos la siguiente notación para la banda cerrada de longitud ϵ sobre la diagonal

$$D(\epsilon) := \{(x_1, x_2) \in \mathfrak{X} \times \mathfrak{X} \mid \mathfrak{p}_{\mathfrak{X}}(x_1, x_2) \leq \epsilon\}.$$

El siguiente es el Teorema 11 de Strassen [28], enunciado para el caso de nuestro interés, que es el conjunto $D(\epsilon)$.

Teorema 70 (Strassen). Sean $\mu, \nu \in \mathcal{M}_1(\mathfrak{X})$. Entonces existe una medida de probabilidad λ en $\mathfrak{X} \times \mathfrak{X}$ con marginales μ y ν tal que

$$\lambda[D(\epsilon)] > 1 - \epsilon, \quad (5.4)$$

si y sólo si para todo subconjunto abierto U de \mathfrak{X} , se cumple que

$$\nu(U) \leq \mu(\pi_1(D(\epsilon) \cap (\mathfrak{X} \times U))) + \epsilon. \quad (5.5)$$

Demostración. Definimos una clase convexa de medidas de probabilidad

$$\Lambda := \{\lambda \mid \lambda[D(\epsilon)] \geq 1 - \epsilon\}.$$

Definimos la función ψ constante $\psi(x) = 1$. En este caso, el espacio $C_\psi(\mathfrak{X})$ consiste de funciones continuas acotadas. Vamos a verificar que la

condición (5.5) implica la condición (5.3), para el par (μ, ν) . Entonces el resultado será consecuencia del Teorema 69.

Sean $f_1, f_2 \in C_\psi(\mathfrak{X})$. Para A un conjunto medible de $\mathfrak{X} \times \mathfrak{X}$, sea

$$N(A) := \sup\{f_1(x) + f_2(y) \mid (x, y) \in A\}.$$

Observemos que el supremo en la desigualdad (5.3), es igual a

$$N = N(D(\epsilon)) \cdot (1 - \epsilon) + N(D^c(\epsilon)) \cdot \epsilon,$$

debido a la definición de la clase Λ .

Las funciones f_1 y f_2 son acotadas y podemos suponer, sin pérdida de generalidad, que son no negativas. Sea $m := \sup f_2$ y denotemos

$$D_x := \{y \in \mathfrak{X} \mid (x, y) \in D(\epsilon)\}.$$

Más aun, para $x \in \mathfrak{X}$ definimos $f^+(x) := \sup\{f_2(y) \mid y \in D_x\}$. Tenemos

$$\begin{aligned} \int f_2 d\nu &= \int_0^m \nu[\{y \in \mathfrak{X} \mid f_2(y) > r\}] dr \\ &\leq \int_0^m 1 \wedge (\epsilon + \mu[\{x \in \mathfrak{X} \mid f^+(x) > r\}]) dr \end{aligned} \quad (5.6)$$

$$\leq \int_0^m (1 - \epsilon) \wedge \mu[\{x \in \mathfrak{X} \mid f^+(x) > r\}] dr + \epsilon \cdot m \quad (5.7)$$

$$= \int_0^{\sup f^+} (1 - \epsilon) \wedge \mu[\{x \in \mathfrak{X} \mid f^+(x) > r\}] dr + \epsilon \cdot m. \quad (5.8)$$

En (5.6) utilizamos la desigualdad (5.5) y la definición de la función f^+ .

En (5.7) se hace uso de la desigualdad elemental

$(g(r) + \epsilon) \wedge 1 \leq g(r) \wedge (1 - \epsilon) + \epsilon$, para una función g arbitraria. En (5.8) utilizamos el hecho de que $m = \sup f^+$, debido a la definición de esta función.

Denotemos por $M^\mu(\epsilon)$, las medidas de probabilidad dominadas por μ y que estan normalizadas a $1 - \epsilon$. Se observa que

$$\int_0^{\sup f^+} (1 - \epsilon) \wedge \mu[\{x \in \mathfrak{X} \mid f^+(x) > r\}] dr = \sup \left\{ \int f^+ d\tilde{\mu} \mid \tilde{\mu} \in M^\mu(\epsilon) \right\}.$$

Como consecuencia, obtenemos la estimación

$$\int f_1 d\mu + \int f_2 d\nu \leq \sup \left\{ \int (f_1 + f^+) d\tilde{\mu} \mid \tilde{\mu} \in M^\mu(\epsilon) \right\} + \epsilon \cdot (\sup f_1 + m) \leq N.$$

□

El siguiente resultado es el Corolario al Teorema 11 de Strassen en [28].

Teorema 71. *Sean $\mu, \nu \in \mathcal{M}_1(\mathfrak{X})$. Entonces la distancia de Prokhorov de medidas satisface*

$$\mathfrak{d}_{\text{prokh}}(\mu, \nu) \leq \epsilon,$$

si y sólo si existe una medida de probabilidad λ , en $\mathfrak{X} \times \mathfrak{X}$, tal que las medidas marginales de λ son precisamente μ y ν . Además,

$$\lambda[D(\epsilon)] > 1 - \epsilon. \quad (5.9)$$

Demostración. *Notemos que si la distancia de Prokhorov entre las medidas μ y ν es menor o igual que ϵ , entonces se satisface la condición (5.5), esto demuestra la implicación de la existencia de la medida de probabilidad λ , con las condiciones requeridas.*

Ahora verificamos que se satisface la implicación recíproca. Sea λ una medida de probabilidad en $\mathfrak{X} \times \mathfrak{X}$ satisfaciendo la desigualdad (5.9) y con marginales μ y ν . Para $A \in \mathcal{B}$ tenemos

$$\nu(A) = \lambda(A \times \Omega) = \lambda(A \times \Omega \cap D(\epsilon)) + \lambda(A \times \Omega \cap D^c(\epsilon)).$$

Como consecuencia,

$$\nu(A) \leq \lambda(A \times \Omega \cap D(\epsilon)) + \epsilon. \quad (5.10)$$

Ahora bien, tenemos las relaciones de conjuntos

$$A \times \Omega \cap D(\epsilon) = A \times A^\epsilon \quad (5.11)$$

$$\subset \Omega \times A^\epsilon \cap D(\epsilon). \quad (5.12)$$

En efecto, para establecer (5.11) probamos la doble contención. En este caso, si $(\omega_1, \omega_2) \in A \times \Omega \cap D(\epsilon)$ tenemos $\mathfrak{p}(\omega_1, \omega_2) < \epsilon$ y como consecuencia $\omega_2 \in A^\epsilon$. La contención recíproca se sigue de la relación $A \times A^\epsilon \subset D(\epsilon)$. La contención (5.12) se establece análogamente.

Tenemos entonces

$$\lambda(A \times \Omega \cap D(\epsilon)) \leq \lambda(\Omega \times A^\epsilon \cap D(\epsilon)) \leq \mu(A^\epsilon). \quad (5.13)$$

Las desigualdades (5.10) y (5.13) implican

$$\nu(A) \leq \mu(A^\epsilon) + \epsilon,$$

estableciendo $\mathfrak{d}_{\text{prokh}}(\mu, \nu) \leq \epsilon$. □

Una variante en la forma de enunciar el Teorema 71 es la siguiente.

Teorema 72 (Teorema de Strassen, segunda versión). *Sean $\epsilon, \delta > 0$. Para $\mu, \nu \in \mathcal{M}_1(\mathfrak{X})$, la condición*

$$\mu(A) \leq \nu(A^\delta) + \epsilon, A \in \mathcal{B}$$

es equivalente a la siguiente propiedad. Existe un espacio de probabilidad $(\Omega', \mathcal{F}', P')$ y variables aleatorias $X, Y : \Omega' \rightarrow \mathfrak{X}$ tales que $L_{P'}(X) = \mu$ y $L_{P'}(Y) = \nu$ y además

$$P'[\omega \in \Omega \mid \mathfrak{p}_{\mathfrak{X}}(X(\omega), Y(\omega)) \leq \delta] > 1 - \epsilon. \quad (5.14)$$

5.4. Teorema de Robustez

El Teorema de Hampel en [13], caracteriza la robustez de un estimador estadístico. Intuitivamente, el teorema relaciona dos formas de continuidad de la funcional: bajo ciertas condiciones la continuidad puntual implica “continuidad en ley”. Para enunciar rigurosamente el teorema es necesario introducir algunos conceptos. Seguiremos la presentación reciente de Krätschmer et al. [18].

Consideremos fija una métrica \mathfrak{d}_1 en la clase $\mathcal{M}_1(\mathfrak{X})$, así como un subconjunto \mathcal{M} . Suponemos que se satisface la contención

$$\mathfrak{F}(\mathfrak{X}) \subset \mathcal{M}.$$

Sea \mathbf{T} un espacio medible en el que se haya definida una métrica $\mathfrak{p}_{\mathbf{T}}$. Denotaremos mediante \mathcal{T} a la σ -álgebra de conjuntos borelianos de \mathbf{T} . Supongamos fija una métrica \mathfrak{d}_2 en la clase $\mathcal{M}_1(\mathbf{T})$. Un estimador es una funcional medible

$$T : \mathcal{M} \rightarrow \mathbf{T}.$$

Recuerde que en la Sección 5.2, introdujimos el espacio de observaciones definido como

$$\mathfrak{D} := \mathfrak{X}^{\mathbb{N}}.$$

Para la distribución F , la medida de probabilidad \mathbb{Q}_F se definió en la ecuación (5.2), como

$$\mathbb{Q}_F := F^{\mathbb{N}}.$$

Utilizaremos esta notación para otras distribuciones en \mathfrak{X} . Un elemento $x \in \mathfrak{D}$ tiene la forma $x = (x_1, \dots, x_n, \dots)$. Definimos una medida empírica

determinada por las primeras n observaciones:

$$m_{x^{(n)}} := \frac{1}{n} \sum_{i=1}^n \delta_{x_i}.$$

Hipótesis 73. *El mapeo*

$$T_n : \mathfrak{D} \rightarrow \mathbf{T},$$

definido por

$$T_n(x) := T(m_{x^{(n)}}),$$

es $(\mathcal{B}(\mathfrak{D}), \mathcal{T})$ -medible.

