



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

GENERACIÓN DE BENCHMARK DE FONDOS PARA EL SISTEMA DE PENSIONES
EN CHILE, UN ENFOQUE BASADO EN OPTIMIZACIÓN ESTOCÁSTICA

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE
OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL
INDUSTRIAL

GIORGIOGIULIO PARRA DE BLASI

PROFESOR GUÍA:
DANIEL ESPINOZA GONZÁLEZ

PROFESOR CO-GUÍA:
NICOLAS FIGUEROA GONZÁLEZ

MIEMBROS DE LA COMISIÓN:
ROBERTO COMINETTI COTTI-COMETTI
FERNANDO ORDÓÑEZ PIZARRO

Este trabajo ha sido parcialmente financiado por CONICYT

SANTIAGO DE CHILE
2017

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE: Ingeniero Civil Industrial y grado de
Magíster en Gestión de Operaciones
POR: Giorgiogiulio Parra De Blasi
FECHA: 25/01/2017
PROF. GUÍA: Daniel Espinoza González

GENERACIÓN DE BENCHMARK DE FONDOS PARA EL SISTEMA DE PENSIONES EN CHILE, UN ENFOQUE BASADO EN OPTIMIZACIÓN ESTOCÁSTICA

En Chile el sistema de pensiones, es un mecanismo de previsión y protección social, que a través del ahorro e inversión, genera el capital para la futura pensión de vejez o potencial pensión de invalidez. Adicionalmente, el sistema ha generado externalidades positivas, por sus regulaciones en composición de inversión, lo que se traduce en un gran volumen de capital que ha dado liquidez al mercado, y en parte, ha facilitado el desarrollo económico. En este sistema, al igual que en todo problema de administración activa de inversiones, es clave el *benchmark* de desempeño, en la actualidad, por regulación el único índice observado es el promedio mensual de los demás participantes en el periodo anterior, y su acumulado, generando incentivos a comportamientos de manada, y no garantizando una gestión activa riesgo-retorno.

En el presente trabajo, utilizando optimización con aversión al riesgo, se propone una metodología de referencia, esta se complementaría con la actual, permitiendo evaluar el desempeño, con un enfoque de optimización de riesgo-retorno enfrentando iguales restricciones, información y condiciones de mercado, cuantificando la habilidad de los agentes, y no condiciones erráticas de mercado, o su habilidad de imitación. Metodológicamente se exploran dos familias de medidas: las *medidas espectrales*, con sus representantes: Valor Esperado, *Conditional Value at Risk* (CVaR) y una combinación convexa de ambas; y la medida Entrópica. CVaR se incorpora como generalización convexa del conocido en la industria *Value at Risk* (VaR). La resolución del problema de optimización que implementa la Medida de Riesgo Entrópica, constituye el principal aporte metodológico. Para su resolución, al igual que para las demás medidas, se empleó la metodología de resolución de problemas estocásticos *Sample Average Approximation* (SAA). En la resolución de la medida Entrópica fue necesario desarrollar una metodología de aproximación y ajuste de cortes locales, como aproximación lineal a la vecindad óptima, que permitió resolver en instancias de prueba, con un *gap* estocástico relativo menor al 0.1 %.

Se realizaron diversos experimentos computacionales evaluando: convergencia y tiempos de resolución, efecto de la correlación entre activos, distribuciones con colas gordas simétricas y asimétricas, y la comparativa frente al clásico Markowitz de 1952 [58]. Adicionalmente, con datos diarios históricos del fondo A, desde el año 2003 al 2013, se resuelve a modo de referencia. Se concluyó que ante instancias simétricas, las soluciones generadas por las diferentes medidas coinciden para algún nivel de aversión. En particular, la noción clásica de dispersión es suficiente. No obstante, ante una instancia asimétrica las diferentes nociones de riesgo se diferencian.

Con mucho amor a mi Familia
Para mi Mamá, Papá, Hermana, Panchita, Abuelitos, Maestros y Amigos
Son mi fuente de energía en la vida

Agradecimientos

Agradecer a mis Padres, Giulia y Jorge, durante toda mi vida me han enseñado, inspirado y apoyado con amor, a cada día ser una mejor persona, soy lo que soy gracias a ustedes, me dieron el mayor de todos los regalos, la vida. A mis abuelitos, que aunque se ya no están aquí, su memoria, enseñanzas, deseos y palabras viven con nosotros.

A mi Hermana, Julianna, por ser siempre mi protectora y apoyo en infancia, compañera en la adolescencia, y ser mi partner y tercer pilar. Desde pequeño siempre has creído en mí, y me has apoyado en mis aventuras e inventos. Eres mi hermanita.

A mi compañera, Panchita, te conocí ya hace nueve años, me cambiaste la vida, hemos crecido juntos, aprendido el uno del otro, y juntos de la vida, me das paz, esperanza y felicidad. Eres mi mejor amiga, confidente, socia y amante. Ya hablamos sin palabras y en los días difíciles, eres como el amanecer.

A mis Maestros y Profesores, tanto de Colegio como Universitarios, sin ellos en el rol motivador, docente y guía, no sería lo que soy. En especial agradecer al Profesor Hernán Quezada, un gran amigo y guía durante el Colegio. A mi comisión por su compromiso. Al Profesor Daniel Epinoza, que desde el primer día siempre me pidió más, me enseñó a no temer, trabajar duro y ser humilde, hoy mirando en retrospectiva lo que he aprendido, desde donde partimos es impresionante, y siempre se ha dado el tiempo, sin usted este trabajo no habría sido posible. Al Profesor Roberto Comminetti, por ser un gran amigo e inspiración, tanto moral como intelectual, desde el primer día que lo conocí, hasta en esta última etapa, alentándome a cerrar el proceso. Al Profesor Richard Weber, por su constante apoyo, y presión a ser mejor.

A mis amigos tanto de Colegio como Universidad, Jorge Godoy, Camilo Rojas, Richard Weber, Eduardo Contreras, Cristian Bravo, José Pizarro, Sebastián Orozco, Marcela Valenzuela, Rodrigo Orellana, Ignacio Ríos, Alexis Pérez, Tito Rivera, Claudio Calderon, y cuantos otros. Por apoyarme, motivarme, ayudarme, creer en mí y consolarme, enseñarme de la vida y estar ahí en las buenas y en las malas. A mis alumnos, Yanxi, Alexis, Andrés y Maricel, me han enseñado y motivado a cada día ser mejor, semillas plantadas en suelo fértil.

Al Centro de Finanzas, en especial al Profesor José Miguel, Soledad Vergara, Soledad Hevia, Carolina, Marcela, Alejandro y Pato, son una gran familia, me han dado la paz y tranquilidad para crecer durante todos estos años. A Fernanda Melis y Linda Valdés, por apoyarme durante el proceso del Master, por ser quienes siempre tienen una solución.

A CONICYT por apoyarme en este gran proyecto. A mi querido Chile, que motiva día a día a ser un mejor país, lleno de gente hermosa, geografía y paraísos naturales que cuidar.

Agradecer sin cometer el error de dejar a alguien fuera, siempre es difícil. La vida es una aventura, en la que conoces una infinidad de almas, cada una de ellas, siempre te marca de alguna forma, y todas esas formas siempre se agradecen, dado que llevan a lo que uno es y será. A todos los que saben que deben estar aquí, muchas gracias.

Tabla de contenido

1. Introducción	1
1.1. Antecedentes y Motivación	1
1.2. Sobre esta tesis	3
1.3. ¿Por qué un sistema de pensiones?	5
1.4. Revisión bibliográfica y estado del arte	6
1.4.1. Teoría Moderna de Portafolios	6
1.4.2. Optimización Estocástica y las Medidas de Riesgo	8
2. Las medidas de riesgo en el problema de portafolio	10
2.1. Teoría moderna de portafolios	11
2.2. Las medidas de riesgo	13
2.2.1. Medida neutra al riesgo: valor esperado	16
2.2.2. Medidas espectrales: CVaR y ECVaR	16
2.2.3. Medida de riesgo entrópica	18
2.3. El problema de portafolio	19
2.3.1. Portafolio de maximización del valor esperado	19
2.3.2. Portafolio de minimización del Conditional Value at Risk (CVaR)	20
2.3.3. Portafolio de minimización de CVaR, con retorno mínimo	20
2.3.4. Portafolio de minimización del riesgo entrópico	21

2.4. Metodología de resolución: Sample Average Approximation (SAA)	21
3. Modelos y metodologías desarrolladas	28
3.1. SAA aplicado a los problemas de portafolio	28
3.1.1. Aproximación del problema de portafolio con valor esperado	29
3.1.2. Aproximación del problema de portafolio con CVaR	29
3.1.3. Aproximación del problema de portafolio con ECVaR	29
3.2. Aproximación del problema de portafolio con la Medida Entrópica	29
3.2.1. Aproximación vía subgradiente del problema de optimización entrópico	30
3.2.2. Algoritmo de generación y ajuste de cortes de aproximación	33
4. Experimentos computacionales	41
4.1. Condiciones generales de la implementación	41
4.2. Visión general de los experimentos realizados	42
4.3. Experimentos monotemporales	45
4.3.1. Análisis transversal de convergencia y tiempos de ejecución	45
4.3.2. Efecto de la correlación entre instrumentos	52
4.3.3. Efecto de colas gordas	58
4.3.4. Las Medidas de Riesgo y el enfoque clásico de Markowitz	67
4.4. Experimento Multitemporal	76
4.4.1. Diseño experimental	79
4.4.2. Resultados y análisis	81
5. Conclusiones y trabajo futuro	86
5.1. Conclusiones	86
5.2. Trabajos futuros	89

Bibliografía	91
6. Apéndices	97
6.1. Resumen extendido	97
6.2. Monotonía de la medida de riesgo entrópica a la aversión al riesgo	100
6.3. Consistencia intertemporal de las medidas de riesgo	101
6.4. Intento de formulación espectral de la medida entrópica	102
6.4.1. Formulación LP de las medidas de riesgo entrópicas	102

Índice de tablas

4.1. Tabla resumen de las distribuciones de los portafolios óptimos distintos, en los experimentos de correlación.	58
4.2. Tabla resumen de los principales estadísticos de los instrumentos utilizados en cola gorda asimétrica	61
4.3. Tabla resumen de los principales estadísticos de los instrumentos utilizados en cola gorda simétrica	62
4.4. Tabla resumen de los principales estadísticos de los portafolios óptimos en la instancia de colas gordas asimétricas	64
4.5. Tabla resumen de los principales estadísticos de los portafolios óptimos en la instancia de colas gordas simétricas	66
4.6. Tabla resumen de los principales estadísticos de los portafolios óptimos extremos . .	75
4.7. Estadística descriptiva de los retornos logarítmicos de los instrumentos	80
4.8. Tabla resumen de estadísticos descriptivos, del proceso de retornos periodo a periodo	85

Índice de ilustraciones

1.1. Ejemplo de la incertidumbre en los tiempos de viajes y la optimización con aversión al riesgo, en el caso de un sistema de despacho de ambulancias.	2
2.1. Ilustración que muestra la cercanía entre $\log(1+x)$ y x entorno a cero, lo que permite la interpretación del retorno logarítmico como una “buena aproximación” del retorno aritmético entorno al cero.	11
2.2. Ilustración del VaR (Value at Risk) y del CVaR (Conditional Value at Risk), en una distribución.	15
3.1. Ilustración de aproximación de exponencial vía cortes	31
3.2. Función de error de aproximación, para un cuarto punto, ya fijados tres anteriores de forma óptima	35
3.3. Puntos de soporte obtenidos por ajuste (triángulos) y obtenidos numéricamente (círculos); para diferente número de soportes	36
3.4. Error total de los puntos iniciales, medido como el área entre las funciones, dependiendo del número de puntos utilizados para aproximar	37
3.5. Ejemplo de función lineal por partes, que aproxima la exponencial, con su respectivo error de aproximación asociado.	38
3.6. Boxplot con el error porcentual de aproximación de la función exponencial, evaluada en los puntos $\{u_i\}_{i \in [N]}$ con 10.000 realizaciones, considerando una instancia de prueba y 100 repeticiones iid.	39
3.7. Diagrama resumen del algoritmo de aproximación, método de ajuste de los puntos de soportes y resolución.	40
4.1. Gap relativo sin (izq.) y con (der.) normalización para las diferentes configuraciones, con los parámetros escogidos para SAA	47

4.2. Valor de la constante de traslación L en $\rho(\bar{w} \cdot (\bar{x} + L))$, para las diferentes configuraciones, ordenadas crecientemente en el eje x por gap relativo no trasladado.	47
4.3. Gap relativo trasladado v.s. ordenamiento creciente por tiempo de las configuraciones (izq.), Tiempo promedio por repeticiones v.s. ordenamiento creciente de configuraciones por gap relativo trasladado (der.)	48
4.4. Evolución del gap relativo sin traslación (izq.) y con traslación (der.) versus el número de repeticiones, para las diferentes configuraciones.	49
4.5. Cambio y magnitud de los mismos en la cota superior, al considerar una repetición adicional	50
4.6. Evolución del gap relativo no normalizado (izq.) y normalizado (der.) versus el tamaño <i>out-sample</i>	51
4.7. Evolución del tiempo promedio de cálculo de la cota superior, versus diferentes tamaños <i>out-samples</i>	52
4.8. Evolución del gap relativo no trasladado (izq.) y trasladado (der.), versus el tamaño <i>in-sample</i> para las diferentes configuraciones	53
4.9. Tiempo promedio por repetición, versus el tamaño <i>in-sample</i>	53
4.10. Simulación con dos variables normales estándar, modificando la correlación para portafolios equi-distribuidos	55
4.11. Densidades de probabilidad de los retornos, para los portafolios óptimos de las diferentes combinaciones medida-aversión, en la instancia con (izq.) y sin correlación (der.)	57
4.12. Perfiles espectrales de los diferentes portafolios óptimos en las instancias con y sin correlación.	58
4.13. Perfiles entrópicos de los diferentes portafolios óptimos en las instancias con y sin correlación.	59
4.14. Ejemplo de la distribución mezcla de tres normales $\sigma = 2$ y $(\mu_1, \mu_2, \mu_3) = (5, 10, 15)$.	60
4.15. Ilustración de instancia con sólo dos estados: normal (n) y crisis (c)	60
4.16. Densidades de los diferentes instrumentos utilizados en la instancia de colas gordas asimétricas	61
4.17. Densidades de los diferentes instrumentos utilizados en la instancia de colas gordas simétricas	62
4.18. Densidades de los diferentes portafolios óptimos en la instancia con colas gordas asimétricas, general (izq.) y con acercamiento a las colas (der.)	63

4.19. Perfil espectral (izq.) y entrópico (der.) de los portafolios óptimos en la instancia de colas gordas asimétricas.	64
4.20. Boxplot de la composición en las diferentes repeticiones para CVaR95 y Entrópica(5).	65
4.21. Densidades y perfiles espectrales de los portafolios óptimos con CVaR95 y Entrópica (5)	65
4.22. Perfiles entrópicos de los portafolios óptimos con CVaR95 y Entrópica (5)	66
4.23. Densidades de los diferentes portafolios óptimos en la instancia con colas gordas simétricas	67
4.24. Perfil espectral (izq.) y entrópico (der.) de los portafolios óptimos en la instancia de colas gordas simétricas.	68
4.25. Histogramas de la distribución de mezcla generada con dos instrumentos.	70
4.26. Histogramas de la distribución estimada.	70
4.27. Comparación de los histogramas de ambas distribuciones	71
4.28. Distribuciones marginales de los instrumentos	71
4.29. Frontera eficiente en espacio varianza-retorno, con portafolios equi-espaciados en parámetro de aversión	73
4.30. Curva de varianza de portafolios CVaR óptimos para diferentes valores de α	74
4.31. Curva de varianza de portafolios Entrópicos óptimos para diferentes valores de θ	75
4.32. Curva de varianza de portafolios Entrópicos óptimos, con aversión menor a 3	76
4.33. Evolución de composición de portafolios Entrópicos óptimos, con aversión menor a 3	77
4.34. Histograma de los tres portafolios óptimos extremos	78
4.35. Gráfico del retorno acumulado, del valor cuota de los fondos A de cada AFP.	81
4.36. Distribuciones de retornos del proceso diario del fondo A, de los diferentes agentes, y del promedio anterior	82
4.37. Gráfico de retornos acumulados, de los diferentes portafolios óptimos, y el promedio de los partícipes.	83
4.38. Distribuciones de los procesos de retornos, periodo a periodo, para los modelos modelos implementados, y el promedio del sistema.	84

4.39. Perfil espectral de los diferentes portafolios, en el problema multiperiodo 85

Capítulo 1

Introducción

1.1. Antecedentes y Motivación

Los recientes avances en optimización con aversión al riesgo, como por ejemplo la formalización y axiomatización de la teoría de medidas de riesgo, con un sólido marco teórico coherente y consistente, que sistematiza y formaliza intuiciones y fenómenos conocidos en el área motivan su aplicación a problemas, donde los agentes deben tomar decisiones con incertidumbre, en procesos donde no sólo importa el valor esperado sino evitar realizaciones con resultados de peor caso para los agentes.

El origen de esta teoría surge en problemas de aprovisionamiento de capital en fondos de garantía colateral. Se han realizado aplicaciones en diferentes áreas donde es relevante el riesgo, por citar alguno de ellos, en problemas de transporte donde es necesario no sólo calcular rutas óptimas en tiempo y/o distancia en valor esperado, sino que encontrar soluciones que garanticen condiciones deseables sobre la distribución de tiempos de viaje, considerando incertidumbre por efectos como la congestión o eventos inciertos.

El enfoque de utilizar este marco conceptual, en problemas financieros de ahorro para la jubilación, donde la interpretación de un fenómeno y su optimización vía aversión actúen como metodología robusta de resolución, no ha sido explorado adecuadamente en la literatura.