Observación 74. *En la Sección 5.5 discutiremos nuestra Hipótesis 73 en el contexto de funciones puntuales simétricas, para el caso especial \mathbb{R}^n .*

Definición 75 (\mathcal{M}_0 -Robustez cualitativa). *Sea \mathcal{M}_0 un subconjunto de \mathcal{M} y sea $\mu \in \mathcal{M}_0$. La sucesión de estimadores $\{T_n\}_{n \in \mathbb{N}}$ se dice que es cualitativamente \mathcal{M}_0 -Robusta en μ con respecto a las métricas $(\mathfrak{d}_1, \mathfrak{d}_2)$ si para cada $\epsilon > 0$ existe alguna $\delta > 0$ y $n_0 \in \mathbb{N}$ tal que para toda $\nu \in \mathcal{M}_0$ y $n \geq n_0$ tenemos*

$$\mathfrak{d}_1(\mu, \nu) \leq \delta \implies \mathfrak{d}_2(\mathbb{Q}_\mu \circ T_n^{-1}, \mathbb{Q}_\nu \circ T_n^{-1}) \leq \epsilon.$$

En el caso especial $\mathcal{M}_0 = \mathcal{M} = \mathcal{M}_1(\mathbb{R})$, la definición coincide con la noción clásica de robustez de Hampel, vista en [13], cuando $\mathfrak{d}_1 = \mathfrak{d}_2 = \mathfrak{d}_{prokh}$. En situaciones concretas, establecer la robustez de un estadístico T puede ser difícil de verificar. El Teorema célebre de Hampel, da una condición suficiente para el caso especial en que la robustez se define en términos de la métrica de Prokhorov.

Introducimos dos conceptos, que serán condiciones suficientes, para asegurar robustez en el sentido de la Definición 75.

Definición 76 (\mathcal{N} -continuidad). *Sea $\mu \in \mathcal{M}$ y \mathcal{N} un subconjunto de \mathcal{M} . El estimador $T : \mathcal{M} \rightarrow E$ se dice \mathcal{N} -continuo en μ con respecto a $(\mathfrak{d}_1, \mathfrak{p}_{\mathbf{T}})$, si para cada $\epsilon > 0$ existe alguna $\delta > 0$ tal que para toda $\nu \in \mathcal{N}$*

$$\mathfrak{d}_1(\mu, \nu) \leq \delta \implies \mathfrak{p}_{\mathbf{T}}(T(\mu), T(\nu)) \leq \epsilon.$$

Definición 77 (Propiedad uniforme de Glivenko-Cantelli). *Sea \mathcal{M}_0 un subconjunto de \mathcal{M} . El espacio métrico $(\mathcal{M}_0, \mathfrak{d}_1)$ satisface la propiedad uniforme de Glivenko-Cantelli (UGC), si para toda $\epsilon > 0$ y $\delta > 0$ existe $n_0 \in \mathbb{N}$ tal que para cada $\mu \in \mathcal{M}_0$ y $n \geq n_0$ se satisface*

$$\mathbb{Q}_\mu [x \in \mathfrak{D} \mid \mathfrak{d}_1(\mu, m_{x^{(n)}}) \geq \delta] \leq \epsilon.$$

Después de los preparativos, estamos listos para enunciar una versión generalizada del Teorema de Hampel, debido a [18], Teorema 2.4.

Teorema 78 (Teorema de Robustez). *Sea \mathcal{M}_0 un subconjunto de \mathcal{M} y sea $\mu_0 \in \mathcal{M}_0$. Suponga que $(\mathcal{M}_0, \mathfrak{d}_1)$ tiene la propiedad UGC. Si el mapeo T es \mathfrak{F} -continuo en μ_0 con respecto a $(\mathfrak{d}_1, \mathfrak{p}_T)$, entonces la sucesión $\{T_n\}_{n \in \mathbb{N}}$ es cualitativamente \mathcal{M}_0 -Robusta con respecto a $(\mathfrak{d}_1, \mathfrak{d}_{prokh})$ en μ_0 .*

Demostración. 1. Para $\epsilon > 0$ existe $\delta > 0$ tal que se cumple que

$$\mathfrak{d}_1(\mu_0, m_{x^{(n)}}) \leq 2\delta \implies \mathfrak{p}_T(T(\mu_0), T(m_{x^{(n)}})) \leq \frac{\epsilon}{2}, \quad (5.15)$$

debido a la \mathfrak{F} -continuidad de T en μ_0 .

2. Existe $n_0 \in \mathbb{N}$ tal que para toda $\nu \in \mathcal{M}_0$ se satisface

$$\mathbb{Q}_\mu [x \in \mathfrak{D} \mid \mathfrak{d}_1(\nu, m_{x^{(n)}}) \geq \delta] \leq \frac{\epsilon}{2},$$

debido a la propiedad UGC de $(\mathcal{M}_0, \mathfrak{d}_1)$. Una forma equivalente de escribir la anterior desigualdad es

$$\mathbb{Q}_\mu [x \in \mathfrak{D} \mid \mathfrak{d}_1(\nu, m_{x^{(n)}}) \leq \delta] \geq 1 - \frac{\epsilon}{2}, \quad (5.16)$$

3. Fijemos $\nu \in \mathcal{M}_0$ tal que $\mathfrak{d}_1(\mu_0, \nu) \leq \delta$. La desigualdad del triangulo justifica la siguiente implicación:

$$\mathfrak{d}_1(\nu, m_{x^{(n)}}) \leq \delta \implies \mathfrak{d}_1(\mu_0, m_{x^{(n)}}) \leq 2\delta. \quad (5.17)$$

4. Tenemos la contención de conjuntos

$$\{x \in \mathfrak{D} \mid \mathfrak{d}_1(\nu, m_{x^{(n)}}) \leq \delta\} \subset \left\{x \in \mathfrak{D} \mid \mathfrak{p}_T(T(\mu_0), T(m_{x^{(n)}})) \leq \frac{\epsilon}{2}\right\},$$

debido a las implicaciones (5.17) y (5.15). Como consecuencia tenemos

$$\mathbb{Q}_\mu \left[\left\{x \in \mathfrak{D} \mid \mathfrak{p}_T(T(\mu_0), T(m_{x^{(n)}})) \leq \frac{\epsilon}{2}\right\} \right] \geq 1 - \frac{\epsilon}{2}, \quad (5.18)$$

debido a la estimación (5.16).

5. Considere el estimador constante $T' : \mathcal{M} \rightarrow \mathbf{T}$ definido por $T'(\mu) := T(\mu_0)$ para toda $\mu \in \mathcal{M}$. La ley de T' en $\mathcal{M}_1(\mathbf{T})$ con respecto a μ_0 es $\delta_{T(\mu_0)}$. La desigualdad (5.18) implica

$$\mathfrak{d}_{\text{prokh}}(\delta_{T(\mu_0)}, \mathbb{Q}_\nu \circ T_n^{-1}) \leq \frac{\epsilon}{2}$$

debido al Teorema de Strassen en la forma del Teorema 72.

6. La demostración concluye al observar la siguiente propiedad. Sean $a, b \in \mathbf{T}$ tales que $\mathfrak{p}_{\mathbf{T}}(a, b) < \epsilon$. Entonces

$$\mathfrak{d}_{\text{prokh}}(\delta_a, \delta_b) < \epsilon.$$

□

5.5. Estructura de espacio de probabilidad para \mathfrak{F}_n

En esta sección, nuestro objetivo es proporcionar de una estructura probabilística al conjunto de medidas de probabilidad empíricas, que se forman con n observaciones. Para este fin, el grupo de permutaciones es crucial y comenzamos con la definición de este concepto.

Definición 79. 1. Denotemos mediante CP_n la clase de permutaciones de n elementos, es decir, de biyecciones $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$.

2. Para $f \in CP_n$ definimos un mapeo $f : \mathfrak{X}^n \rightarrow \mathfrak{X}^n$ definido de la siguiente forma. Para (x_1, \dots, x_n) elemento de \mathfrak{X}^n

$$f(x_1, \dots, x_n) := (x_{f(1)}, \dots, x_{f(n)}). \quad (5.19)$$

3. Sea $A \subset \mathfrak{X}^n$. La simetrización de A es el conjunto definido por

$$[A] := \{x \in \mathfrak{X}^n \mid x = f(a) \text{ para algun } a \in A \text{ y para } f \in CP_n\}.$$

4. Sea $A \subset \mathfrak{X}^n$, decimos que A es un conjunto simétrico si $A = [A]$.

Toda permutación $f \in CP_n$ induce una transformación del espacio \mathfrak{X}^n de acuerdo a la regla (5.19). En particular, para $A \in \mathcal{B}(\mathfrak{X}^n)$ la notación $f(A)$ indica la imagen directa bajo f del conjunto A . El siguiente resultado caracteriza la operación de simetrización. La demostración es simple y la omitimos.

Proposición 80. Para $A \in \mathcal{B}(\mathfrak{X}^n)$

$$[A] = \bigcup_{f \in CP_n} f(A).$$

Para continuar con nuestra discusión haremos la siguiente hipótesis

Hipótesis 81. Para todo $A \in \mathcal{B}(\mathfrak{X}^n)$, la simetrización $[A]$ es un conjunto medible en $\mathcal{B}(\mathfrak{X}^n)$.

Como consecuencia de la Proposición 80, en el caso en que $\mathfrak{X} = \mathbb{R}^d$ se verifica nuestra Hipótesis 81.

Corolario 82. Para $\mathfrak{X} = \mathbb{R}^d$ y $A \in \mathcal{B}(\mathfrak{X}^n)$ tenemos $[A] \in \mathcal{B}(\mathfrak{X}^n)$.

Demostración. Observemos que la función inducida por una permutación $f \in CP_n$ es una transformación lineal y entonces $f(A) \in \mathcal{B}(\mathfrak{X}^n)$. Como consecuencia $[A]$ pertenece a $\mathcal{B}(\mathfrak{X}^n)$ debido a la Proposición 80, ya que una σ -álgebra es estable bajo uniones finitas. \square

Teorema 83. Bajo la hipótesis 81, la clase de conjuntos borelianos simétricos, denotada $\mathcal{B}_{sim}(\mathfrak{X}^n)$, forma una sub- σ -álgebra de $\mathcal{B}(\mathfrak{X}^n)$.

Demostración. Sea $A \in \mathcal{B}(\mathfrak{X}^n)$ un conjunto simétrico. Vamos a demostrar que A^c también es un conjunto simétrico. Por contradicción, suponga que $b := f(a) \notin A^c$ para algún $a \in A^c$ y $f \in CP_n$. Observe que la función inversa f^{-1} también es una permutación, y que $b \in A$ implica que $a = f^{-1}(b) \in A$, una contradicción. Lo anterior demuestra que $A^c \in \mathcal{B}_{sim}(\mathfrak{X}^n)$ siempre que $A \in \mathcal{B}_{sim}(\mathfrak{X}^n)$.

Sea $\{A_i\}_{i \in \mathbb{N}} \subset \mathcal{B}_{sim}(\mathfrak{X}^n)$. La unión de conjuntos simétricos es claramente estable y tenemos entonces $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{B}_{sim}(\mathfrak{X}^n)$. \square

Ahora definimos una σ -álgebra en \mathfrak{F}_n .

Definición 84. Definimos un mapeo $\hat{\mu} : \mathfrak{X}^n \rightarrow \mathfrak{F}_n$ de tal forma que

$$\hat{\mu}(x) := \frac{1}{n} \sum_{i=1}^n \delta_{x_i}, \text{ para } x \in \mathfrak{X}^n.$$

Definimos \mathcal{G}_n la clase de subconjuntos de \mathfrak{F}_n , que son imagen directa, bajo esta transformación, de conjuntos borelianos simétricos, es decir

$$\mathcal{G}_n := \{B \subset \mathfrak{F}_n \mid B = \hat{\mu}(A) \text{ para algún } A \in \mathcal{B}_{sim}(\mathfrak{X}^n)\}.$$

Proposición 85. La clase de conjuntos \mathcal{G}_n es una σ -álgebra.

Demostración. Sea $B \in \mathcal{G}_n$ de la forma $B = \hat{\mu}(A)$, para algun $A \in \mathcal{B}_{sim}(\mathfrak{X}^n)$. Es claro que $\hat{\mu}(A^c) \cap B = \emptyset$, lo cual demuestra que $B^c \in \mathcal{G}_n$, ya que en este caso $B^c = \hat{\mu}(A^c)$.
Sea $\{B_i\}_{i \in \mathbb{N}} \subset \mathcal{G}_n$ de la forma $B_i = \hat{\mu}(A_i)$, para $A_i \in \mathcal{B}_{sim}(\mathfrak{X}^n)$. Tenemos

$$\bigcup_{i \in \mathbb{N}} B_i = \bigcup_{i \in \mathbb{N}} \hat{\mu}(A_i) = \hat{\mu} \left(\bigcup_{i \in \mathbb{N}} A_i \right).$$

Mostrando que $\bigcup_{i \in \mathbb{N}} \hat{\mu}(A_i) \in \mathcal{G}_n$. □

Definición 86. Denotemos mediante $F^{\otimes n}$ la medida de probabilidad definida en \mathfrak{X}^n , que es el producto de n copias independientes de F . El mapeo $\hat{\mu}$ induce la medida de probabilidad $F^{\otimes n}$ a una medida de probabilidad en \mathfrak{F}_n , definida en la σ -álgebra \mathcal{G}_n . La definición es la siguiente

$$Q_n(B) := F^{\otimes n}((\hat{\mu})^{-1}(B)). \quad (5.20)$$

Capítulo 6

Robustez comparativa

El concepto de robustez introducido por Hampel [13] en 1971, ha sido referente en el quehacer estadístico desde entonces. La equivalencia entre la continuidad de un estimador con respecto a la topología débil y la robustez, es el resultado central del citado artículo. En el capítulo anterior, hemos desarrollado una versión actualizada de la prueba de dicho resultado. En este capítulo investigamos una concepción reciente de la idea de robustez. La misma es propuesta por Krätschmmer, Schied y Zähle en [17] y se trata de la robustez comparativa.

Como primer paso, hacemos una revisión general de la teoría de espacios de Orlicz, así como de las funciones que los generan. Seguidamente estudiamos las propiedades de consistencia y continuidad de los estimadores. Con el contexto esclarecido y contando con los resultados básicos necesarios sobre los estimadores, definimos la robustez comparativa. Este nuevo concepto es similar al clásico de Hampel, con la diferencia de que ahora no tomamos en cuenta la continuidad en la topología débil. En su lugar, centramos el análisis en toda una gama de topologías débiles, relacionadas con distintas funciones de Young. Esta flexibilidad nos permite hablar, no de los estimadores que son robustos y los que no, sino que contaremos con diferentes grados de robustez. De hecho, tendremos un índice de robustez, el cual mostraremos es efectivo para categorizar nuestras medidas de riesgo. Finalizamos este apartado con un estudio particular de las medidas de riesgo convexas e invariantes en ley, así como las que son producto de distorsiones cóncavas, de acuerdo a su índice de robustez.

6.0.1. Espacios de Orlicz y consistencia de estimadores

Consideramos un espacio de probabilidad sin átomos $(\Omega, \mathcal{F}, \mathbb{P})$ y al conjunto de variables aleatorias finitas L^0 . En particular, tomamos $\mathcal{X} \subset L^0$ un espacio vectorial que contenga a las constantes. Así, la variable aleatoria $X \in \mathcal{X}$ representará las ganancias o pérdidas de una posición financiera en particular. Llamaremos $\mathcal{M}(\mathcal{X})$ al conjunto de distribuciones generadas por los elementos del espacio vectorial \mathcal{X} . Asimismo, fijamos una medida de riesgo invariante en ley ρ . Observamos que la propiedad de invarianza en ley es equivalente a la existencia de un mapeo $\mathcal{R}_\rho : \mathcal{M}(\mathcal{X}) \rightarrow \mathbb{R}$, donde $\mathcal{R}_\rho(\mathbb{P} \circ X^{-1}) = \rho(X)$. Este mapeo es conocido como el funcional de riesgo asociado con ρ .

Definición 87. *Un mapeo $\Psi : \mathbb{R}_+ \rightarrow [0, \infty]$ se llama función de Young, si es continuo por la izquierda, no decreciente, convexo y se cumple que:*

- $\lim_{x \uparrow \infty} \Psi(x) = \infty$.
- $\lim_{x \downarrow 0} \Psi(x) = \Psi(0) = 0$.

La importancia de las funciones de Young radica en el tipo de espacios que se pueden relacionar con ellas. Dentro de estos espacios, se encuentran los llamados de Orlicz, los cuales se definen a continuación.

Definición 88. *El espacio de Orlicz generado por una función de Young Ψ se define como:*

$$L^\Psi = L^\Psi(\Omega, \mathcal{F}, \mathbb{P}) = \{X \in L^0 : \exists c > 0, \mathbb{E}[\Psi(c|X|)] < \infty\},$$

mientras que el núcleo de Orlicz está dado por:

$$H^\Psi = H^\Psi(\Omega, \mathcal{F}, \mathbb{P}) = \{X \in L^0 : \mathbb{E}[\Psi(c|X|)] < \infty, \forall c > 0\}.$$

Podemos notar que los espacios de Orlicz contienen variables aleatorias que no son necesariamente acotadas. Esto hace que el conjunto de las posiciones financieras que podamos manejar, sea distinto de L^∞ . De hecho, se conoce que un espacio de Orlicz es de Banach, cuando se dota de la norma de Luxemburgo

$$\|Y\|_\Psi = \inf\{\lambda > 0 : \mathbb{E}[\Psi(|Y|/\lambda)] \leq 1\}.$$

De acuerdo con [7], el espacio de Orlicz L^Ψ posee propiedades tales como la propiedad de Fatou, que el núcleo de Orlicz H^Ψ puede no tener. En contraste, características como la densidad de las funciones simples,

pueden ser encontradas en H^Ψ , pero no en L^Ψ . De esta manera, es interesante saber cuándo el espacio coincide con su núcleo. La condición Δ_2 para una función de Young es la siguiente:

$$\exists C \in \mathbb{R}, x_0 > 0 : \Psi(2x) \leq C\Psi(x), \quad \forall x \geq x_0.$$

Un espacio de Orlicz L^Ψ coincide con su núcleo H^Ψ , siempre que la correspondiente función de Young cumpla la condición Δ_2 .

Las variables aleatorias a considerar serán un subconjunto de un espacio de Orlicz, generalmente el núcleo, para una función de Young dada Ψ . Para medir el riesgo de una posición X , es suficiente conocer la distribución de X , lo cual casi nunca sucede. En cambio, trabajamos con muestras, las cuales a su vez generan una distribución empírica, para aproximar la ley de X . En este sentido, denotemos por

$$\hat{m}_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$$

a la distribución empírica de $\{X_i\}_{i=1}^n$ y como $\hat{\rho}_n = \mathcal{R}_\rho(\hat{m}_n)$ al estimador de $\rho(X)$. Otra característica importante del estudio para la estimación del riesgo, es considerar a las variables que generan la muestra, como no necesariamente independientes. En lugar del supuesto de independencia, se puede trabajar con colecciones de mapeos aleatorios que cumplan algún tipo de ley de grandes números. En esta sección consideramos sucesiones estacionarias y ergódicas de variables aleatorias idénticamente distribuidas.

Definición 89. *Un conjunto $A \in \mathcal{F}$ se dice invariante si existe $B \in B(\mathbb{R}^\infty)$ tal que*

$$A = \{\omega \in \Omega : (X_n(\omega), X_{n+1}(\omega), \dots) \in B\}, \quad \forall n \in \mathbb{N}.$$

Definición 90. *Un proceso estacionario $\{X_i\}_{i \in \mathbb{N}}$ es ergódico si todos los conjuntos invariantes tienen probabilidad cero o uno.*

En ocasiones no trabajamos directamente con el proceso, sino con una transformación del mismo. Es importante saber que bajo ciertas condiciones, dicha transformación sigue produciendo un proceso ergódico. La prueba de la siguiente proposición se puede encontrar en Breiman [1].

Proposición 91. *Sea $\{X_i\}_{i \in \mathbb{N}}$ un proceso estacionario y ergódico. Si $\phi(x)$ es un mapeo $B(\mathbb{R}^\infty)$ -medible, entonces el proceso $\{Y_i\}_{i \in \mathbb{N}}$ definido como*

$$Y_k = \phi(X_k, X_{k+1}, \dots)$$

es ergódico.

Como se había mencionado con anterioridad, el proceso con el que trabajamos debe satisfacer algún resultado del tipo de la ley de grandes números. En este caso, el Teorema ergódico de Birkhoff será de utilidad para describir propiedades asintóticas del proceso.

Teorema 92 (Teorema ergódico de Birkhoff). *Sea $\{X_i\}_{i \in \mathbb{N}}$ es un proceso estacionario y ergódico, así como $\phi(x)$ un mapeo $B \in B(\mathbb{R}^\infty)$ -medible. Si $\{Y_i\}_{i \in \mathbb{N}}$ es el proceso definido como en la proposición anterior y $\mathbb{E}[|Y_1|] < \infty$, entonces, casi seguramente*

$$\frac{1}{n} \sum_{k=1}^n Y_k \rightarrow \mathbb{E}[Y_1].$$

Ahora presentamos el teorema de consistencia para el funcional de riesgo \mathcal{R}_ρ , asociado a ρ .