El presente trabajo busca acercar los últimos desarrollos en el área de optimización con aversión al riesgo, con problemas de composición de portafolios de inversión, en particular, se explora su aplicación para la estructuración de las inversiones de los fondos administrados por las *Administradoras de Fondos de Pensión* (AFPs) en Chile. Se plantea y resuelve un problema semejante al que enfrentan estos agentes. Esta propuesta de inversión es comparable en términos de desempeño como un benchmark, basado en una lógica de agente optimizador. Adicionalmente como aporte metodológico, se incorpora y resuelve la Medida de Riesgo Entrópica en el problema de optimización, siendo utilizada esta última como función objetivo del problema, enfrentando su no linealidad a través de un algoritmo de aproximación de la región óptima vía generación de cortes, metodología ampliamente utilizada en la literatura, con trabajos como: [10], [68], [67] y [45].

Un ejemplo que facilita la comprensión de la optimización con aversión al riesgo, es el siguiente: Considere una red de transporte, representada por el grafo de la izquierda en la figura 1.1, donde se tienen dos posibles caminos para un mismo destino. Esta red será utilizada por un sistema de ambulancias. Se considera que si un paciente no es atendido antes de los primeros 10 minutos su probabilidad de sobrevivencia es nula. Si se busca minimizar el tiempo promedio de viaje, la ruta óptima será e_1 con un tiempo promedio de $\bar{t}_{e_1} = 7,7$ minutos¹, pero en promedio morirá el 5.6% de los pacientes, por otro lado, si se busca minimizar el peor tiempo, la ruta escogida será e_2 con un tiempo promedio mayor de $\bar{t}_{e_2} = 8,4$ minutos, pero con ningún viaje sobre los 10 minutos.

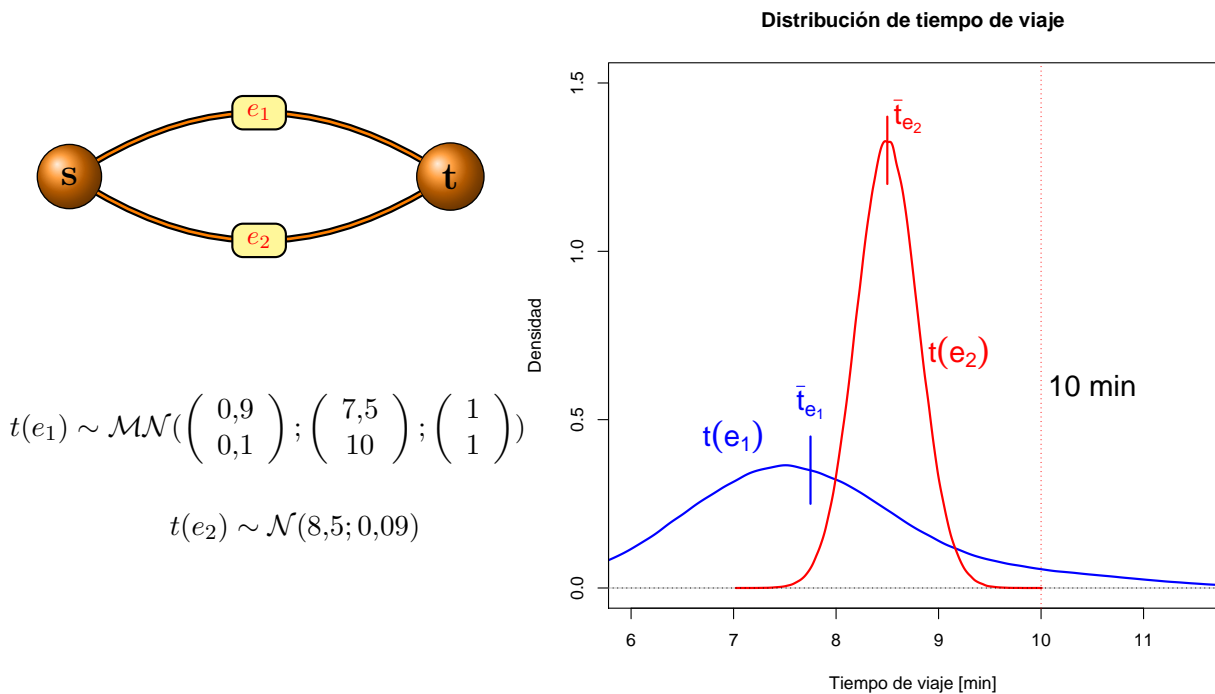


Figura 1.1: Ejemplo de la incertidumbre en los tiempos de viajes y la optimización con aversión al riesgo, en el caso de un sistema de despacho de ambulancias.

En Chile el sistema de pensiones es clasificado como de aporte definido, es decir, es conocido por todos los afiliados el aporte individual que se debe realizar a fin de mes de forma obligatoria, siendo este último un porcentaje del sueldo bruto, este dinero es ingresado a cuentas individuales que son administradas por las AFPs, empresas cuyo único objetivo es administrar dichos fondos y que deben actuar bajo un marco regulatorio especial, debiendo el cotizante pagar por sus servicios una comisión mensual previamente definida. Las AFPs en Chile, en la actualidad, administran cinco diferentes tipos de fondos de inversión, cada uno de ellos con perfiles de riesgo-retorno distintos, coherente a diferentes intereses y perfiles de riesgo de los afiliados.

La principal diferencia entre los fondos de inversión, como se dijo anteriormente, es buscar generar perfiles de riesgo-retorno distintos, implementando este esquema, a través de restricciones en la elegibilidad y límites de inversión por activo y tipo de activo. Esta lógica, permite a los diferentes afiliados tomar la decisión entorno a que tipo de perfil elegir, asumiendo un nivel de volatilidad y retorno distinto. A su vez, este esquema permite introducir una diferenciación en el

¹ $t(e_1)$ sigue una distribución de mezcla de dos normales ($\mathcal{MN}(\vec{\pi}, \vec{\mu}, \vec{\sigma}^2)$), que genera una cola derecha gorda.

tipo de comportamiento de los diferentes afiliados, dada la etapa de vida en que se encuentre. No debemos olvidar que desde un punto de vista pragmático el problema de inversiones en fondos para la vejez, constituye un problema a largo plazo, donde el perfil de riesgo aceptable al inicio de la inversión es completamente opuesta al perfil de salida, donde ante una caída repentina en el valor del fondo, generará que no necesariamente se tendrá el tiempo suficiente de recuperación antes de la jubilación, tomando los conceptos de diversificación y mitigación de peores casos un rol crítico.

Para tener una magnitud del tamaño de los fondos, el 7 de Septiembre de 2015, considerando las seis AFPs y los cinco fondos de cada una, el sistema se encontraba conformado por aproximadamente US\$ 151.9 billones de dólares americanos (10^9), lo que representa aproximadamente un 58.89 % del PIB Chileno del año 2014². A nivel de cobertura, el impacto en número de personas, ha pasado de 1.4 millones de afiliados en diciembre de 1981 a aproximadamente 9.2 millones en septiembre de 2015, con un crecimiento promedio del 7 % anual. Por otro lado, los cotizantes han aumentado de diciembre de 1982, de 1.1 millones, a aproximadamente 5.1 millones en septiembre de 2015, con un crecimiento promedio anual del 5.4 %³.

Como es común en la administración activa de portafolios, las rentabilidades de los fondos de pensión de cada AFP, son evaluados contra un *benchmark*⁴ por fondo, que en este caso es el promedio del mes anterior de los demás partícipes para igual tipo de fondo, teniendo como restricción dura un *tracking error* acumulado menor al 3 % en tres años. Esto entrega un fuerte incentivo al sistema a tomar comportamiento de manadas, tal como expone [49] y [89], generando AFPs líderes del sistema y seguidores, no necesariamente guiando al sistema a una condición riesgo-retorno óptima.

En esta línea, el presente trabajo proporciona una visión diferente, a través de una propuesta metodológica para la generación de *benchmarks* de rentabilidad para los respectivos fondos, adicionalmente entrega una herramienta para evaluar mecanismos regulatorios basados en la administración del riesgo, eliminando incentivos a imitar a otros actores del mercado, tomando un rol de agente optimizador, que obliga a actuar de forma competitiva en rentabilidad ajustada por riesgo. No debemos olvidar, el fin último con que fue creado el sistema, es entregar la máxima rentabilidad posible, ajustado al perfil de riesgo, sobre los fondos aprovisionados por los afiliados para su fondo de jubilación, situación que en la actualidad no queda del todo claro, al no existir un *benchmark* con esta lógica con el cual comparar.

1.2. Sobre esta tesis

Como fue expresado anteriormente, la presente tesis surge como propuesta a la necesidad de un *benchmark* de mercado para los fondos administrados por las AFPs, construido con información de mercado, que tenga su origen bajo la lógica de un agente que busca optimizar con aversión al riesgo la rentabilidad del fondo de inversión, sujeto a las mismas restricciones de inversión que enfrentan las AFPs en términos de régimen de inversión.

Adicionalmente, como aporte metodológico, acorde a lo esperado en un trabajo de postgrado, en la resolución del problema de portafolios, se reconoce el carácter estocástico de los retornos

²El PIB de Chile en 2014, fue obtenido de la página web del Banco Mundial para Chile <http://data.worldbank.org/country/chile>

³Datos obtenidos de la Superintendencia de Pensiones de Chile <http://www.safp.cl>

⁴Un benchmark es un índice o referencia, contra que comparar, tanto en ganancia como en riesgo asumido.

de los activos, incorporando las medidas de riesgo en la función objetivo, como metodología para cuantificar y controlar el mismo. Se consideran diferentes familias de medidas, para comparar sus comportamientos y desempeños, incorporando de forma natural la contraposición entre riesgo-retorno. Entre las medidas incorporadas se encuentra la medida de riesgo entrópica, la que no había sido utilizada en un contexto y rol semejante con anterioridad.

Acorde a lo anterior, se propone como objetivo general, realizar una propuesta metodológica replicable en la generación de un portafolio *benchmark* para los fondos de pensión en Chile, con un enfoque basado en medidas de riesgo y su respectiva optimización estocástica.

Para el cumplimiento de este objetivo, se vuelve necesario cumplir con un conjunto de hitos intermedios, tales como: conocer y comprender el funcionamiento de las diferentes variantes de sistemas de ahorro para la vejez, en particular la realidad desarrollada en Chile, abstrayendo el problema que enfrentan los agentes administradores, a un problema de portafolios, incorporando las medidas de riesgo, como herramienta fundamental para medir y elegir la composición de inversiones.

Lo anterior requiere un conocimiento de la teoría de optimización con aversión al riesgo, y en particular, de las familias de medidas de riesgo a utilizar, implementando posteriormente los modelos desarrollados, en un programa computacional que permita resolver con garantías estadísticas, y en un tiempo razonable los problemas planteados. De forma más específica, dado el interés en formular y resolver el problema de optimización de portafolio con la medida de riesgo entrópica, se hace crítico, el desarrollo de un algoritmo de optimización de esta medida no lineal, vía algún método lineal, como es la generación de cortes para la aproximación de la región óptima.

Se implementan diseños experimentales, que buscan responder los comportamientos explorados de las diferentes medidas, tales como: comportamiento ante correlación de activos, efecto de colas gordas, y la comparación de desempeños en ambientes relativamente realistas (multi y mono período) comparando lo desarrollado, con la clásica y aceptada solución de [58].

Como resultado, se espera obtener una metodología aplicable, replicable y robusta a cambios regulatorios, para la obtención de un índice de rentabilidad por fondo, permitiendo medir el desempeño de los mismos, comparando tanto su nivel de retorno, como de riesgo asumido en la obtención de dichos retornos. Esta metodología incorporará en el proceso de toma de decisión, no sólo las restricciones del régimen de inversión, como límites de inversión y elegibilidad de instrumentos, sino también restricciones de liquidez (en el caso multi período).

Se acota el desarrollado al fondo A, dado el nivel de acceso a la información requerida y al volumen de la misma, siendo decidora la simplicidad en la obtención de las restricciones duras del régimen de inversión, en comparación al resto de los otros fondos, donde el número de circulares que modifican los límites puede ser abrumador, reforzando lo anterior, la extensión de lo desarrollado a los demás fondos, no representa un desafío metodológico. Adicionalmente, dada la disponibilidad de datos, se acota el período de evaluación para las instancias calibradas con datos históricos, entre los años 2003 y 2013. Finalmente, dado el foco del trabajo, se limita a la utilización de modelos sencillos, en la actualización y obtención de retornos futuros en las instancias multiperiodo.

En cuanto a la estructura, el presente trabajo posee cinco capítulos. El primer capítulo busca introducir el tema, explicando y motivando la problemática en estudio, en particular, la necesidad e implicancias de los sistemas de pensiones, con especial cuidado de la creciente necesidad de un benchmark competitivo y representativo de mercado, que permita de forma sencilla a los afiliados,

tener un patrón de comparación base. Finalmente, se concluye con una breve revisión bibliográfica referente a la *Teoría Moderna de Portafolios*, la *Optimización Estocástica* y las *Medidas de Riesgo*.

En el segundo capítulo, se introduce y desarrolla el marco conceptual necesario. Posteriormente, en el tercer capítulo se exponen los modelos desarrollados, y en particular, la metodología desarrollada para la resolución de la medida de riesgo Entrópica.

En el cuarto capítulo, se realizan los experimentos computacionales necesarios, referentes a la implementación de los modelos antes enunciados y desarrollados. En estos experimentos, se responden preguntas e hipótesis formuladas, determinando convergencia de los problemas, tiempos y análisis referente a la caracterización de las distribuciones de los portafolios óptimos, como a su composición en sí.

Finalmente, el cuarto capítulo entrega las conclusiones del trabajo desarrollado, introduciendo los trabajos futuros, extensiones y áreas por investigar.

1.3. ¿Por qué un sistema de pensiones?

Los fondos de pensión nacen como una solución natural, al problema de ahorro para la vejez, y su consiguiente retiro. Desde sus orígenes estos sistemas han sido administrados por el Estado y financiados a través del sistema de impuestos o contribuciones específicas como contraparte de estos beneficios. Tradicionalmente en sus orígenes fueron concebidos como sistemas de reparto, donde trabajadores en actividad financian a quienes se encuentran en etapa pasiva, esta forma de financiamiento fue fuertemente cuestionada por su sustentabilidad, dado el creciente envejecimiento de la población, conllevando a contribuciones cada vez más altas, y por su falta de sentido de justicia contable en el método de cálculo de los beneficios, cálculos basados en mantener un nivel semejante de ingresos al de los últimos años de los trabajadores, situación estructuralmente difícil de sostener, considerando el comportamiento creciente de los ingresos durante la vida del trabajador.

En el caso de las pensiones contributivas por vejez, son ampliamente aceptadas y adoptadas por diferentes países en el mundo, coherente a la ratificación y adopción de la Declaración Universal de los Derechos Humanos de 1948, artículo 25.1, y como solución de protección social frente a la vejez como causa de la pérdida de medios de subsistencia.

En general los sistemas de pensión pueden ser clasificados como públicos, privados o mixtos, dependiendo del rol que toma el Estado en la implementación, es así como se dice que un sistema de pensiones es público cuando el Estado es quien lo administra, como herramienta de garantía de protección social, usualmente estos sistemas se implementan bajo una lógica de reparto, equidad intergeneracional e intrageneracional y la universalidad. Históricamente estos sistemas son financiados vía impuestos o cotizaciones del grupo activo. En la práctica, este tipo de sistemas han mostrado un mal desempeño, en la mayoría de las veces ha implicado que el Estado deba actuar como colateral frente a un sistema quebrado.

El sistema de pensiones privado, se basa fuertemente en la idea de ahorro y rendimientos sobre el capital ahorrado, en la realidad Chilena presenta malas pensiones, dada la existencia de grandes lagunas de cotización y montos insuficientes, principalmente producidos por los bajos sueldos y ausencia de coberturas o seguros ante eventos que impliquen una laguna, no obstante, la existencia

de estos fondos a dinamizado fuertemente la industria, dando gran liquidez de inversión, siendo considerado como una de las principales razones del fuerte crecimiento económico en la región. La principal falencia del sistema privado subyace en la desprotección al segmento más vulnerable de la población. Finalmente en la última categoría, los sistemas de pensiones mixtos (público-privados), realizan una combinación de los anteriores, estableciendo por ejemplo: un sistema con una pensión mínima garantizada por el Estado, donde posterior a la acción de la componente privada, de poseer una jubilación inferior a un límite, el Estado a través de un fondo común, suple la jubilación. En general, estos sistemas mixtos, a través de diferentes medidas buscan cubrir las imperfecciones generadas en el modelo privado.

Continuando en la clasificación de los sistemas de pensión, los sistemas de pensiones privados se dividen en dos grandes grupos, aquellos de beneficio definido, y aquellos de contribución definida. Tal como lo dice su nombre, aquellos de contribución definida se caracterizan por ser conocidos por todos los involucrados del sistema el monto de la cotización a realizar mes a mes, y por lo general se define una edad mínima de jubilación, por el contrario los sistemas de beneficio definido, establecen un monto de jubilación que es calculado como resultado del tamaño del fondo, y en ciertos casos como Estados Unidos y el Reino Unido, denominado *de media final*, donde el monto del salario de los últimos años, a través de su promedio, determina la pensión.

Por ley en Chile los fondos de pensión son administrados por empresas privadas, con regulaciones particulares, denominadas “Administradora de Fondos de Pensión” (AFP), estas empresas administran los fondos en las cuentas de capitalización individual de cada uno de sus afiliados, siendo un sistema de beneficio definido, conformando en la actualidad un sistema con seis AFPs, con un pilar privado de capitalización personal y una reciente componente pública (pilar solidario), como método de garantía social.