Teorema 93. *Supongamos que ρ es una medida de riesgo convexa invariante en ley, definida en H^Ψ . Sea $\{X_i\}_{i \in \mathbb{N}}$ una sucesión estacionaria y ergódica de variables aleatorias idénticamente distribuidas, tales que $X_1 \stackrel{d}{=} X$, con $X \in H^\Psi$ y μ la distribución de X . Entonces $\hat{\rho}_n$ es un estimador consistente para $\rho(X)$, en el sentido de que $\hat{\rho}_n \rightarrow \rho(X)$ casi-seguramente.*

Demostración. *Dado que $X \in H^\Psi$, entonces $\mathbb{E}[\Psi(k|X|)] < \infty$, para toda $k > 0$. Del Teorema 92, para cada $k > 0$ existe $\Omega_1 \subset \Omega$ tal que $\mathbb{P}(\Omega_1) = 1$ y si $\omega \in \Omega_1$,*

$$\int \Psi(k|x|) \hat{m}_n(\omega)(dx) = \frac{1}{n} \sum_{i=1}^n \Psi(k|X_i(\omega)|) \rightarrow \mathbb{E}[\Psi(k|X|)] = \int \Psi(k|x|) \mu(dx). \quad (6.1)$$

De hecho, gracias al Teorema 92, existe $\Omega_2 \subset \Omega$ con $\mathbb{P}(\Omega_2) = 1$ y tal que

$$\hat{m}_n(\omega) \Rightarrow \mu, \quad (6.2)$$

si $\omega \in \Omega_2$. Definimos a Ω_0 , como el conjunto de probabilidad 1 tal que (6.1) y (6.2) se cumplen para todo $\omega \in \Omega_0$ y $k \in \mathbb{N}$. Fijamos $\omega_0 \in \Omega_0$. Por una aplicación del teorema de representación de Skorokhod, garantizamos la existencia de las variables aleatorias X^{ω_0} y $\{X_n^{\omega_0}\}_{n \in \mathbb{N}}$, tales que X^{ω_0} tiene distribución μ y cada $X_n^{\omega_0}$ tiene ley $\hat{m}_n(\omega_0)$. Entonces $X_n^{\omega_0} \rightarrow X^{\omega_0}$ casi-seguramente. Así, para cada $n \in \mathbb{N}$ y por (6.1),

$$\mathbb{E}[\Psi(k|X_n^{\omega_0}|)] = \int \Psi(k|x)\hat{m}_n(\omega_0)(dx) \longrightarrow \int \Psi(k|x)\mu(dx) = \mathbb{E}[\Psi(k|X^{\omega_0}|)].$$

Lo anterior quiere decir que $\Psi(k|X_n^{\omega_0}|) \rightarrow \Psi(k|X^{\omega_0}|)$ en L_1 . Por lo tanto, la sucesión $\{\Psi(k|X_n^{\omega_0}|)\}_{n \in \mathbb{N}}$ es uniformemente integrable para cada $k \in \mathbb{N}$. Por otro lado, sean $a > 0$ y $k \in \mathbb{N}$ tal que $k \geq 2a$. Debido a que Ψ es una función de Young,

$$\begin{aligned} 0 \leq \Psi(a|X_n^{\omega_0} - X^{\omega_0}|) &\leq \frac{1}{2} (\Psi(2a|X_n^{\omega_0}|) + \Psi(2a|X^{\omega_0}|)) \\ &\leq \frac{1}{2} (\Psi(k|X_n^{\omega_0}|) + \Psi(k|X^{\omega_0}|)), \end{aligned}$$

de donde $\{\Psi(a|X_n^{\omega_0} - X^{\omega_0}|)\}_{n \in \mathbb{N}}$ es uniformemente integrable, además de que converge a cero casi-seguramente. Por lo tanto, para cada $a > 0$, se tiene que $\mathbb{E}[\Psi(a|X_n^{\omega_0} - X^{\omega_0}|)] \rightarrow 0$. Entonces de la proposición 2.1.10 en [7],

$$\|X_n^{\omega_0} - X^{\omega_0}\|_{\Psi} \longrightarrow 0.$$

Por otro lado, del teorema 4.1 en [2], ρ es continua con respecto a la norma de Luxemburgo $\|\cdot\|_{\Psi}$, por lo que, para $\omega_0 \in \Omega_0$,

$$\hat{\rho}_n(\omega_0) = \mathcal{R}_{\rho}(\hat{m}_n(\omega_0)) = \rho(X_n^{\omega_0}) \rightarrow \rho(X^{\omega_0}) = \rho(X).$$

□

6.0.2. Continuidad de \mathcal{R}_{ρ}

A lo largo de este escrito, hemos investigado el papel que juega la continuidad débil en la caracterización de robustez para el sentido clásico. Es por ello, que necesitamos estudiar las propiedades de continuidad de nuestros estimadores. Sin embargo, la continuidad en este caso no será con respecto a la topología débil usual. Para cada función de carga ψ , se puede definir una topología relacionada. Esta propiedad nos permite hablar de otra clase de robustez: la comparativa.

Una de las motivaciones para considerar conceptos alternativos de robustez, es que varias medidas de riesgo no son robustas en el sentido de Hampel. Un ejemplo clásico es el mapeo $\mathcal{R}_{\rho}(\mu) = -\int x\mu(dx)$, el cual no es continuo con respecto a la topología débil de medidas estándar.

Para probarlo, fijemos $1 > r > 0$ y definamos una sucesión $\{X_n\}_{n \in \mathbb{N}}$ de variables aleatorias independientes con distribución Bernoulli de la

siguiente forma: $\mathbb{P}[X_n = 0] = 1 - \frac{1}{n^r}$ y $\mathbb{P}[X_n = n] = \frac{1}{n^r}$. Denotemos por μ_n a la distribución de X_n para cada n . Entonces, para $\varepsilon > 0$, se tiene que $\mathbb{P}[X_n > \varepsilon] = \frac{1}{n^r}$. De lo anterior, $X_n \rightarrow 0$ en probabilidad y por lo tanto, $\mu_n \rightarrow \delta_0$. Sin embargo, $\mathbb{E}[X_n] = n^{1-r}$ tiende a infinito, por lo cual, el mapeo $\mu \mapsto \mathcal{R}_\rho(\mu)$ no es débilmente continuo en general.

A lo largo de esta sección, consideramos el conjunto \mathcal{M}_1^Ψ de medidas de probabilidad μ sobre \mathbb{R} , tales que $\int \Psi d\mu < \infty$. El espacio C_Ψ estará formado por las funciones continuas f tales que $\sup_{x \in \mathbb{R}} |f(x)/(1 + \Psi(x))| < \infty$. La topología Ψ -débil es la topología más pequeña tal que los mapeos $\mu \mapsto \int f d\mu$, con $f \in C_\Psi$, son continuos. Dentro de los resultados que se tienen para representar la convergencia con respecto a la topología Ψ -débil, está el siguiente. Se trata de una versión del Teorema de representación de Skorohod. En este caso, en lugar de tratar con la convergencia débil, lo haremos con la Ψ -débil.

Teorema 94. *Para cualquier función de Young finita Ψ , son equivalentes:*

1. *Una sucesión $\{\mu_n\}_{n \in \mathbb{N}}$ converge Ψ -débilmente a alguna μ_0 si y sólo si existe una sucesión $\{X_n\}_{n \in \mathbb{N}_0} \subset H^\Psi$ tal que X_n tiene distribución μ_n para $n \in \mathbb{N}_0$ y $\|X_n - X_0\|_\Psi \rightarrow 0$.*
2. *La función Ψ satisface la condición Δ_2 .*

El resultado anterior no será probado, pero puede encontrarse una prueba en [17]. Podemos empezar a observar la importancia de la condición Δ_2 para garantizar algunas propiedades topológicas importantes. De hecho, el siguiente teorema relaciona la continuidad Ψ -débil con que se cumpla la condición Δ_2 , para una función de Young finita.

Teorema 95. *Para una función de Young finita Ψ , las siguientes condiciones son equivalentes:*

1. *Para toda medida de riesgo convexa ρ que es invariante en ley en H^Ψ , el mapeo $\mathcal{R}_\rho : \mathcal{M}(H^\Psi) \rightarrow \mathbb{R}$ es continuo para la topología Ψ -débil.*
2. *La función Ψ satisface la condición Δ_2 .*

Demostración. *Supongamos que Ψ cumple la condición Δ_2 . Sean ρ una medida de riesgo convexa e invariante en ley en H^Ψ y \mathcal{R}_ρ su mapeo correspondiente. Como H^Ψ es un espacio de Banach, de la proposición 3.1 en [24], ρ es continua con respecto a la norma $\|\cdot\|_\Psi$. Tomamos una sucesión $\{\mu_n\}_{n \in \mathbb{N}}$ tal que $\mu_n \rightarrow \mu$ Ψ -débilmente. Por el Teorema 94, existe una sucesión $\{X_n\}_{n \in \mathbb{N}_0} \subset H^\Psi$ tal que X_n tiene distribución μ_n para $n \in \mathbb{N}_0$ y $\|X_n - X_0\|_\Psi \rightarrow 0$. Entonces*

$$\mathcal{R}_\rho(\mu_n) = \rho(X_n) \rightarrow \rho(X_0) = \mathcal{R}_\rho(\mu_0),$$

de donde (2) \Rightarrow (1).

En la prueba de que (1) \Rightarrow (2) se procede a probar la propiedad contrarrecíproca. Es decir, supondremos que Ψ no cumple la condición Δ_2 y construiremos una medida de riesgo ρ para la cual, el mapeo correspondiente \mathcal{R}_ρ no es Ψ -débilmente continuo. Para tal efecto, consideramos la función de pérdida $l(x) = \Psi(0 \vee (8x))$ y a su correspondiente medida de riesgo, la cual está dada por:

$$\rho(X) = \inf\{m \in \mathbb{R} : \mathbb{E}[l(-X - m)] \leq x_1\},$$

donde x_1 es un punto en el interior de la imagen de l . Puede consultarse en [11] que la medida asociada ρ , es una medida de riesgo convexa en H^Ψ . Notemos que $\mathbb{E}[l(-X - m)]$ está bien definida para $m \in \mathbb{R}$ y $X \in H^\Psi$. Debido a que Ψ no cumple la condición Δ_2 y a que el espacio de probabilidad no tiene átomos, por el teorema 2.1.17 en [7], se tiene que $H^\Psi \neq L^\Psi$. Por lo tanto existe una variable aleatoria $Y \geq 0$ tal que $Y \in L^\Psi$ y $\mathbb{E}[\Psi(2Y)] = \infty$. Tomamos una sucesión $\{a_n\}_{n \in \mathbb{N}}$ para la cual, para cada $n \in \mathbb{N}$,

$$4\mathbb{E}[\Psi(2(Y \wedge a_n))] \geq n + l(n/2), \quad (6.3)$$

y definimos

$$X_n = -\left((Y - n)^+ \wedge a_n\right).$$

Es sencillo notar que para cada n , $X_n \in L^\infty$ y por lo tanto $X_n \in H^\Psi$. Además, puesto que Ψ es creciente y del teorema de convergencia dominada, $\mathbb{E}[\Psi(|X_n|)] \leq \mathbb{E}[l(Y - n)] \rightarrow 0$. Entonces por el Lema 99, si cada X_n tiene distribución μ_n , ocurre que $\mu_n \rightarrow \delta_0$ en la topología Ψ -débil. En dado caso de que \mathcal{R}_ρ fuera continua en la topología Ψ -débil, la sucesión $\mathcal{R}_\rho(\mu_n)$ debería ser acotada, lo cual probaremos que es falso por medio de una contradicción.

Denotamos a z_n como sigue: $z_n = \rho(X_n) = \mathcal{R}_\rho(\mu_n)$ y supongamos que $z^* = \sup_n z_n < \infty$. Notemos que, de la convexidad de la función l ,

$$0 \leq l(-X_n - m) \leq \frac{1}{2}l(-2X_n) + \frac{1}{2}l(-2m) \leq \frac{1}{2}\Psi(2|X_n|) + \frac{1}{2}\Psi(2|m|). \quad (6.4)$$

Fijemos $x_0 \in l(\mathbb{R}) < \infty$. Para cada $n \in \mathbb{N}$, existe una sucesión $\{m_k^n\}_{k \in \mathbb{N}}$ tal que $m_k^n \downarrow m_*^n = \rho(X_n)$. Entonces, de (6.4), el teorema de convergencia dominada y la continuidad por la izquierda de l ,

$$\lim_{k \rightarrow \infty} \mathbb{E}[l(-X_n - m_k^n)] = \mathbb{E}[l(-X_n - m_*^n)].$$

Por lo tanto, $\mathbb{E}[l(-X_n - z_n)] \leq x_0$ se cumple para toda $n \in \mathbb{N}$. Así, por la convexidad de l ,

$$\begin{aligned}
x_0 &\geq \mathbb{E}[l(-X_n - z^*)] = \mathbb{E}\left[l\left(\left((Y - n)^+ \wedge a_n\right) - z^*\right)\right] \\
&\geq 2\mathbb{E}\left[l\left(\frac{1}{2}\left((Y - n)^+ \wedge a_n\right)\right)\right] - l(z^*) \\
&\geq 2\mathbb{E}\left[l\left(\frac{1}{2}(Y \wedge (a_n + n)) - \frac{n}{2}\right)\right] - l(z^*) \\
&\geq 4\mathbb{E}\left[l\left(\frac{1}{4}(Y \wedge a_n)\right)\right] - l(n/2) - l(z^*) \\
&= 4\mathbb{E}[\Psi(2(Y \wedge a_n))] - l(n/2) - l(z^*) \\
&\geq n - l(z^*).
\end{aligned}$$

Lo anterior se debe a la definición de la sucesión a_n y muestra claramente la contradicción deseada. Por lo tanto la primera condición implica la segunda. \square

El teorema anterior, hace referencia a una medida de riesgo que está definida en un núcleo de Orlicz H^Ψ . Sin embargo, en la mayoría de los casos contamos con una clase de variables aleatorias mucho más reducida, como L^∞ . Consideremos una medida de riesgo convexa invariante en ley ρ , la cual está definida en L^∞ . En Filipović y Svindland [10], se muestra que existe una única función $\bar{\rho}$, la cual es convexa, invariante en ley y semicontinua por abajo, definida en L^1 y tal que $\bar{\rho}|_{L^\infty} = \rho$. Cabe mencionar que la función $\bar{\rho}$ tiene como rango el intervalo $[-\infty, +\infty]$, por lo cual las propiedades de continuidad del mapeo $\mathcal{R}_{\bar{\rho}}$ se pueden ver afectadas. El siguiente resultado analiza precisamente la relación que existe entre ρ y su extensión.

Teorema 96. *Supongamos que ρ es una medida de riesgo convexa invariante en ley, definida en L^∞ . Asimismo, sea Ψ una función de Young que satisface la condición Δ_2 . Entonces son equivalentes:*

1. La medida $\bar{\rho}$ es finita en H^Ψ .
2. El mapeo $\mathcal{R}_{\bar{\rho}} : \mathcal{M}(H^\Psi) \rightarrow \mathbb{R}$ es continuo para la topología Ψ -débil.
3. El mapeo $\mathcal{R}_\rho : \mathcal{M}(L^\infty) \rightarrow \mathbb{R}$ es continuo para la topología Ψ -débil.
4. Sea $\{X_n\}$ una sucesión en L^∞ . Si $\|X_n\|_\Psi \rightarrow 0$, entonces $\rho(X_n) \rightarrow \rho(0)$.

Demostración. En el caso en que $\bar{\rho}$ sea finita en H^Ψ , se tiene por la Proposición 3.1 de [24], que $\bar{\rho}$ será continua en H^Ψ respecto a la norma $\|\cdot\|_\Psi$. De esta manera, la función $\bar{\rho}$ es de hecho una medida de riesgo convexa invariante en ley. Lo anterior y el Teorema 95, hacen explícita la equivalencia entre las condiciones (1) y (2). Por otro lado, la condición (2) implica la (3) al ser ρ la restricción de $\bar{\rho}$ al espacio L^∞ .

Probaremos que (3) \Rightarrow (4). Sea $\{X_n\} \subset L^\infty$ tal que $\|X_n\|_\Psi \rightarrow 0$. Entonces, si μ_n es la ley de cada X_n , sucede que $\mu_n \rightarrow \delta_0$ en el sentido Ψ -débil. Como \mathcal{R}_ρ es Ψ -débilmente continua, entonces $\rho(X_n) = \mathcal{R}_\rho(\mu_n) \rightarrow \mathcal{R}_\rho(\delta_0) = \rho(0)$. Finalmente mostramos la equivalencia entre las condiciones (4) y (1). En este caso, será suficiente suponer la condición (4) y probar que 0 tiene una vecindad abierta en H^Ψ tal que la función $\bar{\rho}$ es finita en ella. Lo anterior se asegura gracias al Teorema 4.2 de [2]. Consideramos $\{X_n\}$, una sucesión en L^∞ tal que $\|X_n\|_\Psi \rightarrow 0$. Entonces, como se ha visto, $\rho(X_n) \rightarrow \rho(0)$. Ahora bien, si $K > 0$ está fijo, existe $\varepsilon > 0$ tal que $\rho(X) \leq K$, para alguna $X \in B_\varepsilon \cap L^\infty$. Sin embargo, no siempre $X \in L^\infty$, así que sea $Y \in B_\varepsilon$ cualquiera. Se tiene que la parte negativa de Y , denotada como Y^- , pertenece a B_ε y que $Y^- \wedge K \in B_\varepsilon \cap L^\infty$. Además, $\int |Y^- \wedge K| d\mathbb{P} = \int_{Y^- > K} K d\mathbb{P} + \int_{Y^- \leq K} |Y^-| d\mathbb{P} \rightarrow \mathbb{E}[|Y|]$.

La monotonía de $\bar{\rho}$ implica que $\bar{\rho}(X^-) \leq \bar{\rho}(-X^-)$. De la semicontinuidad por abajo de $\bar{\rho}$ en L^1 y el hecho de que

$$\bar{\rho}(-Y^- \wedge K) = \rho(-Y^- \wedge K),$$

se tiene que

$$\bar{\rho}(-Y^-) \leq \liminf_{k \uparrow \infty} \bar{\rho}(-Y^- \wedge k) = \liminf_{k \uparrow \infty} \rho(-Y^- \wedge k) \leq K.$$

Con lo cual $\bar{\rho}$ es finita en B_ε y por lo tanto en H^Ψ . □

6.1. Robustez comparativa

Una vez analizadas las propiedades de continuidad de nuestros estimadores, es momento de discutir el concepto de robustez comparativa. En el capítulo anterior revisamos el Teorema de Hampel, el cual señala la equivalencia entre la continuidad débil y la robustez cualitativa. En este caso, como su nombre lo indica, este nuevo concepto de robustez nos permite comparar estimadores. Inclusive obtenemos un índice, permitiéndonos clasificar diferentes medidas de riesgo con respecto a su robustez.

El concepto clave de esta sección es el de robustez con respecto a la topología Ψ -débil y se muestra a continuación.

Definición 97. Sea ψ una función de carga. Un conjunto $\mathcal{N} \subset \mathcal{M}_1^\psi$ se dice uniformemente ψ -integrante si

$$\lim_{M \rightarrow \infty} \sup_{\nu \in \mathcal{N}} \int_{\psi \geq M} \psi d\nu = 0.$$

Definición 98. Consideramos la métrica de Ψ -Prokhorov.

$$d_\psi(\mu, \nu) = \mathfrak{d}_{\text{prokh}}(\mu, \nu) + \left| \int \psi d\mu - \int \psi d\nu \right|, \quad \mu, \nu \in \mathcal{M}_1^\psi.$$

Sea ψ una función de carga y $\mathcal{M} \subset \mathcal{M}_1^\psi$. Un funcional \mathcal{R}_ρ se dice ψ -robusto en \mathcal{M} cuando \mathcal{R}_ρ es robusto con respecto a d_ψ y $\mathfrak{d}_{\text{prokh}}$ en cada conjunto uniformemente ψ -integrante $\mathcal{N} \subset \mathcal{M}$.

El siguiente es un resultado que relaciona la convergencia ψ -débil de medidas con algunas nociones usuales de convergencia débil.

Lema 99. Son equivalentes:

1. $\mu_n \rightarrow \mu$ en el sentido ψ -débil.
2. $\int f d\mu_n \rightarrow \int f d\mu$ para toda $f \in C_\psi(\mathbb{R})$.
3. $\int f d\mu_n \rightarrow \int f d\mu$ para toda f continua con soporte compacto. Asimismo se cumple si $f = \psi$.
4. $\mu_n \rightarrow \mu$ en el sentido débil usual se cumple que $\int \psi d\mu_n \rightarrow \int \psi d\mu$.

Ahora probaremos un resultado preparatorio. Esta proposición nos indica la velocidad de convergencia al infinito que debe tener una función de carga Ψ , para que la funcional \mathcal{R}_ρ sea Ψ -robusta.