La actual estructura de multifondos del sistema se obtuvo a través de un proceso iterativo que inicio el año 1981 con el Decreto Ley 3.500 del 6 de diciembre de 1980, remplazando el antiguo sistema de repartos por uno basado en el ahorro obligatorio individual. Inicialmente el sistema contó con un único fondo, conformado a diciembre de 2007 por un 27.47% en renta variable y 72.46% en renta fija. En mayo de 2000 entra en funcionamiento el fondo E, más conservador, compuesto sólo por renta fija. En febrero de 2002 se dicta la Ley 19.795 que introduce tres nuevos fondos, entregando la actual estructura de multifondos, conformado por cinco fondos designados por las letras: A, B, C, D y E donde a través de regulación en la elegibilidad, proporción y cobertura se implementa los perfiles riesgo-retorno de los diferentes fondos, siendo el fondo A el más agresivo y el fondo E el más conservador.

1.4. Revisión bibliográfica y estado del arte

1.4.1. Teoría Moderna de Portafolios

La teoría moderna de portafolios es atribuida al trabajo de [58] y su posterior libro [59], mediante un enfoque centrado en el valor esperado y la varianza como medida de dispersión. Con el enfoque inicial adoptado, donde es minimizada la varianza de la cartera, sujeta a restricciones de rentabilidad mínima, surge de forma natural el concepto de frontera eficiente, como todas aquellas combinaciones de desviación estándar y valor esperado óptimas, donde dado un nivel de retorno

pedido, es asociado a un nivel de volatilidad mínima a enfrentar, como claro paradigma de riesgo-retorno. Surge la contraposición natural de que a mayor valor esperado, implica enfrentar mayores niveles de riesgo, donde la elección de la combinación óptima riesgo-retorno depende únicamente de preferencias individuales. Agregando a la discusión un elemento central y clave, la composición del portafolio óptimo no puede ser guiado únicamente por atributos individuales de cada activo, sino también por su comportamiento conjunto, permitiendo generar coberturas ante escenarios extremos conjuntos, agregando instrumentos que minimizan el riesgo conjunto total.

Desde su origen a la fecha han surgido diferentes variantes de desarrollo, algunas de ellas centradas en grupos de investigación de origen más económico, buscando validar y/o expandir el supuesto de que un enfoque puramente de media-varianza sea suficiente, como es el caso de [91], donde evalúa las condiciones necesarias sobre la función de utilidad de los agentes y de los procesos generadores de retornos para que un enfoque puramente media-varianza sea óptimo. Trabajos posteriores como [60] y [50] buscan estimar el efecto de agregar momentos de orden superior al problema, otros en cambio, buscan mejorar el entendimiento y precisión de la distribución de retornos futuros de los instrumentos, como es el caso de [29], [26], [6] y [24], o el comportamiento conjunto de los instrumentos, donde el supuesto de relaciones de covarianza estables en el tiempo no es válido, más aún en crisis, al igual que distribuciones gaussianas simétricas sin colas gordas.

Por la variante de comprender de mejor forma la naturaleza de los procesos generadores de retornos futuros, en particular su valor esperado, como precio fundamental de equilibrio, surge el *Capital Asset Pricing Model* CAPM con trabajos como: [94], [93], [84], [54] y [64], donde bajo condiciones de equilibrio en un mercado eficiente y líquido es posible describir el comportamiento individual a partir de características propias del mercado y su relación con el comportamiento colectivo del mercado, todo lo anterior premiado con el premio Nobel en Economía del año 1990, para un desarrollo histórico más detallado al lector se recomienda leer [90]. Posteriores avances en esta misma línea, al introducir otros factores para buscar explicar retornos futuros, logran resultados como el conocido, y premiado con el premio Nobel en Economía del año 2013, modelo de [31] y [30]. Otros enfoques en esta misma línea, bajo condiciones puramente de no arbitraje de mercado, abren campos nuevos de investigación como es la Teoría de valorización por arbitraje (*Arbitrage Pricing Theory*), donde se recomienda leer [75] y [17]. Para una revisión histórica más detallada de la teoría moderna de portafolios, se le recomienda [25] al lector.

En otra línea, más cercana a Matemáticas Financieras han surgido variantes asociadas a considerar el costos de transacción, su implementación y tratamiento computacional, aún cuando usualmente estas modificaciones pueden afectar la convexidad del problema. Adicionalmente posterior al trabajo inicial [8] y su aplicación en problemas de aprovisionamiento [7], surge nueva literatura ligada a la medición del riesgo y sus implicancias en depósitos, diferenciando el concepto de medida de dispersión, originalmente utilizada como la varianza, con las medidas de riesgo propiamente tal. Este nuevo enfoque es rápidamente extendido con aplicaciones en finanzas, abordando inicialmente el problema de los depósitos colaterales en el sistema bancario, para posteriormente extenderse de forma natural al problema de portafolios de inversión.

Finalmente, algunas de las últimas innovaciones en el área, surgen del cuestionamiento de los supuestos de racionalidad y eficiencia de mercado, con trabajos como [85] y [57]. Adicionalmente, en lo referente a medición de riesgo y estructuración de portafolios, desde el trabajo [76], se habla de la época *Post-modern Portfolio Theory*, donde toma especial realce el *Downside Risk* como medida de riesgo de caída, con trabajos como [88], [86] y [87]. En aspectos normativos referido a los fondos de pensiones en el mundo, han surgido cuestionamientos a la gobernabilidad y buenas prácticas,

principalmente cuestionando los supuestos de racionalidad y consistencia, con trabajos como [65], [21] y [15] muchos aspectos de ellos, en línea con lo ya analizado en su origen por [22] y su reimpresión [23], aspectos cubiertos por la nueva línea de investigación llamada *Behavioral Finance*.

1.4.2. Optimización Estocástica y las Medidas de Riesgo

La optimización bajo incertidumbre u optimización estocástica hace referencia al proceso de optimización de modelos, donde existen parámetros que no son fijos, por el contrario, actúan de modo aleatorio bajo una cierta ley de probabilidad conjunta. De lo anterior, la optimización estocástica actúa como una extensión natural de la optimización determinista, en particular, para el tipo de modelos lineales con que se trabaja en este trabajo, se basa fuertemente en la programación lineal matemática, lo que permite que dada la posibilidad de obtener realizaciones de los parámetros del universo de eventos posibles, el problema pueda ser tratado.

Reconociendo lo anterior, en la presente revisión bibliográfica, se deja de lado lo referente a la evolución y estado del arte en la optimización lineal y mixta que han tenido su fuerte auge y desarrollo durante la Segunda Guerra Mundial, desde los trabajos iniciales de [46] con la introducción de problemas lineales, y posteriores trabajos como [42] con una de las primeras formulaciones de problemas en redes de transporte, que abrió la puerta al método Simplex que es propuesto y formalizado por el trabajo de [18], posteriormente con el trabajo de von Neumann y la conjetura de la teoría de dualidad, se abre la puerta al desarrollo de nuevos métodos, para finalmente, con el trabajo de [47] poder ser demostrado, a través del método de punto interior, que los problemas lineales pueden ser resueltos en tiempo polinomial.

En la optimización bajo incertidumbre existen principalmente dos enfoques metodológicos y de modelamiento. El primero de ellos la optimización robusta, que busca entregar soluciones robustas desde el punto de vista de los parámetros con estocasticidad, mitigando y cubriendo sobre los peores casos realizables. Y el segundo enfoque, comúnmente llamado optimización estocástica a secas, donde la incertidumbre es introducida tanto en el poliedro de restricciones vía: restricciones en probabilidad, esperanza o momentos superiores; como en la función objetivo con parámetros de costo que siguen una cierta función de distribución, o se tiene acceso a un generador del mismo.

Es importante añadir, que si bien el enfoque de optimización robusta es muy diferente al empleado en optimización bajo incertidumbre, ambos enfoques se encuentran fuertemente ligados, pudiendo ser planteados problemas equivalentes, es así como en [66] y [13] se muestra y utiliza dicha equivalencia. Para una mayor revisión entorno a optimización robusta, y sus últimos avances, se le recomienda [39] al lector.

Dentro de la optimización estocástica se encuentra la optimización con aversión al riesgo, que como su nombre lo dice busca cuantificar y entregar soluciones óptimas ajustadas en riesgo, al existir un proceso con resultados inciertos. En esta línea, con el trabajo seminal de la teoría de portafolios [58], se puede hablar de los orígenes de las medidas de riesgo, una metodología que busca generar un mapeo inequívoco y coherente entre variables aleatorias y el riesgo inducido por las mismas, generando una noción de orden. Dado que en la era pre Markowitz, el riesgo era considerado sólo como un factor de corrección del retorno esperado.

Posteriormente, la necesidad de cuantificar nuevamente riesgo, esta vez para establecer el

monto necesario en un fondo colateral de garantía, para que el riesgo del mismo sea aceptable, hace renacer la noción de medida de riesgo, esta vez desde un punto de vista axiomático, exigiendo ciertos comportamiento básicos deseables, es así como con el trabajo de [8], se introduce el concepto de *medida de riesgo coherente*, cuestionando y desplazando fuertemente la noción de Value at Risk (VaR), que entrega la peor pérdida posible observada dado un nivel de confianza, ampliamente utilizada en la industria, principalmente al no cumplir con el principio básico de la diversificación del riesgo. A su vez, se entrega como opción coherente el Conditional Value at Risk (CVaR) [72], [4] y [44], que gracias a su tratabilidad en el proceso de optimización, por su equivalencia como problema lineal [73], permite y fomenta su auge. A su vez, el CVaR mantiene su interpretación, como la pérdida promedio observando las peores α % realizaciones.

Desde sus inicios, hasta la actualidad, se han realizado diversas aplicaciones y extensiones naturales, tales como los conceptos de Convex Risk Measures [36], Spectral Risk Measures [3], y Deviation Measures [71]. Y aplicaciones como: [12] en problemas de planificación de transporte, [78] en diseño de cadena de suministros, [62] en el tratamiento mediante descomposición de un problema en dos etapas.

A su vez, CVaR ha tomado mayor importancia, dado su rol como base de descomposición a diferentes niveles de riesgo, como una representación de Kusuoka [80], de las llamadas medidas de distorsión, medida de riesgo coherentes ley invariante, conocidas también como medidas de riesgo espectrales [13], donde ambos conceptos, provenientes de dos desarrollos muy diferentes, son equivalentes [41], al lector interesado en profundizar respecto de esta descomposición se le recomienda revisar [2] y [3]. Esta descomposición tiene grandes implicancias prácticas en el área, dado que resolver de forma eficiente y rápida CVaR, pasa a ser suficiente para optimizar cualquier medida de riesgo coherente ley invariante, para lo que ya son conocidos métodos óptimos, como lo desarrollado en [28].

Todos los resultados anteriores, sin un método claro y sencillo para resolver dichos problemas, tendrían poca importancia práctica, en el último tiempo la utilización del enfoque *sample average approximation* (SAA), donde el problema aproximado mediante realizaciones de los parámetros inciertos actúa como un estimador del problema original, donde bajo ciertas condiciones generales, y ampliamente satisfechas en los problemas estándar se obtienen garantías de convergencia tanto en la función objetivo como de la solución misma, a los interesados en este método, se les recomienda fuertemente revisar lo desarrollado en [81], [48] y [96]. Adicionalmente, para aquellos interesados en profundizar en las propiedades de esta metodología, así como sus garantías en la inferencia estadística, se les recomienda consultar el libro [82].

Uno de los temas aún en investigación, corresponde a la importancia de ciertas propiedades por sobre otras, es el caso de la convexidad y la coherencia, donde la medida de riesgo entrópica, es un ejemplo de medida convexa, no coherente [69]. El origen de esta medida, se relaciona con la función de utilidad exponencial ampliamente utilizada en economía [77], donde a través del ajuste del parámetro de aversión, es posible modelar diferentes agentes tomadores de decisión con diferentes niveles de aversión. Un resultado relativamente reciente e importante relacionado a esta medida, corresponde a lo desarrollado en [16], donde se muestra que esta medida de riesgo es la única medida que cumple consistencia en la aditividad, siendo éste resultado muy importante al buscar una solución global razonable, mediante rolling forward de periodos sucesivos independientes, para posteriormente, a través del retorno acumulado obtener el resultado del ejercicio total.

Capítulo 2

Las medidas de riesgo en el problema de portafolio

En el presente capítulo se cubre el marco conceptual necesario para el desarrollo del trabajo, introduciendo y desarrollando los conceptos de: Teoría de portafolios, las Medidas de Riesgo, problemas de optimización en portafolios, y finalmente para la resolución de los problemas estocásticos que estos generan, la metodología Sample Average Approximation (SAA), con su respectiva algorítmica e inferencia estadística sobre la calidad de las soluciones.

Es importante aclarar que el propósito de las secciones iniciales del capítulo, no son entregar una extensa y detallada descripción de la administración activa de portafolios, ni ser un tratado sobre las diferentes medidas de riesgo existentes, su axiomática, propiedades y metodologías de resolución estocástica. Por el contrario, se busca introducir de forma rigurosa y clara, los conceptos necesarios para el desarrollo del trabajo. Para aquellos lectores interesados en profundizar en cada uno de los temas, a medida que se desarrolla cada uno de ellos, se entregarán las respectivas referencias recomendadas.

Se inicia brevemente con la Teoría Moderna de Portafolios, centrando el foco en el contexto y estructura de los portafolios, abstrayendo, formulando y transformando los mismos a un conjunto de problemas de programación matemática, con función objetivo y restricciones lineales. Estos problemas, son equivalentes a los que se tendrán que resolver por escenario, en su versión determinista equivalente, en los modelos implementados. Se aborda la presentación del problema, introduciendo inicialmente la lógica de inversión, posteriormente se define las restricciones que definen los portafolios, para posteriormente introducir de forma general, la función objetivo en cada variante como una cuantificación del riesgo del portafolio.

Como continuación de la Teoría de portafolios, en la siguiente sección se introducen las Medidas de Riesgo utilizadas, tanto con su axiomática, como la intuición que formaliza cada axioma, presentando cada una de las familias de medidas estudiadas: medidas espectrales (valor esperado, CVaR y ECVaR) y la medida entrópica.

Concluida la sección de medidas de riesgo, a manera de síntesis, con todos los elementos necesarios, se introducen los problemas de portafolios que en el siguiente capítulo son desarrollados y resueltos. En particular, el caso de la medida de riesgo Entrópica, donde fue necesario el desarrollo

de una metodología de aproximación del problema no lineal.

2.1. Teoría moderna de portafolios

Antes de introducir el problema de portafolios como tal, es necesario introducir y formalizar la noción de retorno sobre los activos financieros que lo definen. El ciclo de inversión lo simplificaremos como el proceso iniciado con la adquisición o retención de un activo, a un precio inicial (P_0), y un segundo precio, posterior al periodo de inversión (Δt), que llamaremos precio final ($P_{\Delta t}$) buscando medir la variación de precio del subyacente. Es importante acotar que usualmente el precio final, es el inicial del siguiente período. Es directo observar que la real observación de importancia es la variación del precio período a período, con su respectivo acumulado, descontados consideran todos los costos adicionales de transacción, comisiones y otros.

Considerando $I \subset \mathbb{N}$ al conjunto de instrumentos de inversión, para cada instrumento (activo) $i \in I$, en los periodos de decisión de inversión $t \in T$, con precios netos de transacción P_i^t , se define la rentabilidad aritmética como: $R_i^t = \frac{P_i^t - P_{i,t-1}}{P_{i,t-1}}$, siendo esta la variación porcentual del precio con respecto al precio del periodo anterior. Esta noción de rentabilidad es la más conocida, utilizada y de fácil interpretación, en particular es la utilizada por normativa para definir el interés máximo convencional. Esta metodología genera lo que llamaremos interés (o rentabilidad) simple.

Es conocido que la obtención de retornos aritméticos, extrema las observaciones, generando distribuciones extremas e irregulares de forma artificial, adicionalmente carece de ciertas propiedades deseables, como la simetría y la aditividad para obtener el retorno acumulado. Adicionalmente es aceptado, no sin críticas, la lognormalidad de los retornos en condiciones teóricas ideales de mercado, para mayor detalle ver el trabajo fundador de [9] y más recientemente, los trabajos de [1], [61] y [63], respecto de las criticas a esta idea ver [74]. Por lo anterior, se introduce el concepto de retorno logarítmico $r_i^t = \log\left(\frac{P_i^t}{P_{i,t-1}}\right)$, que puede ser interpretado como llevar en el caso del interés compuesto, el periodo de inversión infinitesimalmente a cero ($\Delta t \rightarrow 0$), generando una tasa de composición continua. Para observar lo anterior, basta con tender el número de periodos de composición a infinito, en un proceso de tasa compuesta, obteniendo: $\lim_{n \rightarrow \infty} \left(1 + \frac{r}{n}\right)^{nt} = e^{rt}$. Otra interpretación aceptada, es que actúa como una muy buena aproximación de la rentabilidad aritmética entorno a cero, lo que podemos ver en la figura 2.1. La aditividad de los retornos, como medio para obtener el retorno acumulado ($\sum_{t=0}^T r_t = \log\left(\frac{P_T}{P_0}\right)$) es una propiedad extremadamente útil cuando se resuelve un problema período a período, con real interés en el resultado acumulado final, situación que no ocurre con los retornos aritméticos.

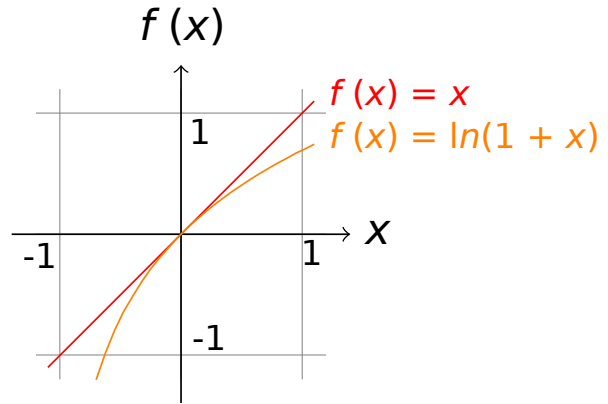


Figura 2.1: Ilustración que muestra la cercanía entre $\log(1+x)$ y x entorno a cero, lo que permite la interpretación del retorno logarítmico como una “buena aproximación” del retorno aritmético entorno al cero.

El problema central en la administración activa de portafolios es la asignación de recursos en

los diferentes activos factibles, a fin de generar un portafolio óptimo por el período de inversión, con condición de optimalidad coherente a la política de inversión de quien lo genera, ya sea con fines de rentabilidad o de cobertura a ciertos escenarios. Se distingue de la administración pasiva, por seguir un proceso continuo-activo de revisión y corrección de composición. En este proceso, se debe elegir como es distribuido el dinero, reconociendo el carácter aleatorio de los precios futuros de los instrumentos, guiados en creencias subyacentes sobre el valor conjunto de los activos, para posterior al proceso de elección, conocer las realizaciones y el resultado neto del ejercicio, comportamiento que genera de forma natural la noción de riesgo.

Al pensar en soluciones que consideren el riesgo, es fundamental considerar el proceso conjunto de rentabilidad de los activos del portafolio, con potenciales exposiciones a escenarios desfavorables, donde es claro que el criterio de elección deberá considerar tanto el comportamiento individual de los activos, como su comportamiento aleatorio agregado, con coberturas a peores casos y una diversificación frente a instrumentos semejantes, por el efecto de la incertidumbre.

En cada período se busca determinar el porcentaje de inversión en los activos factibles. En forma genérica, se identifica al conjunto de instrumentos como $I \subset \mathbb{N}$, definiendo los pesos de inversión sobre cada uno de ellos, como $\{w_i\}_{i \in I}$.

Adoptando un enfoque de programación matemática, podemos escribir el problema genérico de optimización para un período como: $\max_{\vec{w} \in \mathcal{P}} \mathbb{E}_{\vec{r}}(F(\vec{w}, \vec{r}))$ donde debemos tener una descripción de la distribución conjunta del proceso generador de los retornos futuros \vec{r} , o en su efecto un mecanismo para generar realizaciones independientes e idénticamente distribuidas.

La variante a cada uno de los modelos estudiados, dependerá directamente de la elección de la función objetivo $F(\cdot)$, del conjunto de soluciones factibles \mathcal{P} , y de la distribución conjunta de retornos sobre \vec{r} . Todos los elementos antes enunciados, serán caracterizados a continuación, partiendo por la región factible \mathcal{P} , para posteriormente, en la siguiente sección, cubrir en detalle las diferentes funciones objetivos $F(\cdot)$ vía medidas de riesgo, y finalmente en el capítulo de Experimentos Computacionales, poner especial cuidado en la función generadora de retornos.

Una restricción básica que debe cumplir un problema de asignación, es no asignar aquello que no tiene, es decir, las variables $\{w_i\}_{i \in I}$, deben cumplir la restricción de inversión máxima ($1^t w \leq 1$), que impone no invertir más de lo que se tiene, siendo conocida en economía como *restricción presupuestaria*.

La restricción anterior, sumada a la presencia de un instrumento base en el conjunto de inversión, que elimina incentivos a no invertir, a través de una rentabilidad libre de riesgo, completa la igualdad en la restricción $\sum_{i \in I} w_i = 1$, a esta nueva restricción la llamaremos *restricción de inversión total*.

Adicional a la restricción anterior, se agregan restricciones propias del régimen de inversión de los fondos, todas modelables como restricciones lineales en un poliedro de la forma $\mathcal{P}' = \{\vec{w} \in \mathbb{R}^{|I|} : A^t \vec{w} \leq b\}$, con familias de restricciones de la forma:

1. Prohibición de ventas cortas, o positividad de las variables: $w_i \geq 0 \forall i \in I$.
2. Límites de exposición o concentración por activo: $w_i \leq \delta_i \forall i \in I$.
3. Límite de exposición o concentración por grupo de activos: $\sum_{i \in G \subset I} w_i \leq \delta_G \forall G \in \mathbb{G}$.

Al poliedro solución que resulta de la intersección de \mathcal{P}' y la restricción de inversión total, lo llamaremos \mathcal{P} , conjunto que contiene todas las posibles composiciones porcentuales válidas del portafolio de inversión. En particular, se observa de la positividad y la restricción de inversión total, que es un politopo, y en efecto $\mathcal{P} \subset [0, 1]^{|I|}$.

$$\mathcal{P} := \{\vec{w} \in [0, 1]^{|I|} : A\vec{w} \leq b \ \wedge \ 1^t w = 1\} \quad (2.1)$$

Teniendo adecuadamente descrito el politopo de soluciones \mathcal{P} , nos centramos en las diferentes variantes de la función $F(\cdot)$, siendo necesario en ocasiones agregar ciertas restricciones adicionales a \mathcal{P} , para implementar la medida de riesgo o función $F(\cdot)$. Es importante observar que este enfoque generaliza lo introducido en [58], donde se interpreta la función como $F(\vec{r}, \vec{w}) = -w^t \mathbb{E}_r(rr^t)w$ de varianza del portafolio, donde adicionalmente se le agrega la restricción $\mathbb{E}_r(\vec{w} \cdot \vec{r}) \geq \bar{r}_{min}$ de retorno mínimo esperado, encontrando el portafolio de mínima varianza, sujeto a un nivel de rentabilidad esperada mínima, introduciendo el concepto de frontera eficiente de inversiones.

De forma natural, las medidas de riesgo actúan adecuadamente en esta generalización, donde a través de un marco conceptual sólido, vía exigencias axiomáticas mínimas que formalizan conceptos de sentido común en la administración del riesgo, se formula un nuevo enfoque de problemas de portafolio, esta vez minimizando la noción misma de riesgo, y no la de dispersión, como es el caso de la varianza.

2.2. Las medidas de riesgo

Las medidas de riesgo entregan, como su nombre lo dice, una medida inambigua y consistente del riesgo asociado a una variable aleatoria, generando un orden total sobre las variables aleatorias, incorporando en su construcción intuiciones de sentido común, cuantificando el riesgo asociado a un proceso aleatorio. En esta línea, su utilización en el problema de optimización de portafolios, surge de forma natural, buscando generar inversiones donde se controla el riesgo asociado al retorno del portafolio.

Existen diferentes familias de medidas, como se dijo anteriormente, inducidas por el cumplimiento de propiedades deseables diferentes. En este trabajo se consideran dos de ellas, siendo estas las más conocidas y utilizadas en la industria y academia, incluyendo la medida de riesgo entrópica, que recientemente adquirió un fuerte realce, principalmente al demostrarse en el trabajo de [16], que es la única medida de riesgo que cumple consistencia en la aditividad. La consistencia en aditividad resulta ser crítica para problemas multiperíodo, donde el resultado final del ejercicio, resulta de la adición de resultados parciales de cada período, como es el caso de los retornos logarítmicos del portafolio, para obtener el retorno acumulado, al resolver el problema de portafolio multiperíodo mediante una estrategia *rolling-forward*.

Para la formalización de conceptos, definiciones, propiedades y teoremas relevantes, se considera un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, y se trabaja con variables aleatorias en \mathcal{L}^2 dada su implicancia práctica como variables aleatorias de varianza acotada, aún cuando todo lo desarrollado sigue siendo cierto si se toman variables en \mathcal{L}^∞ , para revisar más detalles, se recomienda revisar [82]

Para lo desarrollado se mantiene la convención de variables de pérdida o costo, por lo que en el caso de los portafolios de inversión, se considera la variable aleatoria $-\vec{w}^t \vec{r}$, siendo preocupación el lado izquierdo de la distribución de retornos del portafolio.

Diremos que un mapeo $\rho: \mathcal{L}^2 \rightarrow \bar{\mathbb{R}}$ es una medida de riesgo si cumple las siguientes propiedades:

- A1) Normalización: $\rho(0) = 0$
- A2) Traslación: Si $a \in \mathbb{R}$ y $Z \in \mathcal{L}^2$, entonces $\rho(Z + a) = \rho(Z) + a$.
- A3) Monotonía: Si $Z_1, Z_2 \in \mathcal{L}^2$ y $Z_1 \leq Z_2$, entonces $\rho(Z_1) \leq \rho(Z_2)$.

La normalización, permite la comparación de las medidas de diferentes variables, con un punto de referencia común. La translación, constituye una propiedad clave, no cumplida por las medidas de dispersión, como es el caso de la conocida varianza, y formaliza la intuición clásica del activo libre de riesgo, de que agregar una posición segura a un portafolio, debe reducir su riesgo. Finalmente, la propiedad de monotonía, captura la intuición de orden y la formaliza, estableciendo que si una variable aleatoria de pérdida, siempre tiene valores menores o iguales, entonces su medida de riesgo debe ser menor o igual que la variable con valores consistentemente no menores.

En 1999 con el trabajo de [8], se introduce la noción de *coherencia* de una medida de riesgo, formalizando el concepto clásico de diversificación, por ello, diremos que una medida de riesgo es *coherente* si satisface las siguientes condiciones:

- A4) Convexidad: Si $Z_1, Z_2 \in \mathcal{L}^2$ y $\lambda \in [0, 1]$, entonces $\rho(\lambda Z_1 + (1 - \lambda)Z_2) \leq \lambda\rho(Z_1) + (1 - \lambda)\rho(Z_2)$.
- A5) Homogeneidad positiva: Si $\lambda \geq 0$ y $Z \in \mathcal{L}^2$, entonces $\rho(\lambda Z) = \lambda\rho(Z)$.

Como fue mencionado anteriormente, la convexidad introduce el concepto de diversificación, en efecto, tomando $X = \lambda Z_1$ e $Y = (1 - \lambda)Z_2$ obtenemos subaditividad, que nos dice que siempre diversificar en un portafolio, tendrá un riesgo igual o menor que mantener posiciones separadas concentradas ($\rho(X + Y) \leq \rho(X) + \rho(Y)$). La homogeneidad positiva, nos entrega un orden de crecimiento lineal del riesgo, ante incrementos lineales positivos en la posición. En la formulación anterior, existe una definición equivalente, que remplazada la convexidad por la subaditividad, lo que sumado a la homogeneidad positiva, implica la convexidad.

Es importante mencionar en este punto, que la principal medida de riesgo utilizada por el mercado, el VaR (Value at Risk) no es una medida de riesgo coherente en general. Esta medida de riesgo nos entrega la pérdida posible a un cierto nivel de confianza, no obstante, es conocido que sin el supuesto de normalidad, no es coherente. Como principal alternativa coherente al VaR, se introduce en trabajos como [8] y [4] el CVaR (Conditional Value at Risk) o ES (Expected Shortfall), medida de riesgo que actúa como el valor esperado de la cola de pérdidas, definida a partir de un nivel de confianza que determina el corte.

Resultados importantes en la teoría de medidas de riesgo, son las medidas de riesgo espectrales y las de distorsión, introducidas por [51] y [3], y por [20] y [97] respectivamente. Ambas medidas, se motivan y provienen de formulaciones completamente distintas, para posteriormente en [41] ser demostrada su fuerte relación y equivalencia. Posterior a la formulación de estas medidas, se han

Diagrama explicativo del promedio, VaR y CVaR de una variable de pérdidas a un 95% de confianza

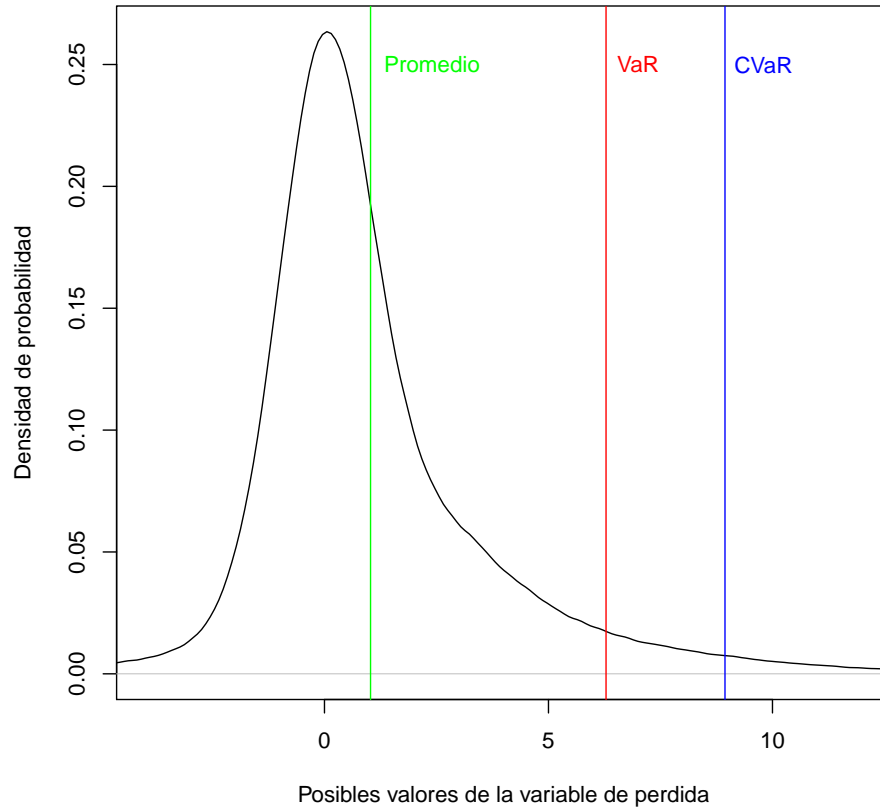


Figura 2.2: Ilustración del VaR (Value at Risk) y del CVaR (Conditional Value at Risk), en una distribución.

realizado diversos trabajos y aplicaciones relacionadas, como los trabajos de [95], [43] y [79] con aplicaciones en Finanzas, y [98], [11], [5] con un enfoque más metodológico. Para un mayor detalle se le recomienda al lector [99]. No obstante, un trabajo que generaliza lo anterior, es lo realizado en [80] donde se muestra que toda medida de riesgo coherente ley invariante, puede ser representada como una combinación espectral de CVaR a diferentes niveles de confianza, a través de una representación de Kusuoka. Este resultado es importante para efectos prácticos, dado que optimizando eficientemente CVaR, con metodos acelerados como [28], teniendo una adecuada descripción de la *función de mezcla*, sería suficiente para optimizar aproximadamente medidas de riesgo espectral de forma eficiente.

Motivados por lo anterior, se dice que una medida de riesgo es *espectral*, si es coherente y satisface las siguientes dos propiedades:

- A6) Comonotonía: Si $(Z_1(w) - Z_1(w'))(Z_2(w) - Z_2(w')) \geq 0 \quad \forall w, w' \in \Omega$, entonces $\rho(Z_1 + Z_2) = \rho(Z_1) + \rho(Z_2)$.
- A7) Ley-invariante: Si $Z_1, Z_2 \in \mathcal{L}^2$ tienen la misma distribución bajo \mathbb{P} , entonces $\rho(Z_1) = \rho(Z_2)$.

La comonotonía nos entrega un refinamiento a la idea de diversificación, agregando que dos variables aleatorias que se comportan de forma idéntica en el signo de la variación, para todo evento

$w, w' \in \Omega$, no puede generar una cobertura de riesgo con el otro activo, por ello, si componen un portafolio, no diversificarán su riesgo.

Por otro lado, la ley-invariante, asigna el mismo riesgo a dos posiciones que son idénticamente distribuidas con respecto a su ley de probabilidad \mathbb{P} , formalizando la intuición que dos variables idénticamente distribuidas, deben tener el mismo riesgo asociado. Esta última propiedad es crítica para efectos prácticos, dado que da sentido a distribuciones desde datos históricos.

A fin de permitir entrar en más detalle, a continuación se detalla las familias de medidas de riesgo consideradas en este trabajo. En particular, es conocido que el valor esperado es el caso más sencillo de medida espectral, no obstante se prefirió dejar separado.

2.2.1. Medida neutra al riesgo: valor esperado

El valor esperado, es considerado la medida neutra al riesgo de una variable aleatoria X . El valor esperado, con cambios de medida de probabilidad, es utilizado en Finanzas, para la valoración de instrumentos, vía martingala equivalente con *risk-neutral measures*. Es el primer momento de una variable aleatoria, un elemento básico estadístico, entregando la tendencia central de la variable. Considerando una variable continua, discretizable según una grilla D :

$$\rho^{exp}(X) = \mathbb{E}_{\mathbb{Q}}(X) := \int_{x \in \mathbb{R}} x d\mathbb{Q}(x) \quad (2.2)$$

2.2.2. Medidas espectrales: CVaR y ECVaR

Adicional a la definición anterior en 2.2, se considera que una medida $\rho_{\phi}(\cdot)$, es una medida de riesgo *espectral*, si se puede caracterizar por:

$$\rho_{\phi}(X) = \int_0^1 \phi(p) q_X(p) dp \quad (2.3)$$

Donde $\phi(p)$ es la función de pesos, sobre los cuantiles de pérdida definidos por $q_X(p) = F_X^{-1}(p)$, con $F_X(x) = \mathbb{P}(X \leq x)$ la función de distribución, siendo exigido que ϕ sea un *espectro de riesgo admisible*, vale decir, satisface las siguientes condiciones:

- 1) ϕ es positiva.
- 2) ϕ es decreciente.
- 3) $\int_0^1 \phi(p) dp = 1$.

Para mayor detalle referente a las medidas espectrales, y su definición vía espectro de riesgo admisible, se le recomienda al lector revisar [3].