Proposición 100. Sea \mathcal{R}_ρ el funcional de riesgo asociado a una medida de riesgo ρ convexa e invariante en ley en L^∞ . Si $\Psi : \mathbb{R}_+ \rightarrow (0, \infty)$ es una función no decreciente tal que \mathcal{R}_ρ es Ψ -robusta en $\mathcal{M}(L^\infty)$, entonces tiene al menos un crecimiento lineal, es decir:

$$\liminf_{x \uparrow \infty} \frac{\psi(x)}{x} > 0.$$

Demostración. Realizaremos la demostración por contradicción, para lo cual sea ψ una función de carga tal que $\lim_{x \uparrow \infty} \psi(x)/x = 0$ y \mathcal{R}_ρ es ψ -robusta en $\mathcal{M}(L^\infty)$. El Teorema 101 y la consistencia de \mathcal{R}_ρ , dada por el Teorema 93, implican que el mapeo es ψ -débil continuo. La contradicción

se realizará de manera similar al Teorema 95. Sea Y una variable aleatoria no negativa tal que $\mathbb{E}[\psi(Y)] < \infty$ y $\mathbb{E}[Y] = \infty$. Escogemos una sucesión $a_n > 0$ para cada $n \in \mathbb{N}$ tal que $\mathbb{E}[Y \wedge a_n] \geq 2n$. Si definimos $X_n = -(Y - n)^+ \wedge a_n$, con μ_n la ley de X_n para cada n , observamos que $X_n \rightarrow 0$ c.s. Así, dado que $\mathbb{E}[\psi(X_n)] \rightarrow 0$ por el teorema de convergencia dominada, sabemos que $\mu_n \rightarrow \delta_0$, gracias al Lema 99. Por otro lado,

$$\mathbb{E}[X_n] \leq -\mathbb{E}[Y \wedge a_n] + n \leq -n.$$

Entonces, utilizando el Lema 2.3 de [25],

$$\mathcal{R}_\rho(\mu_n) = \rho(X_n) \geq \rho(\mathbb{E}[X_n]) \geq \rho(0) + n,$$

lo cual contradice la continuidad de \mathcal{R}_ρ . □

Antes de dar paso a uno de los resultados principales, enunciamos el recíproco del Teorema de Hampel. La prueba se puede encontrar en [17].

Teorema 101. *Supongamos que $T : \mathcal{M} \rightarrow \mathbb{R}$ es un funcional estadístico donde $\mathcal{M} \subset \mathcal{M}_1^\psi$. Sean $\mu \in \mathcal{M}$ y $\delta_0 > 0$ dados, a la vez que suponemos que T es débilmente consistente para toda $\nu \in \mathcal{M}$ tal que $d_\psi(\nu, \mu) \leq \delta_0$. Entonces, si T es ψ -robusta en μ , implica que T es ψ -débilmente continua en μ .*

Notemos que la definición de ψ -robustez no considera funciones de Young estrictamente. Incluso los resultados de esta sección, hasta ahora, están basados en un a función de carga ψ . Sin embargo, recordemos que si Ψ es una función de Young, son equivalentes el cumplimiento de la condición Δ_2 y la continuidad Ψ -débil, para medidas de riesgo convexas invariantes en ley. En realidad, cuando hablamos de continuidad Ψ -débil, no nos hemos enfocado a las medidas de riesgo en si, sino a sus mapeos de riesgo correspondientes. En el siguiente resultado, relacionamos la robustez de \mathcal{R}_ρ , cuando contamos con una medida de riesgo ρ convexa e invariante en ley, y el cumplimiento de la condición Δ_2 .

Teorema 102. *Para una función de Young finita Ψ , las siguientes condiciones son equivalentes:*

1. *Para toda medida de riesgo ρ convexa e invariante en ley, definida en H^Ψ , el mapeo \mathcal{R}_ρ es Ψ -robusto en $\mathcal{M}(H^\Psi)$.*
2. *Ψ satisface la condición Δ_2 .*

Demostración. Se conoce del Teorema 93 que \mathcal{R}_ρ es consistente para toda $\mu \in \mathcal{M}(H^\Psi)$. Si \mathcal{R}_ρ es robusto, por el Teorema 101, el mapeo \mathcal{R}_ρ es Ψ -débilmente continuo en $\mathcal{M}(H^\Psi)$. Por lo tanto, del Teorema 95, la función de Young Ψ satisface la condición Δ_2 . Ahora bien, del Teorema de Hampel y de nuevo el Teorema 95, se deriva la robustez de \mathcal{R}_ρ , cuando se cumple la condición Δ_2 por Ψ . \square

Las medidas de riesgo que hemos considerado hasta ahora, se han definido en el espacio de Orlicz L^Ψ , para una función de Young Ψ . Sin embargo, las posiciones financieras se consideran como un subconjunto de L^∞ , en algunos casos. Entonces es natural preguntarse cómo se comporta la robustez cuando cambiamos la clase de variables aleatorias a considerar.

Teorema 103. *Sea Ψ una función de Young que satisface la condición Δ_2 . Para una medida de riesgo ρ convexa invariante en ley en L^∞ , sea $\bar{\rho}$ la única extensión de ρ en L^1 que es convexa, monótona y semicontinua por abajo. Entonces las siguientes condiciones son equivalentes:*

1. $\mathcal{R}_{\bar{\rho}}$ es Ψ -robusto en \mathcal{M}_1^Ψ .
2. \mathcal{R}_ρ es Ψ -robusto en $\mathcal{M}(L^\infty)$.
3. $\bar{\rho}$ es finita en H^Ψ .

Demostración. Al ser ρ la restricción de $\bar{\rho}$ al espacio L^∞ , se tiene que la condición (1) implica la (2). En el caso cuando \mathcal{R}_ρ es Ψ -robusta en, de los Teoremas 93 y 101, el mapeo es Ψ -débilmente continuo en $\mathcal{M}(L^\infty)$. Del Teorema 96, $\bar{\rho}$ es finita en H^Ψ . Finalmente, si se tiene (3), se ha discutido que $\bar{\rho}$ es una medida de riesgo convexa invariante en ley. Por lo tanto, del Teorema 103 se obtiene (1). \square

Como hemos visto, la Ψ -robustez de las medidas de riesgo se mantiene aún en L^1 , si la extensión de la medida de riesgo es finita en el núcleo de Orlicz. Ahora que conocemos la caracterización de ψ -robustez, podemos comparar a dos medidas de riesgo, con respecto a su robustez. Cabe recalcar que, al no contar con una función de carga ψ fija, la robustez que consideramos es relativa a dicha función. A diferencia del concepto clásico de Hampel, donde un estimador es robusto, o no lo es, esta concepción es más flexible y nos permite explorar con otro enfoque a nuestros estimadores. En este momento presentamos la definición de robustez comparativa, entre dos medidas de riesgo. De igual manera, se enuncia una consecuencia directa de la definición y nuestro análisis previo en esta sección.

Definición 104 (Robustez comparativa). Sean ρ_1 y ρ_2 dos medidas de riesgo convexas e invariantes en ley en L^∞ . Diremos que ρ_1 es al menos tan robusta como ρ_2 si la siguiente implicación se cumple. Cuando Ψ es una función de Young que satisface la condición Δ_2 y \mathcal{R}_{ρ_2} es Ψ -robusta en $\mathcal{M}(L^\infty)$, entonces \mathcal{R}_{ρ_1} es Ψ -robusta en $\mathcal{M}(L^\infty)$. Si no es así, diremos que ρ_1 es más robusta que ρ_2 .

Corolario 105. Para dos medidas de riesgo ρ_1 y ρ_2 convexas e invariantes en ley definidas en L^∞ , las siguientes proposiciones son equivalentes:

1. ρ_1 es al menos tan robusta como ρ_2 .
2. La función de Young Ψ satisface la condición Δ_2 y $\bar{\rho}_2$ es finita en H^Ψ . Entonces $\bar{\rho}_1$ es también finita en H^Ψ .

Como hemos estado comentando, la robustez comparativa abre una nueva posibilidad de clasificar estimadores. En particular, podemos estudiar a las medidas de riesgo convexas e invariantes en ley de acuerdo a su robustez. Para facilitar el análisis, se define un índice de robustez cualitativa. Este índice nos permite proporcionar una noción de intensidad en la robustez de la medida de riesgo en cuestión. Con ayuda de la robustez comparativa, mostraremos que este índice en verdad estratifica adecuadamente a los estimadores considerados.

Definición 106. Sea ρ una medida de riesgo convexa invariante en ley en L^∞ . El índice de robustez cualitativa asociado a ρ se define como

$$iqr(\rho) = (\inf\{p \in (0, \infty) : \mathcal{R}_\rho \text{ es } \Psi_p\text{-robusto en } \mathcal{M}(L^\infty)\})^{-1}.$$

Conocemos que el crecimiento de la función de Young debe ser a lo menos lineal, para que el estimador de riesgo sea robusto, si la medida de riesgo es convexa. Entonces, para una medida de riesgo ρ convexa e invariante en ley, el índice de robustez cualitativa toma esta forma:

$$iqr(\rho) = (\inf\{p \geq 1 : \bar{\rho}(X) < \infty, X \in L^p\})^{-1}.$$

De hecho, el índice de robustez cualitativa está en el intervalo $[0, 1]$, cuando la medida de riesgo es convexa. Por otro lado, la relación entre los espacios L^p y los espacios de Orlicz, puede parecer clara en un principio. Dado que la función $\Psi(x) = |x|^p$ es de Young para $x \in \mathbb{R}_+$, los espacios de Orlicz se consideran generalizaciones de los espacios L^p . A pesar de ello, cuando la función de Young es fija, existen algunas interacciones interesantes entre los espacios L^p y el espacio de Orlicz en cuestión, sobre

todo con su núcleo H^Ψ . Por ejemplo, podemos preguntarnos cuándo un espacio L^p está contenido en un núcleo de Orlicz dado. De hecho, una cantidad importante a considerar en este análisis es la siguiente:

$$p_\Psi = \inf\{p \geq 1 : L^p \subset H^\Psi\}.$$

Los dos resultados siguientes relacionan la cantidad anterior con el hecho de que $L^p \subset H^\Psi$, o viceversa. Estos lemas serán de utilidad para mostrar la efectividad del índice de robustez cualitativa, en cuanto a la clasificación de medidas de riesgo. Las pruebas de estas propiedades pueden encontrarse en [17].

Lema 107. *Si se cumple que:*

$$\frac{2^p}{\alpha(x_0)} \geq \frac{\Psi(x_0)}{x_0^p},$$

para alguna $x_0 \in \alpha^{-1}(\mathbb{R})$, entonces $L^p \subset H^\Psi$.