Las medidas de riesgo espectrales, inicialmente fueron concebidas como una medida de riesgo coherente, que a través del espectro de riesgo admisible, permite la flexibilidad para capturar aversiones al riesgo subjetivas de los agentes.

A modo de ejemplo, una medida espectral que usualmente es utilizada como base de representación, es el promedio de las peores pérdidas obtenibles a un α nivel de confianza, que llamaremos *Conditional Value at Risk* a un nivel α ($CVaR_\alpha$). Esta medida de riesgo coherente, es representada espectralmente, a través de la función de pesos $\phi(p) = \frac{1}{1-\alpha}$ en el dominio $[\alpha, 1]$ de la cola de pérdidas, y cero en el resto del dominio:

$$CVaR_\alpha(x) = \frac{1}{1-\alpha} \int_\alpha^1 q_X(p) dp$$

Un aspecto interesante a considerar, es que dado que el espacio de medidas de riesgo espectrales es cerrado ante combinaciones convexas, y $CVaR$ es representable espectralmente, en vez de usar $q_X(p)$ como la base para las representaciones espectrales, como fue anteriormente enunciado, puede ser ocupado $CVaR$ a distintos niveles de confianza.

En línea con lo anterior, en este trabajo se consideran dos medidas espectrales, tomando como base de representación $CVaR$ y $q_X(p)$, ellas son:

1. **Conditional Value at Risk** ($CVaR_\alpha$): Representable con base $CVaR$: $\phi(p) = 1_{\{p=\alpha\}}$, y con base $q_X(p)$: $\phi(p) = 1_{\{p \in [1, \alpha]\}} \frac{1}{1-\alpha}$.
2. **Expected Value - CVaR** ($ECVaR_{\alpha, \epsilon}$): Combinación convexa con peso ϵ de valor esperado y $1-\epsilon$ $CVaR$, que es representable con base $CVaR$: $\phi(p) = \epsilon 1_{\{p=1\}} + (1-\epsilon) 1_{\{p=\alpha\}}$, y con base $q_X(p)$: $\phi(p) = \epsilon 1_{\{p \in [0, 1]\}} + \frac{1-\epsilon}{1-\alpha} 1_{\{p \in [1, \alpha]\}}$.

Conditional Value at Risk (CVaR)

Medida de riesgo ampliamente estudiada y utilizada en el último tiempo, adicionalmente en el último tiempo a tomado un papel relevante por ser una medida coherente utilizable como base para las medidas espectrales. El valor entregado por esta medida de riesgo es interpretable como el caso promedio, si sólo se observan los $100\alpha\%$ peores casos de la distribución, razón por la que también es llamada Expected Shortfall, Average Value at Risk y también Expected Tail Loss, formalmente el CVaR a un nivel de confianza de α se define como:

$$\rho^{cvar}(X) = CVaR_\alpha(X) := \mathbb{E}_{\mathbb{Q}}(X | X \geq VaR(X, \alpha)) = \frac{1}{1-\alpha} \int_\alpha^1 VaR(X, \gamma) d\gamma \quad (2.4)$$

Donde $VaR(X, \alpha)$ es el Value-at-Risk (VaR) de X a un nivel α de confianza, es decir la peor pérdida a un cierto nivel de confianza. Para el caso de distribuciones continuas es equivalente al percentil α de los peores valores.

$$VaR(X, \alpha) := \min\{\beta \in \mathbb{R} : \mathbb{F}_X(\beta) \geq \alpha\} \quad (2.5)$$

En el trabajo de [72], se establece una formulación equivalente, donde se obtiene CVaR de x a un nivel de confianza α de resolver el siguiente problema, fácilmente linealizable:

$$\min_t \left(t + \frac{1}{1-\alpha} \mathbb{E}_{\mathbb{Q}}(\max\{x-t, 0\}) \right) \quad (2.6)$$

Combinación convexa entre valor esperado y CVaR (ECVaR)

Medida de riesgo generada a través de la combinación convexa entre el valor esperado y el CVaR, esta medida a diferencia del CVaR, no es *ciega* al $\{[0, \alpha]\}$ de la distribución, considera el efecto de observaciones fuera de la cola de pérdidas definida. El efecto relativo de la importancia entre el valor esperado, y CVaR es controlado por el parámetro de mezcla ϵ . Por las razones anteriores, se conjetura que esta medida debería tender a tener una mejor convergencia estocástica vía SAA que CVaR por sí sólo.

$$\rho^{ecvar}(X) = ECVaR_{(\alpha, \epsilon)}(X) := \epsilon \mathbb{E}(X) + (1-\epsilon) CVaR_{\alpha}(X) \quad \epsilon \in [0, 1] \quad (2.7)$$

Adicionalmente, la formulación anterior es un caso particular de una medida espectral, y también puede ser interpretada y relacionada con un problema equivalente a la lógica media-varianza de la forma:

$$\begin{aligned} & \max_{\vec{x} \in X} \mathbb{E}_Z[\vec{x} \cdot Z] \\ s.t. \quad & CVaR_{\alpha}(\vec{x} \cdot Z) \leq \bar{\mu} \end{aligned}$$

donde de forma explícita se intenta maximizar los retornos, que es equivalente a minimizar las pérdidas, sujeto a que el nivel de riesgo de la distribución sea menor a un cierto límite $\bar{\mu}$ que dependerá de ϵ . Tomando $\epsilon = \frac{1}{1+\lambda}$ con $\lambda \in [0, \infty)$ se obtiene un problema equivalente en solución a plantear el problema desde un enfoque Lagrangiano, con penalización por restricción, del tipo: $\max_{\vec{x} \in X} \mathbb{E}_Z[\vec{x} \cdot Z] - \lambda CVaR_{\alpha}(\vec{x} \cdot Z)$ con incertidumbre sobre la variable Z .

Una ventaja práctica de este enfoque, será controlar de forma directa mediante ϵ , la exposición a mayores riesgos con el fin de obtener rentabilidades sobre el portafolio, adicionalmente es importante recalcar una vez más, que esta medida es igualmente coherente y por ende cumple las mismas propiedades deseables sobre la medición del riesgo.

2.2.3. Medida de riesgo entrópica

Las medidas de riesgo entrópicas, constituye el ejemplo por excelencia de una medida convexa (no lineal), pero no coherente. Dado un nivel de aversión $\theta > 0$, el riesgo entrópico de la variable aleatoria X , en su representación robusta y explícita, está definido por:

$$\rho^{ent}(X) := \sup_{\mathbb{Q} \in \mathcal{M}_1} \left\{ \mathbb{E}_{\mathbb{Q}}[-X] - \frac{1}{\theta} H(\mathbb{Q}|\mathbb{P}) \right\} = \frac{1}{\theta} \log(\mathbb{E}(e^{-\theta X})) \quad (2.8)$$

Donde $H(\mathcal{Q}|\mathcal{P})$ denota la entropía relativa de la medida \mathcal{Q} con respecto a \mathcal{P} , donde \mathcal{Q} es absolutamente continua con respecto a \mathcal{P} ($\mathcal{Q} \ll \mathcal{P}$). Esta medida de riesgo, surge de considerar agentes con aversión al riesgo, modelada por la función de utilidad exponencial $u(c) = 1 - e^{-ac}$, única dentro de la familia de funciones de utilidad CARA (Constant Absolute Risk Aversion), es decir, que marginalmente la tasa de cambio de la utilidad, con respecto al cambio en la utilidad es constante, formalmente dada una constante a se cumplen que $-\frac{u''(c)}{u'(c)} = a$.

El parámetro $\theta > 0$, controla el nivel de aversión de la medida, cumpliendo en sus límites que $\lim_{\theta \downarrow 0} \rho_{\theta}^{ent}(X) = \mathbb{E}(-X)$, es neutra al riesgo y en el otro sentido $\lim_{\theta \uparrow \infty} \rho_{\theta}^{ent}(X) = \text{ess sup}(-X)$ se observa el peor caso probable de $-X$.

Una forma que resulta interesante de escribir esta medida, es considerar descomponer X , ya sea en su promedio $\langle X \rangle$ o $\text{ess sup}(X)$, y su variación con respecto al elegido, que llamaremos ξ_X , en cuyo caso:

$$\rho^{ent}(X) = \langle -X \rangle + \frac{1}{\theta} \log(\mathbb{E}(e^{-\theta \xi_X})) \quad (2.9)$$

Esta observación resulta crítica tanto en la interpretación de la medida, donde queda en evidencia el control vía θ , del efecto de las desviaciones de X .

Esta medida puede ser deducida, como se dijo anteriormente, vía maximización de utilidad esperada de agentes o a través del principio de aditividad de premios por riesgo, para profundizar en esta formulación se invita al lector a revisar [38], [37] y [35].

2.3. El problema de portafolio

Con los elementos introducidos en las secciones anteriores, es posible definir detalladamente los problemas de optimización de portafolios, que serán resueltos. Se acota que en todos los problemas de optimización, el poliedro de soluciones factibles \mathcal{P} definido en 2.1 será el mismo.

Es importante destacar que todos los problemas desarrollados, son problemas de optimización estocástica con incertidumbre sólo en la función objetivo, donde la función objetivo $\mathbb{E}(F(\vec{w}, \vec{r}))$ para una cierta función $F(\cdot)$. Para este tipo de problemas, es necesario métodos que resuelvan los problemas, y a su vez, entreguen garantías estadísticas sobre las soluciones. Este aspecto metodológicamente será desarrollado en la siguiente sección. A continuación, se presentan de forma específica los problemas desarrollados.

2.3.1. Portafolio de maximización del valor esperado

Este portafolio maximiza el valor esperado del portafolio, este enfoque es ciego a la distribución de las colas de la distribución. La principal crítica a este portafolio, es su alta concentración e inestabilidad en las soluciones, en particular, ante instrumentos con valor esperado igual o semejante, donde por aspectos de error numérico concentrará toda la inversión en un único activo.

El problema queda completamente definido por:

$$\min_{\vec{w} \in \mathcal{P}} -\mathbb{E}(\vec{w} \cdot \vec{r}) \quad (2.10)$$

De donde la aleatoriedad se presenta en los retornos \vec{r} , pudiendo transformar la función objetivo, dada la linealidad de los elementos a $-\vec{w} \cdot \mathbb{E}(\vec{r})$, con \vec{w} el vector de pesos sobre los instrumentos.

2.3.2. Portafolio de minimización del Conditional Value at Risk (CVaR)

Este portafolio minimiza el promedio de las peores realizaciones, que se producen a un nivel α de confianza. De este modo, el problema determina la composición del portafolio \vec{w} , consideran que en la distribución del portafolio óptimo, el promedio de la cola izquierda de los retornos (pérdidas) bajo un nivel α de confianza, sea el mínimo posible.

Como fue mencionado en 2.2.2, utilizando la equivalencia propuesta por [72], el problema de optimización que define el portafolio, es el siguiente:

$$\min_{t \in \mathbb{R}, \vec{w} \in \mathcal{P}} \left(t + \frac{1}{1 - \alpha} \mathbb{E}(\max\{\vec{w} \cdot \vec{r} + t, 0\}) \right) \quad (2.11)$$

Este portafolio de optimización, surge como solución convexa y coherente, al *Value at Risk* (VaR) ampliamente utilizado en la industria, por su simplicidad en significado. Es interesante acotar que CVaR y VaR se encuentran íntimamente ligados, en particular, en la solución del problema, el término t , es exactamente el VaR, marcando el punto desde donde CVaR promedia las realizaciones inferiores.

Las principales críticas asociadas a este portafolio son: (1) la necesidad de una distribución o un número grande observaciones para su cálculo. (2) que el portafolio óptimo no distingue entre soluciones con igual promedio en la cola izquierda pero con comportamiento completamente distintos posterior al nivel α de corte.

En este contexto, es importante acotar que la formulación equivalente vía dualidad, permite resolver en términos prácticos CVaR de forma más eficiente, en particular al observar la estructura de los cortes de este problema. Al lector interesado, se le recomienda revisar [28].

2.3.3. Portafolio de minimización de CVaR, con retorno mínimo

Portafolio que surge como punto medio al CVaR y al valor esperado. Este portafolio, mediante el parametro $\epsilon \in [0, 1]$, controla el efecto del CVaR o el valor esperado. Este portafolio permite

minimizar el riesgo asociado a la cola izquierda de la distribución de retornos, teniendo cuidado con el retorno esperado del portafolio.

Una interpretación equivalente al mismo, considerando lo expuesto anteriormente en 2.2.2, en la presentación de la medida de riesgo, permite interpretar este portafolio como una minimización de CVaR con restricción de retorno mínimo esperado, lo que a su vez, de forma dual equivalente, se traduce en la maximización del retorno esperado, sujeto a una restricción de riesgo en CVaR.

El problema de optimización que define el problema, es el siguiente:

$$\min_{t \in \mathbb{R}, \vec{w} \in \mathcal{P}} -\epsilon \mathbb{E}(\vec{w} \cdot \vec{r}) + (1 - \epsilon) \left(t + \frac{1}{1 - \alpha} \mathbb{E}(\max\{\vec{w} \cdot \vec{r} + t, 0\}) \right) \quad (2.12)$$

Nuevamente la crítica asociada a este portafolio, es que la medida es ciega, en la distribución óptima, a las realizaciones que se encuentran sobre el valor esperado o el nivel α de confianza definido en el corte. Por lo que no puede diferenciar entre dos soluciones con igual promedio y $CVaR_\alpha$, pero que la cola de ganancias es completamente distinta.

2.3.4. Portafolio de minimización del riesgo entrópico

Portafolio que surge de minimizar el riesgo entrópico, a un nivel de aversión θ . La solución a este portafolio, cumplen aditividad en las soluciones, como fue mencionado en 2.2.3 por la medida de riesgo entrópica, lo que se traduce en que dividir un problema global, en sub problemas aditivos, obtienen la misma solución

Si bien la medida de riesgo entrópica es $\frac{1}{\theta} \log(\mathbb{E}(e^{-\theta X}))$, en la práctica se calcula $\mathbb{E}(e^{-\theta X})$, dado que la función $\frac{1}{\theta} \log(\cdot)$ es monótona creciente, por ellos, el problema de optimización a resolver, que define el portafolio entrópica, es el siguiente:

$$\min_{\vec{w} \in \mathcal{P}} \mathbb{E}(e^{-\theta \vec{w} \cdot \vec{r}}) \quad (2.13)$$

A diferencia de las medidas anteriores, esta considera en su totalidad la distribución del portafolio óptimo. La principal interpretación al efecto de esta medida, es la transformación de la medida de probabilidad sobre los peores eventos, efecto controlado por el parametro de aversión θ .

2.4. Metodología de resolución: Sample Average Approximation (SAA)

Para la resolución de los problemas de optimización planteados, se utiliza la técnica Sample Average Approximation (SAA), por lo que se considera relevante su introducción. Esta presentación

se basa en [83], específicamente en las secciones 5.1.1 y 5.1.2 para la consistencia estadística y propiedades asintóticas de los estimadores, y en la sección 5.6.1 para la evaluación de la solución a través del gap de optimalidad. Inicialmente en 2.4 se introduce SAA, para posteriormente verificar consistencia y convergencia para las medidas implementadas.

Bajo exigencias razonables sobre la función objetivo a optimizar y la ley de probabilidad sobre el conjunto de incertidumbre, esta técnica provee una herramienta sencilla y versátil para la resolución de problemas de programación estocástica que reúnen las siguientes tres condiciones: (1) la resolución mediante un método cerrado no es viable, (2) la evaluación de la función objetivo, en un escenario fijo, es sencilla, y (3) el conjunto de posibles soluciones, es lo suficientemente grande de modo que resulta inviable evaluar todas las posibilidades.

La técnica se enfoca en resolver repeticiones de problemas equivalentes ciertos, a partir de realizaciones del conjunto de incertidumbre, resolviendo dichas “repeticiones del problema”, para posteriormente vía inferencia estadística sobre las soluciones, cuantificar el gap estocástico, obteniendo un candidato a solución con garantías estadísticas sobre el mismo.

Esta metodología es aplicable a problemas de optimización, que pueden ser formulados de la forma:

$$\vartheta^* := \min_{x \in \mathcal{P}} \mathbb{E}_\xi [F(x, \xi)] \quad (2.14)$$

donde $\mathcal{P} \subset \mathbb{R}^n$ es un politopo no vacío, llamaremos a $f(x) := \mathbb{E}_\xi [F(x, \xi)]$, ξ es el vector aleatorio con ley de probabilidad \mathbb{P} soportada en $\Xi \subset \mathbb{R}^d$, y $F : \mathcal{P} \times \Xi \rightarrow \mathbb{R}$ una función Carathéodory, es decir, continua en x y medible en ξ . Se asume que $f(x)$ esta bien definida y es finita para todo $x \in \mathcal{P}$.

Suponemos que se tiene una muestra $\{\xi_i\}_{i \in \{1, \dots, N\}}$ independiente e idénticamente distribuida (iid), de realizaciones del vector ξ , las que pudieron ser generadas vía un generador de números pseudo aleatorios con alguna técnica de Monte Carlos.

Para cualquier valor de $x \in \mathcal{P}$ se puede aproximar el valor de $f(x)$, vía el estimador del promedio de los valores $F(x, \xi_i)$ con $i \in \{1, \dots, N\}$, en particular, se puede plantear el problema aproximado de 2.14, siguiente:

$$\hat{\vartheta}_N^* := \min_{x \in \mathcal{P}} [\hat{f}_N(x)] := \min_{x \in \mathcal{P}} \left[\frac{1}{N} \sum_{i=1}^N F(x, \xi_i) \right] \quad (2.15)$$

Las realizaciones iid $\{\xi_i\}_{i \in \{1, \dots, N\}}$ no sólo pueden ser vistas como realizaciones particulares de la variable ξ , sino que también puede ser interpretados como escenarios considerados en el proceso de optimización, escenarios con probabilidad $\frac{1}{N}$, el desarrollo de esta idea permite desarrollar técnicas de *importance sampling*, tema que no será tratado en este trabajo, pero son importantes de acotar.

Un aspecto a considerar, es que la función aproximada $\hat{f}_N(x)$ y por ello la solución del problema 2.15, dependen de la muestra escogida, en este sentido son elementos aleatorios en si.

Por la Ley de los Grandes Números (LGN), bajo ciertas condiciones de regularidad que se discutirán más adelante, $\hat{f}_N(x)$ converge puntualmente a $f(x)$ c.s cuando $N \rightarrow \infty$, considerando una muestra iid. Adicionalmente se tiene que el estimador $\hat{f}_N(x)$ es insesgado, es decir $\mathbb{E}[\hat{f}_N(x)] = f(x)$,

por ende es esperable que tanto la solución como el valor óptimo de 2.15 converjan a los valores del problema real 2.14 c.s. cuando $N \rightarrow \infty$. Denotaremos por S al conjunto de soluciones óptimas del problema verdadero 2.14, y por \hat{S}_N al conjunto de soluciones óptimas del problema aproximado 2.15.

Antes de discutir garantías estadísticas de convergencia de SAA en los problemas implementados, se definirán las nociones de consistencia, convergencia asintótica y validación (gap de optimalidad de una solución) de los estimadores \hat{S}_N y $\hat{\vartheta}_N^*$. Es importante destacar que las siguientes nociones se basa íntegramente en [83, 5.1.1 Consistency of SAA Estimators], [83, 5.1.2 Asymptotics of the SAA Optimal Value] y [83, 5.6.1 Estimation of the Optimality Gap] siendo expuestas en el presente trabajo por completitud, recomendando fuertemente al lector a recurrir a la fuente citada.

Consistencia Se dice que un estimador $\hat{\vartheta}_N^*$ de un parámetro ϑ^* es *consistente* si $\hat{\vartheta}_N^*$ converge c.s. a ϑ^* cuando $N \rightarrow \infty$. Por otro lado, referirnos a la consistencia del conjunto \hat{S}_N , hace referencia a que c.s. $\mathbb{D}(\hat{S}_N, S) \rightarrow 0$ cuando $N \rightarrow \infty$, donde \mathbb{D} denota la distancia usual entre conjuntos, talque $\mathbb{D}(A, B) := \sup_{x \in A} \text{dist}(x, B)$ y $\text{dist}(x, B) := \inf_{y \in B} \|x - y\|$ con $\|\cdot\|$ la norma euclidiana en \mathbb{R}^n

Comenzando por la consistencia del estimador $\hat{\vartheta}_N^*$ con respecto al valor óptimo ϑ^* , se sabe que para todo $x \in \mathcal{P}$ fijo, $\hat{\vartheta}_N^* \leq \hat{f}_N(x)$, y por ello:

$$\lim_{N \rightarrow \infty} \sup \hat{\vartheta}_N^* \leq \lim_{N \rightarrow \infty} \hat{f}_N(x) = f(x) \quad (2.16)$$

Si adicionalmente se tiene puntualmente la LGN, entonces:

$$\lim_{N \rightarrow \infty} \sup \hat{\vartheta}_N^* \leq \vartheta^* \quad (2.17)$$

Sin condiciones adicionales, la desigualdad anterior puede cumplirse de forma estricta, por ello a continuación se expone un teorema 1, que nos entregan condiciones suficientes.

Teorema 1. (Theorem 5.3, [83]) *Supongamos que existe un conjunto compacto $C \subset \mathbb{R}^n$ tal que:*

1. *El conjunto S de soluciones óptimas del problema verdadero es no vacío y está contenido en C .*
2. *La función $f(x)$ es a valores finitos y continua en C .*
3. *$\hat{f}_N(x)$ converge c.s. uniformemente en $x \in \mathcal{P}$ a $f(x)$, cuando $N \rightarrow \infty$.*
4. *Con probabilidad 1 para N suficientemente grande el conjunto \hat{S}_N es no vacío y $\hat{S}_N \subset \mathcal{P}$.*

Entonces $\hat{\vartheta}_N^* \rightarrow \vartheta^*$ y $\mathbb{D}(\hat{S}_N, S) \rightarrow 0$ c.s. cuando $N \rightarrow \infty$.

El teorema anterior será de gran ayuda cuando se analicen los problemas implementados, donde por lo general el compacto C será un subconjunto de $[0, 1]^n$.

Convergencia asintótica Dado que \hat{f}_N es condicional a la muestra escogida, para establecer un grado de convergencia, es crítico conocer su comportamiento asintótico como distribución. Considerando una muestra iid de $\{\xi_i\}_{i \in [N]}$, y una solución $x \in \mathcal{P}$, sabemos que el estimador $\hat{f}_N(x) = \frac{1}{N} \sum_{i=1}^N F(x, \xi_i)$ de $f(x) = \mathbb{E}[F(x, \xi)]$ es insesgado y tiene varianza $\frac{\sigma^2(x)}{N}$, donde se asume $\sigma^2(x) := \mathbb{V}[F(x, \xi)]$ acotado. Adicionalmente, por el Teorema Central del Límite (TCL) se tiene que:

$$N^{\frac{1}{2}} [\hat{f}_N(x) - f(x)] \xrightarrow{\mathcal{D}} Y_x \quad (2.18)$$

donde $\xrightarrow{\mathcal{D}}$ denota convergencia en distribución, e $Y_x \sim \mathcal{N}(0, \sigma^2(x))$, es decir $\hat{f}_N(x)$ tiene asintóticamente distribución $\mathcal{N}(f(x), \frac{\sigma^2(x)}{N})$.

El resultado asintótico anterior es extremadamente importante, dado que permite definir para un N suficientemente grande, el siguiente intervalo de confianza para $f(x)$:

$$\left[\hat{f}_N(x) - z_{\alpha/2} \frac{\hat{\sigma}(x)}{\sqrt{N}}, \hat{f}_N(x) + z_{\alpha/2} \frac{\hat{\sigma}(x)}{\sqrt{N}} \right] \quad (2.19)$$

donde $z_{\alpha/2} := \Phi^{-1}(1 - \alpha/2)$ define el corte por el nivel de confianza, y $\hat{\sigma}^2$ es el estimador muestral de la varianza $\sigma^2(x)$.

$$\hat{\sigma}^2 := \frac{1}{N-1} \sum_{i=1}^N [F(x, \xi_i) - \hat{f}_N(x)]^2 \quad (2.20)$$

Un aspecto importante a tener presente, referente a la convergencia de \hat{f}_N hace referencia a la dirección de la misma, donde tenemos que $\forall x' \in \mathcal{P}$ se tiene que $\hat{f}_N(x') \geq \inf_{x \in \mathcal{P}} \hat{f}_N(x)$, de donde tomando valor esperado y empujando el lado izquierdo minimizando, obtenemos:

$$\hat{\vartheta}^* = \inf_{x \in \mathcal{P}} f(x) = \inf_{x \in \mathcal{P}} \mathbb{E}[\hat{f}_N(x)] \geq \mathbb{E}[\inf_{x \in \mathcal{P}} \hat{f}_N(x)] = \mathbb{E}[\hat{\vartheta}_N^*] \quad (2.21)$$

Este resultado es extremadamente útil, y será utilizado más adelante para la definición de la cota inferior del gap de optimalidad.

Validación de una solución Un problema no tratado aún hace referencia a dado un candidato a solución $\bar{x} \in \mathcal{P}$, obtenido con la aproximación SAA, conocer su gap de optimalidad, es decir, dado un nivel de confianza el posible error en el valor óptimo, inducido por el método aproximado de resolución, dado el carácter aleatorio del problema.

Como elemento ideal, tenemos que dado un candidato \bar{x} su gap será $gap(\bar{x}) := f(\bar{x}) - \hat{\vartheta}^*$, no obstante, dado que a priori desconocemos el valor real de $\hat{\vartheta}^*$, debemos estimarlo vía cotas estadísticamente válidas.

Cota Inferior Utilizando la relación mostrada en 2.21, se tiene que $\hat{\vartheta}^* \geq \mathbb{E}[\hat{\vartheta}_N^*]$, es decir $\hat{\vartheta}_N^*$ entrega una cota inferior del valor de $\hat{\vartheta}^*$, utilizando como estimador de la esperanza de $\hat{\vartheta}_N^*$, el promedio de M realizaciones independientes, con N muestras iid en cada repetición, generando

$\{\hat{\vartheta}_{N,j}^*\}_{j \in \{1, \dots, M\}}$ valores óptimos de cada repetición. De lo anterior, tenemos como estimador insesgado de $\mathbb{E}[\hat{\vartheta}_N^*]$ a:

$$\bar{v}_{N,M} = \frac{1}{M} \sum_{r=1}^M \hat{\vartheta}_{N,r}^* \quad (2.22)$$

Como las muestras son iid, y por ello los valores $\{\hat{\vartheta}_{N,j}^*\}_{j \in \{1, \dots, M\}}$, se tiene que podemos estimar su varianza por:

$$\hat{\sigma}_{N,M}^2 := \frac{1}{M} \left[\frac{1}{M-1} \sum_{m=1}^M (\hat{\vartheta}_{N,m}^* - \bar{v}_{N,M})^2 \right] \quad (2.23)$$

Por el TCL la distribución de probabilidad de $\bar{v}_{N,M}$ se hace aproximadamente normal con un M suficientemente grande, en consecuencia podemos usar de forma conservadora a un nivel de confianza α , como cota inferior (LB):

$$LB_{N,M}^\alpha := \bar{v}_{N,M} - z_\alpha \hat{\sigma}_{N,M} \quad (2.24)$$

Cota Superior Como cota superior, dado resolvemos un problema de minimización, podemos utilizar una estimación de $f(\bar{x})$ con un muestreo N' , con $N' \gg N$ (en el sentido de mucho mayor) buscando aproximar la integral, completamente iid al utilizado para computar $LB_{N,M}$, y por ello \bar{x} , tal que el $\hat{f}_{N'}(\bar{x})$ sea el estimador muestral del promedio, con estimador de varianza muestral $\hat{\sigma}_{N'}^2$:

$$\hat{\sigma}_{N'}^2(\bar{x}) := \frac{1}{N'-1} \sum_{j=1}^{N'} [F(\bar{x}, \xi_j) - \hat{f}_{N'}(\bar{x})]^2 \quad (2.25)$$

Como las muestras anteriores son iid, y por el TCL se obtiene que la cota superior (UB) propuesta es:

$$UB_{N'}^\alpha(\bar{x}) := \hat{f}_{N'}(\bar{x}) + \frac{z_\alpha}{\sqrt{N'}} \hat{\sigma}_{N'}(\bar{x}) \quad (2.26)$$

Gap de optimalidad Como gap de optimalidad, se entiende el posible error en el valor óptimo candidato, donde de los resultados anteriores, dado un nivel de confianza α y tamaños muestrales suficientes N , N' y repeticiones M se cumple que $UB_{N'}^\alpha(\bar{x}) \geq LB_{M,N}^\alpha$, y en particular $[UB_{N'}^\alpha(\bar{x}), LB_{M,N}^\alpha]$ es el intervalo donde puede encontrarse ϑ^* con al menos un α nivel de confianza, por ello se define como *gap absoluto* al tamaño de este intervalo, es decir:

$$gap_A(\bar{x}) := UB_{N'}^\alpha(\bar{x}) - LB_{M,N}^\alpha \quad (2.27)$$

adicionalmente para tener una noción relativa porcentual, se define el concepto de *gap relativo*, relativizando esta magnitud al valor absoluto de la cota inferior, es decir:

$$gap_R(\bar{x}) := \frac{gap_A(\bar{x})}{|LB_{M,N}^\alpha|} \quad (2.28)$$

donde en realidad, por asuntos computacionales se relativiza con respecto a $\max\{10^{-6}, |LB_{M,N}^\alpha|\}$ para evitar dividir por cero.

Al trabajar con magnitudes muy cercanas a cero, o idénticamente iguales a cero, estas medidas de convergencia quedan mal definidas, distorsionando la medida del gap relativo, mostrando sólo un juego de velocidades de convergencia a cero.

Para solucionar esto, utilizando la propiedad de traslación de las medidas de riesgo, se define el *gap relativo trasladado* ($gap_{R,L}(\bar{x})$) como el gap relativo de la medida, calculando la misma con un desplazamiento en L , para posteriormente corregir el mismo, es decir: $\rho(-\vec{w} \cdot \vec{r})$ a través de su traslación $\rho(-\vec{w} \cdot \vec{r} + L) - L$ donde L es escogido tal que $LB_{M,N}^\alpha + L$ se normalice a 100, para conservar su interpretación porcentual. Es así como se define:

$$gap_{R,L}(\bar{x}) := \frac{(UB_{N'}^\alpha(\bar{x}) + L) - (LB_{M,N}^\alpha + L)}{|LB_{M,N}^\alpha + L|} = \frac{gap_A(\bar{x})}{|LB_{M,N}^\alpha + L|} \quad (2.29)$$

Es importante observar que un desplazamiento de la variable aleatoria en L , es lineal tanto en la cota superior como en la inferior, eliminando el efecto de la traslación en el gap absoluto.

Algoritmo de búsqueda y validación de soluciones SAA Los elementos definidos anteriormente, entregan implícitamente un algoritmo para la resolución de problemas que pueden ser escritos como un valor esperado, como se dijo anteriormente problemas con la estructura de 2.14. Este algoritmo para la utilización de SAA, permite entregar el menor gap como garantía de optimalidad, y en particular el mejor candidato solución \bar{x} y su valor óptimo asociado, dado tamaños específicos de realizaciones.

Dado M repeticiones, con N muestras iid cada repetición, y un adicional de N' realizaciones adicionales iid para la validación de resultados, a un nivel de confianza $\alpha \in [0, 1)$, el algoritmo es el siguiente:

Un aspecto importante a señalar en este punto, hace referencia a la composición del gap estocástico, y el control que se puede ejercer en estas diferentes componentes con los tamaños muestrales M , N y N' . Podemos descomponer el gap estocástico en tres partes: (1) el primero de ellos asociado a la incertidumbre de la cota superior, y su estimación (2) el segundo de ellos, asociado a la estimación de la cota inferior con realizaciones del problema cierto, y finalmente (3) el tercero de ellos, asociado a reducir el *expected gap*, es decir la distancia entre los valores esperados de ambas cotas.

Por un lado, es claro que N' reduce la incertidumbre del estimador de la cota superior, reduciendo directamente $\hat{\sigma}_{N'}^2(\bar{x}_r)$ y por ello el primer componente del gap. Adicionalmente, M y N controlan la dispersión del estimador de la cota inferior, afectando directamente la segunda componente del gap por $\hat{\sigma}_{N,M}^2$, e indirectamente la tercera componente, a través de nuevos candidatos a solución (M) y una mayor calidad en cada solución (N).

Algorithm 1 Utilización de SAA

1. Con las M repeticiones de tamaño N , con las $\{\hat{\xi}_i^r\}_{r \in [M]}^{i \in [N]}$ realizaciones de la variable incierta, se resuelven los M problemas ciertos equivalentes:

$$\hat{\vartheta}_{N,r}^* := \min_{x \in \mathcal{P}} [\hat{f}_N(x) := \frac{1}{N} \sum_{i=1}^N F(x, \xi_i^r)]$$

Se obtienen los M candidatos a solución \bar{x}_r , y candidatos a valores óptimos $\hat{\vartheta}_{N,r}^*$ con $r \in [M]$

2. Con los elementos $\{\hat{\vartheta}_{N,r}^*\}_{r \in [M]}$ obtenidos anteriormente, se obtiene la cota inferior para $\hat{\vartheta}^*$, a través del calculo de $\bar{v}_{N,M}$, $\hat{\sigma}_{N,M}^2$ y finalmente $LB_{N,M}^\alpha$.
 3. Sabemos que todo punto $x \in \mathcal{P}$, es una cota superior valida para el problema, en particular buscamos la mejor (con menor gap inducido), de los M candidatos \bar{x}_r generados en el paso (1), para este fin, con el sample de tamaño N' iid generamos las cotas superiores calculando $\hat{f}_{N'}(\bar{x}_r)$, $\hat{\sigma}_{N'}^2(\bar{x}_r)$ y finalmente $UB_{N'}^\alpha(\bar{x}_r)$.
 4. Del paso anterior, se escoge $r' \in [M]$ tal que induce la menor cota superior (menor gap estocástico), entregando así el candidato a solución, adicionalmente con $UB_{N'}^\alpha(\bar{x}_{r'})$ y $LB_{N,M}^\alpha$ se pueden obtener las medidas de convergencia a través de $gap_A(\bar{x}_{r'})$, $gap_R(\bar{x}_{r'})$ y $gap_{R,L}(\bar{x}_{r'})$.
 5. De buscar una mejor convergencia, se deben ajustar los tamaños muestrales, volviendo al paso (1) con los nuevos tamaños de realizaciones.
-

Capítulo 3

Modelos y metodologías desarrolladas

Concluida la sección de medidas de riesgo, se introduce y desarrolla la metodología utilizada para resolver los problemas de optimización estocástica planteados, con especial cuidado en el caso de la medida de riesgo entrópica, donde dada la naturaleza no lineal de la misma, se diseña e implementa un algoritmo de aproximación vía generación y ajuste de cortes.

Posteriormente, se desarrollan con detalle los diferentes modelos, por medida de riesgo, para finalmente concluir con la propuesta metodológica desarrollada e implementada para la resolución de la medida de riesgo entrópica.