Lema 108. *Si $p \in [1, p_\Psi)$, entonces $H^\Psi \subset L^p$.*

Ahora presentamos el resultado que muestra la conveniencia del índice de robustez cualitativa. Observamos que es posible clasificar a las medidas de riesgo convexas, de acuerdo a su robustez cualitativa.

Teorema 109. *Si ρ_1 y ρ_2 son dos medidas de riesgo convexas invariantes en ley tales que $iqr(\rho_1) > iqr(\rho_2)$, entonces ρ_1 es más robusta que ρ_2 .*

Demostración. *Debido a que $iqr(\rho_1) > iqr(\rho_2)$, se tiene que existe $r > 0$ tal que $\bar{\rho}_1$ es finita en L^r , pero $\bar{\rho}_2$ no lo es. Por tanto, solamente resta mostrar que ρ_1 es al menos tan robusta que ρ_2 . Sea Ψ cualquier función de Young finita que cumple la condición Δ_2 , para la cual $\bar{\rho}_2$ es finita en H^Ψ . Por el Lema 107, $L^p \subset H^\Psi$ para $p > p_\Psi$ y $H^\Psi \subset L^q$ para $q \in [1, p_\Psi)$ por el Lema 108. Por hipótesis y de la definición de $iqr(\rho_2)$,*

$$iqr(\rho_1) > iqr(\rho_2) \geq \frac{1}{p_\Psi},$$

de lo cual $\bar{\rho}_1$ es finita en L^q para alguna $q \in [1, p_\Psi)$. Sin embargo, como $H^\Psi \subset L^q$, se tiene que $\bar{\rho}_1$ es finita en H^Ψ . Por el Corolario 105, ρ_1 es al menos tan robusta como ρ_2 . \square

6.1.1. Índice de robustez cualitativa para distorsiones cóncavas de medidas de riesgo

Hemos estudiado el índice de robustez cualitativa para las medidas de riesgo convexas. En este apartado mostraremos otra representación de dicho índice, para el caso de medidas de riesgo producto de distorsiones cóncavas. Asimismo, se mostrará que esta clase de medidas de riesgo cubren el espectro $(0, 1]$, en cuanto al índice de robustez cualitativa. Uno de los teoremas principales en el análisis funcional, es el Teorema de Banach-Steinhaus. El teorema establece que para un conjunto de operadores lineales continuos, en un espacio de Banach, la finitud de la norma de cada operador evaluado en todos sus puntos, es equivalente a la finitud de la norma de los operadores. Este resultado se utilizará para obtener una representación del índice de robustez cualitativa, por lo cual se enuncia a continuación. Una demostración del mismo puede ser hallada en Conway, [5].

Teorema 110 (Teorema de Banach-Steinhaus). *Sea \mathcal{X} un espacio de Banach y \mathcal{Y} un espacio normado. Sea $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ el conjunto de operadores lineales y acotados, con dominio \mathcal{X} y rango \mathcal{Y} . Si $\mathcal{A} \subset \mathcal{B}(\mathcal{X}, \mathcal{Y})$ es tal que, para cada $x \in \mathcal{X}$, se satisface que $\sup\{\|Ax\| : A \in \mathcal{A}\} < \infty$, entonces $\sup\{\|A\| : A \in \mathcal{A}\} < \infty$.*

Recordemos del Teorema 24, que las medidas de riesgo derivadas de distorsiones cóncavas, para las medidas de probabilidad definidas en $[0, 1]$, tienen la siguiente representación.

$$\rho_g(X) = g(0+) \text{ ess sup } (-X) + \int_0^1 VaR_t(X) g'_+(t) dt, \quad X \in L^\infty. \quad (6.5)$$

Por último, mostramos la forma particular del índice de robustez cualitativa que se obtiene al considerar medidas de riesgo del tipo de (6.5).

Proposición 111. *Supongamos que g es cóncava y continua. Entonces para $p \in [1, \infty)$ y q tal que $\frac{1}{p} + \frac{1}{q} = 1$, la extensión $\bar{\rho}_g$ de ρ_g es finita en L^p si y sólo si $g'_+ \in L^q((0, 1))$. En particular,*

$$iqr(\rho_g) = \frac{q^* - 1}{q^*}, \quad \text{donde } q^* = \sup \left\{ q \geq 1 \mid \int_0^1 (g'_+(t))^q dt < \infty \right\}.$$

En particular, si $q^* = \infty$, consideramos que $iqr(\rho_g) = 1$.

Demostración. Debido a que g es continua, se observa que

$$\rho_g(X) = \int_0^1 \text{VaR}_t(X) g'_+(t) dt,$$

en (6.5). De hecho, la extensión $\bar{\rho}_g$ cumple (6.5) cuando $X \in L^1$. Ahora, notemos de la definición de VaR_t ,

$$\begin{aligned} \int_0^1 \text{VaR}_t(X) g'_+(t) dt &= \int_0^1 q_{-X}(1-t) g'_+(t) dt \\ &= \int_0^1 q_{-X}(t) f(t) dt, \end{aligned}$$

donde $f(t) = g'_+(1-t)$. Supongamos que $g'_+ \in L^q(0,1)$. Del Lema A.23 en [11], se tiene lo siguiente:

$$\begin{aligned} \int_0^1 |q_{-X}(t)|^p dt &= \int_0^{F(0)} |-q_{-X}(t)|^p dt + \int_{F(0)}^1 |q_{-X}(t)|^p dt \\ &= \int_0^{F(0)} q_{|X|^p}(t) dt + \int_{F(0)}^1 q_{|X|^p}(t) dt \\ &= \mathbb{E}[|X|^p]. \end{aligned} \tag{6.6}$$

Entonces, de la desigualdad de Hölder y (6.6), para toda $X \in L^p$,

$$\begin{aligned} \bar{\rho}_g(X) &= \int_0^1 q_{-X}(t) f(t) dt \leq \left(\int_0^1 |q_{-X}(t)|^p dt \right)^{1/p} \left(\int_0^1 f(t)^q dt \right)^{1/q} \\ &= \mathbb{E}[|X|^p]^{1/p} \|f\|_{L^q} < \infty. \end{aligned}$$

De lo cual, $\bar{\rho}_g$ es finita. Ahora supongamos que $\bar{\rho}_g < \infty$. Debido a que el espacio de probabilidad en cuestión no tiene átomos, existe una variable aleatoria uniforme U en $(0,1)$, definida en dicho espacio. Para probar que $g'_+ \in L^q$, se mostrará que $Y = f(U) \in L^q$.

Lo anterior es equivalente a probar que $\mathbb{E}[(-X)Y] < \infty$ para toda $X \in L^p$. Si $Y \in L^q$, entonces, de la desigualdad de Hölder, se cumple la desigualdad deseada. Por otro lado, si se cumple que $\mathbb{E}[(-X)Y] < \infty$ para toda $X \in L^p$, el Teorema de Banach-Steinhaus, tomando el operador $F : L^p \rightarrow \mathbb{R}$ como $F(X) = \mathbb{E}[(-X)Y]$, implica que

$$\sup_{\|X\|_{L^p} \leq 1} |\mathbb{E}[XY]| < \infty, \quad X \in L^p. \tag{6.7}$$

Es sencillo notar que el operador F es lineal, de ahí que (6.7) implique que F sea continuo. Más aún, tomamos en cuenta que $(L^p)^* = L^q$, donde $(L^p)^*$

denota el dual analítico de L^p , con identificación

$Z \in L^q \mapsto T_Z(X) = \mathbb{E}[(-X)Z]$. Así, dado que $F \in (L^p)^*$, existe $Z_0 \in L^q$ tal que $F(X) = \mathbb{E}[(-X)Z_0]$. Como $\mathbb{E}[(-X)Z_0] = \mathbb{E}[(-X)Y] < \infty$ para toda $X \in L^p$, entonces $Z_0 = Y$ c.s. Por lo tanto $Y \in L^q$.

Lo anterior pone de manifiesto que es suficiente con probar que $\mathbb{E}[(-X)Y] < \infty$ para toda $X \in L^p$. Debido a la distribución de U , se observa que f es una función cuantil para Y . Entonces, de la desigualdad de Hardy – Littlewood,

$$\mathbb{E}[(-X)Y] \leq \int_0^1 q_{-X}(t)q_Y(t)dt = \int_0^1 q_{-X}(t)f(t)dt = \rho(X) < \infty,$$

obteniendo el resultado deseado. □

Se muestra en [11], que $CVaR_\lambda$ puede ser representado por medio de la función de distorsión cóncava $g(t) = (t/\lambda) \wedge 1$. Debido a que g' es acotada, se sigue del resultado anterior que $iqr(CVaR_\lambda) = 1$. De hecho, si consideramos $\beta \in (0, 1]$ y la función de distorsión cóncava $g_\beta(t) = (t/\lambda) \wedge 1$, se cumple que $iqr(\rho_{g_\beta}) = \beta$. Observamos entonces que las medidas de riesgo producto de distorsiones cóncavas pueden expresar casi cualquier magnitud posible, en el índice de robustez cualitativa.

Capítulo 7

Conclusiones

En primer lugar, hemos hecho una revisión exhaustiva de las definiciones y resultados principales sobre medidas de riesgo monetario convexas, y su representación. En el Capítulo 2, se han definido las medidas Valor en Riesgo (VaR) y Valor en Riesgo Condicionado (CVaR), los cuales son conceptos fundamentales a lo largo del trabajo.

El estudio de las medidas de riesgo convexas, comonótonas y de distorsiones cóncavas también ha sido realizado. Como resultado, hemos obtenido un teorema de representación para medidas de riesgo convexas que son comonótonas. La importancia de este teorema radica en que cualquier medida de riesgo convexa comonótona, puede ser vista como una mezcla de CVaR y alguna medida de probabilidad en $[0, 1]$.

Seguidamente, hemos hecho una revisión de las distribuciones α -estables, concentrándonos en algunas propiedades de su función característica. El resultado principal del Capítulo 3, es el teorema de representación integral de la función de densidad, para estas leyes. Hemos utilizado la integración compleja para lograr dicho objetivo. Esta representación es relevante, debido a que es fácilmente manejable numéricamente.

Gracias al resultado principal del Capítulo 3, es posible hallar una representación integral para la medida de riesgo CVaR, en el caso de leyes estables. El uso de este tipo de distribuciones, para modelar los rendimientos de una posición, representa varias ventajas. La primera es que captura el efecto de las colas pesadas en la distribución. Esto nos permite, a través de CVaR, que nuestras mediciones de riesgo sean sensibles a la probabilidad de grandes ganancias o pérdidas.