Con todos los conceptos antes introducidos, en esta sección se presentan los modelos desarrollados y sus transformaciones en el contexto de portafolios y su aproximación vía SAA, así como en el caso particular de la medida de riesgo entrópica, el algoritmo para su resolución vía aproximación por generación de cortes, y ajuste iterativo de los soportes de los cortes para la reducción del error de aproximación.

3.1. SAA aplicado a los problemas de portafolio

En el contexto de portafolios, la incertidumbre proviene del vector de retornos futuros \vec{r} . La variable a resolver x es el vector de composiciones factibles $\vec{w} \in \mathcal{P}$. Como se dijo anteriormente en 2.1, se entenderá \mathcal{P} como el poliedro que cumple las restricciones de inversión: $\vec{a}_j \cdot \vec{w} \leq b_j$ para cada restricción $j \in R$, y que cumple la restricción de inversión total: $\vec{1} \cdot \vec{w} = 1$ sin venta corta $\vec{w} \geq \vec{0}$.

Como función de pérdidas, trabajamos con la variable $-\vec{w} \cdot \vec{r}$ que representa los retornos negativos de un portafolio, de este modo las pérdidas se encuentran en el lado derecho de la distribución de probabilidad.

3.1.1. Aproximación del problema de portafolio con valor esperado

En el caso del valor esperado, el problema aproximado vía SAA resuelto en cada repetición, asumiendo N samples iid, es el siguiente:

$$\hat{\vartheta}_{N,r}^* := \min_{w \in \mathcal{P}} \frac{1}{N} \sum_{i \in [N]} -\vec{w} \cdot \hat{r}_i^r \quad (3.1)$$

3.1.2. Aproximación del problema de portafolio con CVaR

En el caso del valor condicional en riesgo, como se dijo anteriormente en 2.2.2, por el trabajo de [72], se ocupa su formulación equivalente discreta, que es una linealización vía variables auxiliares positivas, resolviendo finalmente:

$$\begin{aligned} \hat{\vartheta}_{N,r}^* := \min_{\vec{w}, t} \quad & t + \frac{1}{N(1-\alpha)} \sum_{i \in \{1, \dots, N\}} v_i \\ & v_i \geq -\vec{w} \cdot \hat{r}_i^r - t \quad \forall i \in \{1, \dots, N\} \\ & v_i \geq 0 \quad \forall i \in \{1, \dots, N\} \\ & \vec{w} \in \mathcal{P} \end{aligned} \quad (3.2)$$

Es importante acotar, que es conocido que la resolución de este problema es más costoso y demoroso que resolver su formulación dual, o versiones aceleradas de la misma como es lo presentado en [28].

3.1.3. Aproximación del problema de portafolio con ECVaR

El problema de optimización de la combinación convexa entre valor esperado y CVaR, corresponde a una variante del problema resuelto con CVaR, siendo el modelo a resolver el siguiente:

$$\begin{aligned} \hat{\vartheta}_{N,r}^* := \min_{\vec{w}, t} \quad & \frac{1-\epsilon}{N(1-\alpha)} (tN(1-\alpha) + \sum_{i \in \{1, \dots, N\}} v_i) - \frac{\epsilon}{N} \sum_{i \in \{1, \dots, N\}} \vec{w} \cdot \hat{r}_i^r \\ & v_i \geq -\vec{w} \cdot \hat{r}_i^r - t \quad \forall i \in \{1, \dots, N\} \\ & v_i \geq 0 \quad \forall i \in \{1, \dots, N\} \\ & \vec{w} \in \mathcal{P} \end{aligned} \quad (3.3)$$

3.2. Aproximación del problema de portafolio con la Medida Entrópica

En el caso de la medida de riesgo entrópica, fue necesario realizar transformaciones por aspectos computacionales en su implementación.

La primera transformación que se realiza es la omisión de $\frac{1}{\theta} \log(\cdot)$ dado que lo anterior es una función creciente y no altera el argumento de la solución óptima, no obstante facilita su tratamiento, y permite su tratamiento por SAA. El trabajar con funciones exponenciales, presenta problema de escala ante exponentes muy grande (o pequeños), por esta razón se escala el problema (función objetivo) en un factor K , propagando esta modificación en lo desarrollado en SAA.

Finalmente, es necesario realizar una aproximación vía cortes para aproximar la exponencial involucrada en el modelo, este aspecto es cubierto con detalle de la siguiente subsección, no obstante, en aspectos generales el algoritmo en la selección del soporte de cortes, esta compuesta por tres etapas: (1) Inicialmente se utilizan puntos pre-establecidos para el soporte de los cortes, posteriormente desde la solución anterior (2) se insertan iterativamente minimizando el peor error de aproximación y (3) finalmente se evalúan mejoras moviendo los puntos de aproximación en sus vecindades con una búsqueda local. Esta última etapa de aproximación por cortes realiza iterativamente un ajuste de K , imponiendo que los puntos de interés se encuentren en el intervalo de aproximación impuesto.

El problema de optimización que se busca resolver es:

$$\min_{w \in \mathcal{P}} \frac{1}{\theta} \log(\mathbb{E}[e^{-\theta \bar{w} \cdot \bar{r}}])$$

este problema en solución es equivalente a resolver:

$$\min_{w \in \mathcal{P}} \mathbb{E}[e^{-\theta \bar{w} \cdot \bar{r}}]$$

que en su versión aproximada vía SAA, considerando un un sample iid de tamaño N de \bar{r} , y escalando el problema en θK , que a lo más es $\theta \bar{K}$, para controlar el tamaño de $-\theta \bar{w} \cdot \bar{r}$ evitando problemas numéricos, se obtiene:

$$e^{\theta K} \min_{w \in \mathcal{P}} \frac{1}{N} \sum_{i \in [N]} e^{-\theta(\bar{w} \cdot \hat{r}_i + K)}$$

Escribiendo el problema anterior, con variable auxiliares reales (u_i) y positivas (v_i), para explicitar cada restricción exponencial a aproximar, obtenemos:

$$\begin{aligned} \min_{\bar{w} \in \mathcal{P}, \bar{v}, \bar{u}} \quad & \sum_{i \in [N]} v_i \\ u_i \quad & \geq -\theta(\bar{w} \cdot \hat{r}_i + K) \quad \forall i \in [N] \\ v_i \quad & \geq 0 \\ Nv_i \quad & \geq e^{u_i} \quad \forall i \in [N] \end{aligned}$$

Donde observamos que la última restricción, resume la no-linealidad exponencial del problema, la que será aproximada de forma lineal por partes.

3.2.1. Aproximación vía subgradiente del problema de optimización entrópico

La idea central de la generación de cortes para aproximar las restricciones donde se involucra la función no lineal, es una idea ampliamente explotada, permite linealizar localmente la zona

Diagrama explicativo de la aproximación de la exponencial por cortes

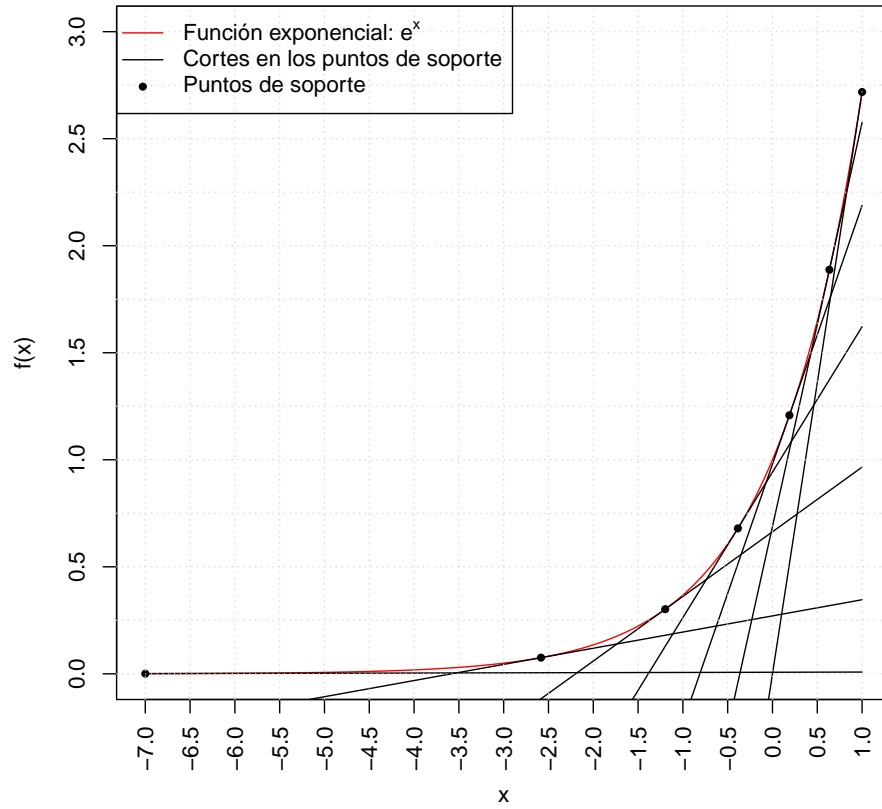


Figura 3.1: Ilustración de aproximación de exponencial vía cortes

óptima, la complejidad que enfrenta este problema es reconocer en que área factible se encuentra la zona a linealizar con mayor precisión, en este caso, nos referimos como zona óptima al conjunto de variables \vec{w} y K , que dado una realización iid definen $u_i(\vec{w}, K)$, y por ello $e^{u_i(\vec{w}, K)}$. Para tener una idea ilustrativa más clara, se muestra en la figura 3.1 el enfoque aplicado con siete cortes.

Un elemento crítico que define a los cortes, son sus puntos de soporte. Consideremos el conjunto de puntos de apoyo $U := \{\rho_1, \rho_2, \dots, \rho_k\}$ con k elementos. Para el caso de funciones convexas diferenciables, se tiene que la aproximación subgradiente inferior es $f(x) - f(x_0) \geq \partial f(x_0) \cdot (x - x_0)$ con $\partial f(x_0) = \frac{df}{dx}(x_0)$, dada la diferenciable de $f(\cdot)$. Aplicando lo anterior a la exponencial, se obtiene la forma funcional de la aproximación buscada, centrando el problema sólo en encontrar los adecuados puntos de soporte para los cortes.

De lo anterior, con los $\{\rho_j\}_{j \in k}$ soportes, se aproximan las N restricciones exponenciales, por Nk restricciones de la forma:

$$\begin{aligned} v_i &\geq e^{u_i} \geq e^{\rho_j} (u_i - \rho_j) + e^{\rho_j} & \forall (i, j) \in [N] \times [k] \\ v_i &\geq e^{\rho_j} (u_i - \rho_j) + e^{\rho_j} & \forall (i, j) \in [N] \times [k] \end{aligned}$$

Es decir:

$$v_i - e^{\rho_j} u_i \geq (1 - \rho_j) e^{\rho_j} \quad \forall (i, j) \in [N] \times [k]$$

Con los cortes para la aproximación de la exponencial, el problema a resolver en su forma primal, considerando las restricciones que definen \mathcal{P} resumidas en la matriz y vector (A, \vec{b}) , es el siguiente:

$$\begin{aligned}
& \underset{\vec{w} \in \mathcal{P}, K, \vec{v}, \vec{u}}{\text{mín}} && \sum_{i \in [N]} v_i \\
& u_i + \theta(\vec{w} \cdot \hat{r}_i + K) & \geq & 0 \quad \forall i \in [N] \\
& K & \leq & \bar{K} \\
& v_i - e^{\rho_j} u_i & \geq & (1 - \rho_j) e^{\rho_j} \quad \forall (i, j) \in [N] \times [k] \\
& A^t \vec{w} & \leq & \vec{b} \\
& v_i & \geq & 0 \quad \forall i \in [N] \\
& \vec{w} & \geq & 0
\end{aligned}$$

Este problema tiene $N(1+k) + 1 + \text{rang}(A)$ (entendiendo $\text{rang}(A)$ como el máximo número de filas linealmente independientes en la matriz A) restricciones y $|I| + 1 + 2N$ variables, de las cuales $N + 1$ se pueden fijar imponiendo igualdad en la primera y segunda restricción, por ello, reconociendo que $|I| \ll N$ y que el desempeño empírico de Simplex implementado en CPLEX (con pre-rutinas previas), funciona en ordenes de magnitud lineal al número de restricciones, y raíz cuadrado en el número de variables, para mayor detalle sobre esto revisar [19] y [92], se pasa el problema a su formulación dual.

La formulación dual del problema anterior, sustituyendo las variables que se despejan de las restricciones duales, asociadas a las variables primales libres (K y u_i), es la siguiente:

$$\begin{aligned}
& \underset{\lambda_{(i,j)} \forall (i,j) \in [N] \times [k], \lambda_0, \vec{y}}{\text{máx}} && \sum_{(i,j) \in [N] \times [k]} \lambda_{(i,j)} (1 - \rho_j - \bar{K}) e^{\rho_j} + \vec{b} \cdot \vec{y} \\
& && \sum_{j \in [k]} \lambda_{(i,j)} \leq \frac{1}{N} \forall i \in [N] \\
& A^t \vec{y} + \theta \sum_{(i,j) \in [N] \times [k]} \hat{r}_i e^{\rho_j} \lambda_{(i,j)} \leq \vec{0}
\end{aligned}$$

Este problema fue el implementado para su resolución, se observa consistente a lo discutido anteriormente que este problema tiene $Nk + \text{rang}(A)$ variables y sólo $N + |I|$ restricciones.

Un problema aún no tratado de esta formulación, es el conjunto de puntos de soporte $U := \{\rho_j\}_{j \in [k]}$ a utilizar por los cortes, lo que será tratado a continuación.

3.2.2. Algoritmo de generación y ajuste de cortes de aproximación

Los cortes buscan aproximar la exponencial de la mejor forma posible en la región de interés, entendiendo esto último como minimizando el error de aproximación de $e^{u_i(\vec{w}, K)}$ dado \vec{w} , K y un sample de realización $\{\hat{r}_i\}_{i \in [N]}$. De este punto de vista, el problema que se esta buscando resolver dados \vec{w} , K y $\{\hat{r}_i\}_{i \in [N]}$ es:

$$\begin{aligned} \min_{\rho_j \quad j \in [k]} \quad & e^{u_i(\vec{w}, K)} - v_i \\ v_i - e^{\rho_j} u_i(\vec{w}, K) \geq & (1 - \rho_j) e^{\rho_j} \quad \forall (i, j) \in [N] \times [k] \\ v_i \geq & 0 \quad \forall i \in [N] \end{aligned}$$

$$\text{con } u_i(\vec{w}, K) = -\theta(\vec{w} \cdot \hat{r}_i + K) \quad \forall i \in [N]$$

Para enfrentar este problema, en vez de resolver el problema anterior, se diseña e implementa un algoritmo de tres pasos. La lógica central del algoritmo para generar y ajustan los puntos de soporte de los cortes, involucra el ajuste de \bar{K} para escalar y asegurar siempre estar en el intervalo de aproximación buscado, ajustar U con los nuevos puntos de soporte, y volver resolver el problema de optimización para obtener \vec{w} (y por ello $u_i(\vec{w}, K)$) candidatos de la iteración, y observar si la solución cambia, al cambiar los puntos de soporte. Todos los elementos anteriores, serán descritos con cuidado a continuación.

Como fue introducido anteriormente K es utilizado para ajustar la escala, idea que toma especial importancia al estar aproximando la exponencial en un intervalo fijo $[L_{min}, L_{max}]$. Con lo anterior, y dado el problema es de minimización, en el proceso de ajuste de \bar{K} se impone que se cumpla la condición de que: dada la muestra iid de la repetición $\{\hat{r}_i\}_{i \in [N]}$, K sea tal que $\max_{i \in [N]} u_i \in B(L_{max}, 10^{-6})$, lo que es logrado ajustando \bar{K} vía resolución de problemas e iteración con búsqueda binaria sobre \bar{K} .

Para efectos prácticos en la implementación se asume el intervalo $[-7, 1]$. Un aspecto importante a observar, es la importancia de dejar todos los puntos u_i en o por debajo de L_{max} , sin ser un elemento crítico la cota inferior L_{min} , dada la dirección de minimización, y por ser considerados irrelevantes puntos con un escalamiento relativo inferior a e^{-7} .

Para el algoritmo se considera dado el número de puntos del soporte, por lo que dado un número $|U| = k$ de puntos de aproximación, en el primer paso del algoritmo se utilizan puntos iniciales pre-definidos fijos.

Para la generación de los puntos iniciales fijos, se resuelve el problema de optimización para diferentes número de puntos de soporte, considerando el error total de aproximación de ambas funciones, es decir, como si los u_i se distribuyen homogéneamente en $[L_{min}, L_{max}]$, para facilitar ideas:

$$\min_{\rho_j \quad j \in [k]} \int_{L_{min}}^{L_{max}} (e^x - L(x, \vec{\rho}_j)) dx \quad (3.4)$$

con $L(x, \vec{\rho}_j) = \max_{j \in [k]} \{L(x, \rho_j)\}$ y $L(x, \rho_j)$ el corte generado por el subgradiente de la exponencial en el punto ρ_j , es decir $L(x, \rho_j) = e^{\rho_j}(1 - \rho_j + x)$.