Otra ventaja es que las distribuciones estables cumplen con propiedades del tipo del Teorema del Límite Central, a diferencia de otras leyes con

colas pesadas. Dentro de esta clase de variables aleatorias, podemos encontrar algunas muy familiares, como las de tipo exponencial, incluyendo la gaussiana y la de Cauchy. Sin embargo, hay otras varias distribuciones que tienen colas pesadas y se han propuesto para modelar rendimientos en finanzas. A diferencia de tales alternativas, las distribuciones estables cumplen con la propiedad de autosimilitud. Desde los trabajos de Mandelbrot, la propiedad fractal, o autosimilitud, ha sido motivo de estudio en finanzas. Algunos trabajos con datos empíricos, han ratificado las ideas de Mandelbrot. Entonces, las leyes estables se vuelven muy atractivas para modelar el rendimiento de una posición. De igual manera, el hecho de tener una expresión cerrada y manejable numéricamente, para el CVaR de estas leyes, en el Capítulo 4, es destacable.

A pesar de que los resultados, para el cálculo de CVaR de leyes estables, se tienen desde 2006, aparecen algunas interrogantes acerca de su uso práctico. Uno de los principales cuestionamientos que se le hacen a CVaR, es que no es robusto. La concepción clásica de robustez fue acuñada por Hampel en 1971. Este concepto se basa en la idea de que si la distribución estimada, es ligeramente diferente a la que origina nuestra muestra, entonces los estimadores no deberían variar demasiado.

En el Capítulo 5, realizamos una revisión de una versión actualizada del resultado clásico de Hampel. Cabe destacar que esta versión es más general, además que utiliza argumentos más recientes de la teoría probabilista. Este análisis, además de situarnos en el contexto de robustez, es conveniente porque da la pauta al estudio realizado en el capítulo final. En el Capítulo 6, trabajamos una nueva noción de robustez: la robustez comparativa. En el caso del resultado de Hampel, un estimador podía catalogarse como robusto o no robusto. Bajo esta nueva concepción, los estimadores tienen un grado de robustez, dependiendo del espacio de variables aleatorias donde estén definidas. Como se ha mencionado, el CVaR no es robusto en el sentido de Hampel.

Sin embargo, bajo este nuevo paradigma, CVaR vuelve a cobrar relevancia en este rubro, al tener cierto grado de robustez. El capítulo concluye con el establecimiento de un índice, el cual estratifica a las medidas de riesgo por su robustez. Se encuentran expresiones para dicho índice, en el caso de las medidas de riesgo convexas y las de distorsiones cóncavas.

En conclusión, este trabajo sirve como una referencia de algunos aspectos de medidas de riesgo monetario. Hemos enfocado el análisis en el CVaR y en la relevancia de las leyes α -estables en este contexto. Bajo la nueva categorización de robustez, encontramos que CVaR sigue siendo una

medida de riesgo conveniente para la modelación de rendimientos en finanzas. Contar con una expresión fácilmente calculable de CVaR, para una familia tan importante de distribuciones como las α -estables, es de gran relevancia. Como un proyecto a futuro, se pueden investigar algunas otras leyes utilizadas para modelar los rendimientos de un portafolio. Lo anterior con el fin de conocer, estimar o aproximar su cercanía a alguna ley estable e investigar la conveniencia de la robustez de CVaR. Este y otros aspectos pueden ser estudiados en oportunidades posteriores.

Apéndice A

Implementación de CVaR

Para ilustrar la aplicabilidad del Teorema de Representación 61, desarrollamos un programa en el software R, que implementa el cálculo de riesgo CVaR para distribuciones α -estables. Para este fin, será crucial el cálculo de funciones cuantiles de la distribución estable y haremos un uso intensivo de la paquetería “stabledist” desarrollada por Nolan. Una exposición sistemática de algoritmos de simulación para variables y procesos estocásticos estables se puede encontrar en Janicki y Weron [16].

```
#Función para el cálculo de VaR de nivel "e"
var1 = function(e,a,b){stabledist::qstable(p=e, alpha=a, gamma=1,
beta=b, delta = 0,pm=1)}

#Funciones beta y theta cero
bg = function(var,b){-sign(var)*b}
t0 = function(a,bg){(1/a)*atan(bg*tan((a*pi)/2))}

#Para la funcion vt (v de theta) necesitamos la funcion auxiliar h
h = function(a,t0,t){cos(a*t0 + (a-1)*t)/(cos(t))}

vt = function(a,t0,t){(cos(a*t0))^(1/(a-1)) *
(cos(t)/(sin(a*(t0+t))))^(a/(a-1)) *h(a,t0,t)}

g = function(a,t0,t){sin(a*(t0+t)-2*t)/sin(a*(t0+t))
- (a*(cos(t))^2)/((sin(a*(t0+t)))^2)}

vta = function(a,t0,t){(cos(a*t0))^(1/(a-1)) *
(cos(t)/(sin(a*(t0+t))))^(a/(a-1))}
```

```

#Definición de la función CVaR
cvar = function(e,a,b){
v = -var1(e,a,b)
bgo = bg(v,b)
tt0 = t0(a,bgo)
vtt = function(t){vt(a,tt0,t)}
gt = function(t){g(a,tt0,t)}
inter = function(t){gt(t)*exp(-(abs(v))^(a/(a-1))*vtt(t))}

cv = (a/(1-a))*(abs(v)/(e*pi))*integrate(inter,-tt0,pi/2)[1]$value
return(cv)
}

#Obtención de CVaR de nivel "e"
e = .1
x = seq(1.04,2,by=.03)
y = c(-.2,-.1,-.05,0,.05,.1,.2)

f = matrix(0,1,length(x)*length(y))
w = matrix(0,length(x)*length(y),3)

i=1
z=1
while(i<=length(x))
{
j = 1
while(j<=length(y))
{
f[z] = cvar(e,x[i],y[j])
w[z,1] = x[i]
w[z,2] = y[j]
w[z,3] = f[z]
z = z+1
j = j+1
}
i = i+1
}

persp(w, phi=20, theta=45)

```

Bibliografía

- [1] L. Breiman, *Probability*, Classics in Applied Mathematics, vol. 7, SIAM, Philadelphia, PA., 1991.
- [2] P. Cheridito and T. Li, *Risk Measures On Orlicz Hearts*, *Mathematical Finance* **19** (2009), no. 2, 189–214.
- [3] A.S. Chernobai, S.T. Rachev, and F.J. Fabozzi, *Operational risk: a guide to Basel II capital requirements, models, and analysis*, Frank J. Fabozzi series, Wiley, 2007.
- [4] R. Cont, R. Deguest, and G. Scandolo, *Robustness and sensitivity analysis of risk measurement procedures*, *Quantitative Finance* **10** (2010), no. 6, 593–606.
- [5] J.B. Conway, *A course in functional analysis*, Graduate Texts in Mathematics, Springer, 1990.
- [6] N. Dunford and J. Schwartz, *Linear Operators. Part I: General Theory*, Interscience Publishers, New York, 1958.
- [7] G.A. Edgar and L. Sucheston, *Stopping times and directed processes*, Cambridge University Press, Cambridge, 1992.
- [8] R. F. Engle and S. Manganelli, *CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles*, *Journal of Business and Economic Statistics* **22** (2004), no. 4, 367–381.
- [9] Carl J. G. Evertsz, *Self-similarity of high-frequency USD-DEM exchange rates*, 1995.
- [10] D. Filipović and G. Svindland, *The Canonical Model Space for Law Invariant Convex Risk Measures is L^1* , *Mathematical Finance* **22** (2012), no. 3, 585–589.

- [11] H. Föllmer and A. Schied, *Stochastic Finance. An introduction in discrete time*, third edition ed., de Gruyter, Berlin, 2011.
- [12] Markus Haas and Christian Pigorsch, *Financial economics, fat-tailed distributions*, Complex Systems in Finance and Econometrics (Robert A. Meyers, ed.), Springer New York, 2011, pp. 308–339.
- [13] F.R. Hampel, *A general qualitative definition of robustness*, Annals of Mathematical Statistics **42** (1971), 1887–1896.
- [14] P. Huber, *Robust statistics*, Wiley Series in Probability and Mathematical Statistics, 1974.
- [15] J. Jacod and P. Protter, *Probability essentials*, 2nd ed., Springer-Verlag, 2002.
- [16] A. Janicki and A. Weron, *Simulation and chaotic behavior of α -stable stochastic processes*, Marcel Dekker, New York, 1994.
- [17] V. Krättschmer, A. Schied, and H. Zähle, *Comparative and qualitative robustness for law-invariant risk measures*, ArXiv e-prints (2012).
- [18] V. Krättschmer, A. Schied, and H. Zähle, *Qualitative and infinitesimal robustness of tail-dependent statistical functionals*, Journal of Multivariate Analysis **103** (2012), 35–47.
- [19] B. Mandelbrot, *The variation of certain speculative prices*, The Journal of Business **36** (1963), 394–419.
- [20] ———, *Self-similar random process and the range*, IBM research [paper], 1964.
- [21] John P. Nolan, *Numerical calculation of stable densities and distribution functions*, Communications in Statistics. Stochastic Models **13** (1997), no. 4, 759–774.
- [22] Basel Committee on Banking Supervision, *Basel II: International Convergence of Capital Measurement and Capital Standards. A Revised Framework Comprehensive Version*, Bank for International Settlements, 2006.
- [23] K. R. Parthasarathy, *Probability measures on metric spaces*, Probability and Mathematical Statistics, vol. 3, Academic Press Inc., New York-London, 1967.

- [24] A. Ruszczyński and A. Shapiro, *Optimization of convex risk functions*, Mathematics of Operations Research **31** (2006), no. 3, 433–452.
- [25] A. Schied, *On the Neyman-Pearson problem for law-invariant risk measures and robust utility functionals*, Quantitative finance papers, arXiv.org, July 2004.
- [26] E. M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.
- [27] S. Stoyanov, G. Samorodnitsky, S.T. Rachev, and S. Ortobelli, *Computing the portfolio conditional value-at-risk in the α -stable case*, Probability and Mathematical Statistics **26** (2006), 1–22.
- [28] V. Strassen, *The existence of probability measures with given marginals*, Annals of Mathematical Statistics **36** (1965), no. 2, 423–439.
- [29] Zolotarev, V. M., *One-dimensional stable distributions*, Amer. Math. Soc. Transl. of Math. Monographs, vol. 65, Providence, RI. (Transl. of the original 1983 Russian), 1986.