Inicialmente se busca resolver el problema anterior de forma analítica, para lo anterior se utiliza que la función $L(x, \vec{\rho}_j)$ que es lineal por partes en x , con lo que el problema anterior es separable por los puntos donde se intersecan cortes sucesivos. Dado dos cortes sucesivos con soportes ρ_i y ρ_j , el punto $t(i, j) := t(\rho_i, \rho_j) := \frac{e^{\rho_j(1-\rho_j)} - e^{\rho_i(1-\rho_i)}}{(e^{\rho_i} - e^{\rho_j})}$ es el quiebre en la función lineal por partes. Para facilitar la notación se extiende la definición de $t_{i,j}$ en los casos de borde a:

$$t(i, j) := \begin{cases} L_{min} & \text{si } i < 0 \\ \frac{e^{\rho_j(1-\rho_j)} - e^{\rho_i(1-\rho_i)}}{(e^{\rho_i} - e^{\rho_j})} & \text{si } x \geq 0 \\ L_{max} & \text{si } j > k - 1 \end{cases}$$

Con lo que el problema 3.4 puede ser escrito como:

$$\min_{\rho_j \quad j \in [k]} \left\{ \sum_{i \in [k]} \int_{t(i-1, i)}^{t(i, i+1)} (e^x - L(x, \rho_j)) dx \right\} \quad (3.5)$$

Desarrollada la expresión anterior, se constata que la misma no es lineal ni convexa en $\{\rho_i\}_{i \in [k]}$, en particular, para el caso con cuatro puntos, la función de error para el tercer y cuarto punto, ya fijados de forma óptima secuencial los dos y tres puntos anteriores respectivamente, es la que se muestra en la figura 3.2.

Es relevante observar que los cortes tienen mayor concentración en el lado derecho del intervalo de aproximación, comportamiento coherente a los esperado con un crecimiento exponencial y minimización de error absoluto, donde se enfoca en la zona de mayor crecimiento.

En la implementación, se escoge para independizar la aplicación de un archivo de puntos externos, utilizar una aproximación de la ubicación de los puntos. Para este fin se escoge una distribución parabólica (pre imagen de una cuadrática), por su saturación en entorno al punto final, ajustando la misma tal que el punto inicial sea L_{min} , y el punto final L_{max} , de lo anterior se obtiene la secuencia inicial para los k puntos iniciales de soporte:

$$\rho_j = (L_{max} - L_{min}) \sqrt{\frac{i}{k-1}} + L_{min} \quad \forall j \in [k] := \{0, \dots, k-1\}$$

Donde se verifica que $\rho_0 = L_{min}$, $\rho_{k-1} = L_{max}$ y $\lim_{j \rightarrow \infty} \Delta \rho_{j+1} = \rho_{j+1} - \rho_j = 0$.

Este modelo es escogido de un conjunto mayor, por su ajuste con las soluciones numéricas obtenidas anteriormente, y por su simplicidad de cálculo. Es importante destacar que dadas las tres etapas del algoritmo: la inicial y dos de ajuste, y considerando que lo relevante es ajustar la aproximación dependiente del sample iid de la realización, y no de una distribución de puntos homogénea, partir de puntos cercanos a un óptimo puede reducir el número de iteraciones de ajuste,

Evolución iterativa del error con dos y tres puntos fijos

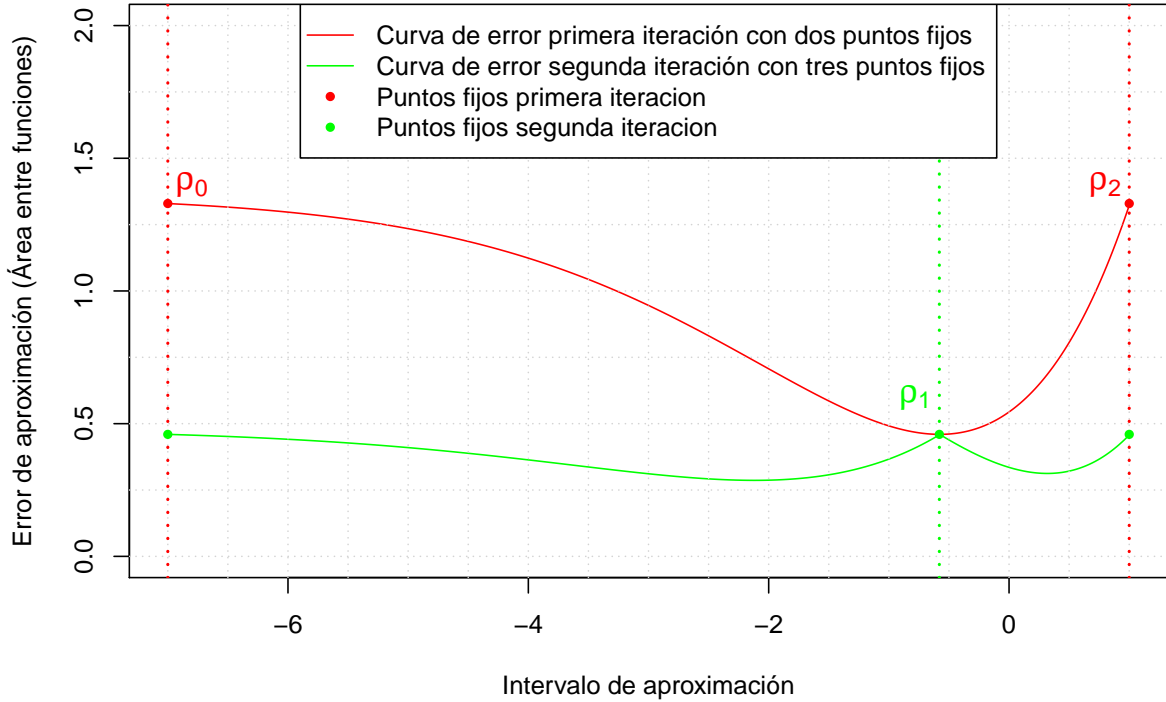


Figura 3.2: Función de error de aproximación, para un cuarto punto, ya fijados tres anteriores de forma óptima

pero en ningún caso es exigencia de optimalidad en los puntos iniciales, para la adecuada resolución de las siguientes dos etapas.

En el gráfico 3.3 se puede observar para diferente número de puntos, los soportes obtenidos numéricamente y los generados por el ajuste.

Un segundo elemento a considerar, es el número de puntos a utilizar para aproximar, como se observa en la figura 3.4, se observa un fuerte decaimiento del error condicional al número de cortes, siendo razonable un punto de corte entorno a los diez cortes, no obstante, por seguridad se escoge aproximar con veinte puntos.

Una vez cargados los puntos iniciales pre-fijados, se resuelve el problema de optimización iterativamente, ajustando \bar{K} hasta cumplir la condición $\max_{i \in [N]} u_i(\bar{w}^*, K^*) \in B(L_{max}, 10^{-6})$, cumplida esta condición, podemos asegurar que los $\{u_i\}_{i \in [N]}$ se encuentran en $(-\infty, L_{max}]$, donde la aproximación tiene sentido.

Con el adecuado K^* , y el candidato a solución \bar{w}^* se computan los valores de $\{u_i(\bar{w}^*, K^*)\}_{i \in [N]}$, los que serán utilizados en las restantes dos etapas sucesivas de ajuste del soporte $\{\rho_j\}_{j \in [k]}$. De las restantes dos etapas de ajuste, en la primera de ellas de forma iterativa, se define la nueva ubicación de un punto del soporte, insertando este último en el punto medio del *bin*, de una división en 2^{16} *bins* del intervalo de búsqueda $[L_{min}, L_{max}]$, donde se tiene el máximo error de aproximación de la exponencial con la función lineal por partes $L(U', x)$, la que es construida iterativamente agregando

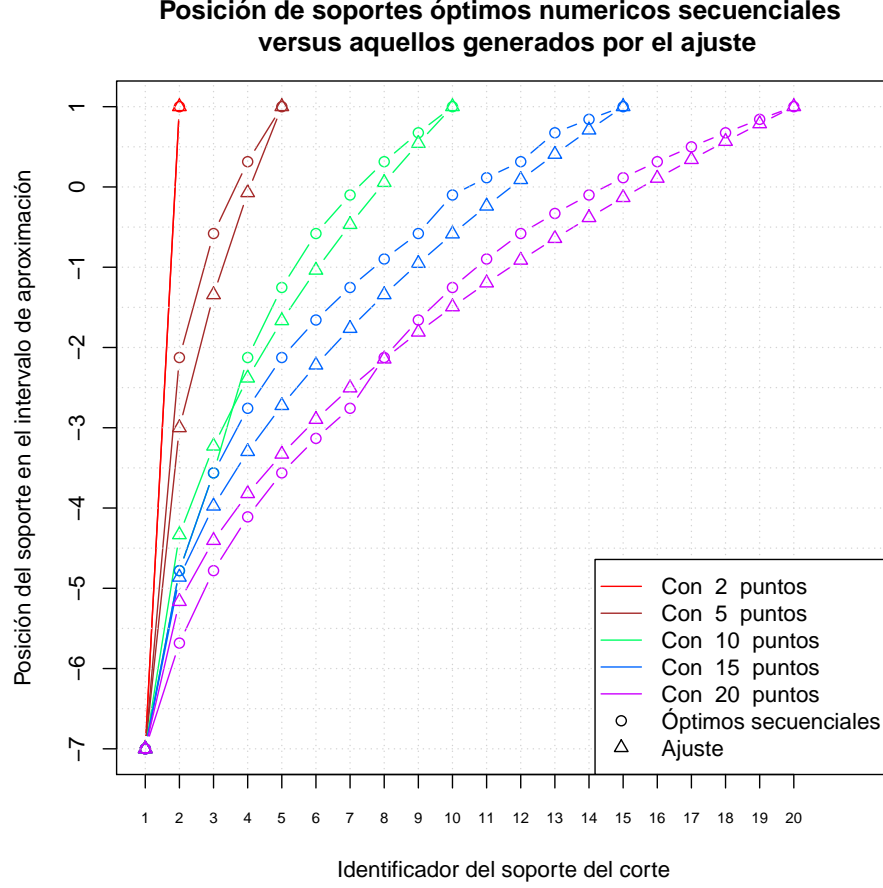


Figura 3.3: Puntos de soporte obtenidos por ajuste (triángulos) y obtenidos numéricamente (círculos); para diferente número de soportes

los nuevos puntos soportes al conjunto U' , partiendo este último vacío.

Se define el error de aproximación (AE), dado un *bin* b de límites $I_b = [l_b, u_b]$, y el conjunto $N_b := \{u_i : u_i \in I_b\}$, como:

$$AE_b(U', \{u_i\}_{i \in [N]}) = \sum_{u_i \in N_b} e^{u_i} - L(U', \{u_i\}_{i \in [N]}, u_i) \quad (3.6)$$

donde idéntico a lo desarrollado en la primera parte, para encontrar los puntos iniciales, pero esta vez considerando la real distribución de los puntos $\{u_i\}_{i \in [N]}$, se define la función lineal por partes $L(U', x) := \max\{\max_{j \in [U']} l_j(u_i), 0\}$ inducida por los cortes $l_j(u_i) := e^{\rho_j(1 - \rho_j + u_i)}$ con los nuevos puntos del conjunto U' , donde en el caso de $U' = \emptyset$ se define $L(\cdot) := 0$.

En la implementación del método anterior, dado un punto u_i , se generó una rutina que identifica el corte activo en la función lineal por partes, computando sólo esa función lineal asociada. Adicionalmente, para garantizar tiempos de cómputo razonables, y explotar la estructura del problema (orden del error), se utilizó búsqueda binaria en la obtención del *bin* con máximo error.

Para facilitar la idea de la función lineal por partes, y la función de error de aproximación

Error absoluto de aproximación dependiendo del número de puntos

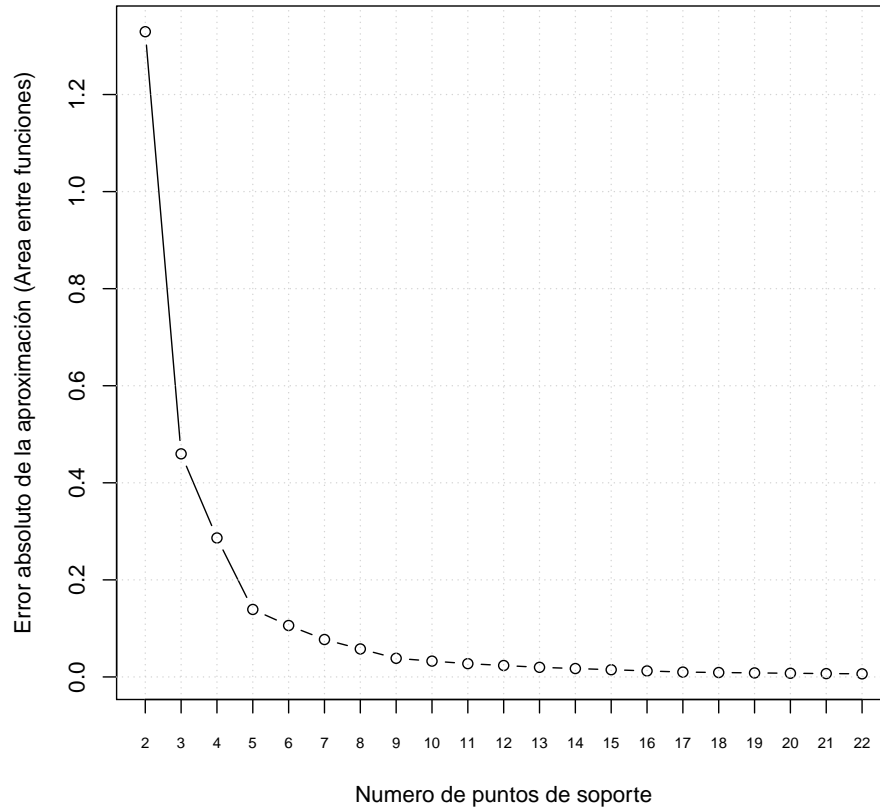


Figura 3.4: Error total de los puntos iniciales, medido como el área entre las funciones, dependiendo del número de puntos utilizados para aproximar

asociada, se deja un ejemplo en la figura 3.5, donde se utiliza el soporte inicial antes explicado, y un conjunto de u_i uniformemente distribuidos.

En la última etapa del algoritmo, para mejorar el ajuste de la función lineal por partes, se revisa localmente si existen mejoras en la función de error, revisando individualmente cada punto de soporte con la posibilidad de ser movido dentro del intervalo generado entre el punto de soporte anterior y el siguiente. Este proceso se detiene cuando posterior a una iteración completa sobre los k elementos, no existe un ajuste de al menos uno de ellos.

Concluido el proceso de ajuste de los puntos de soporte, con los K^* y \vec{w}^* iniciales, surge la posibilidad de que estas soluciones dado el cambio en la aproximación, ya no sean óptimos. Por lo anterior, con los nuevos puntos de aproximación, y ocupando como punto de partida el escalamiento anterior ($\bar{K} = K^*$), se vuelve a resolver el problema de optimización, obteniendo un nuevo factor de escala \bar{K}^* , y \vec{w}^* óptimo candidato con los que se puede volver a repetir el proceso de ajuste de los puntos del soporte, y así sucesivamente iteración a iteración.

Para concluir con la definición del algoritmo, es necesario definir la condición de término de los ciclos de optimización y ajuste. El ajuste de la grilla de aproximación afecta los valores de \vec{w}^* , lo que queda en evidencia al considerar que el conjunto de vértices del poliedro cambia. Dada una

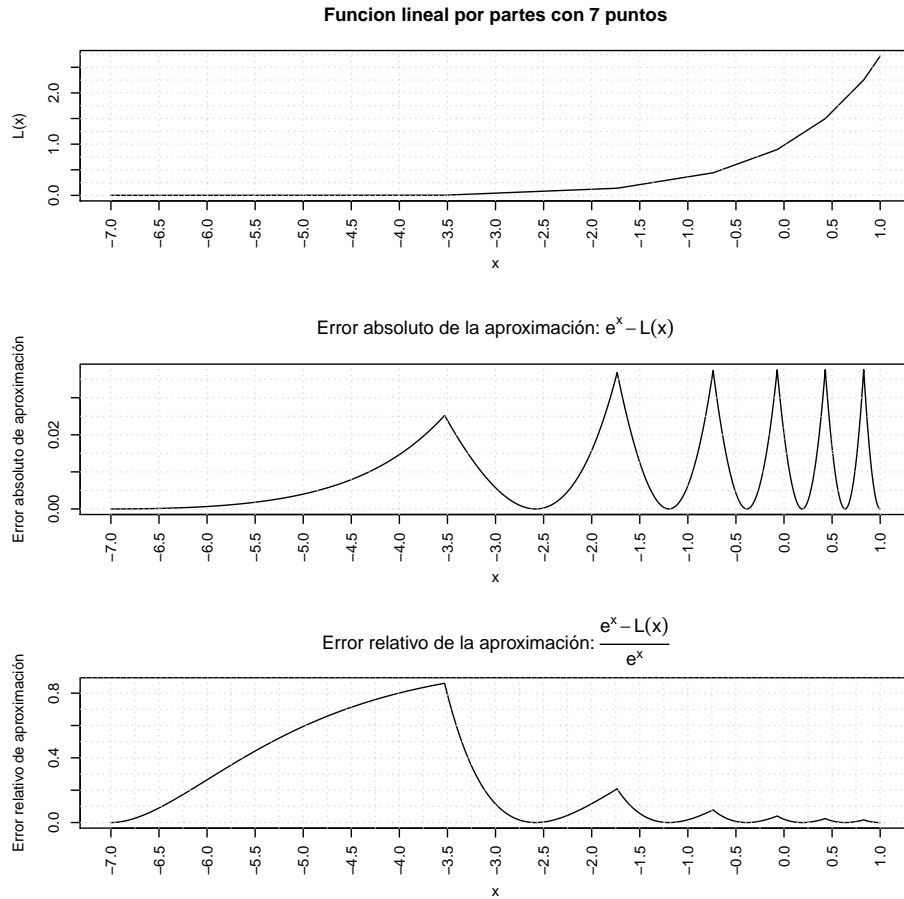


Figura 3.5: Ejemplo de función lineal por partes, que aproxima la exponencial, con su respectivo error de aproximación asociado.

misma muestra iid de una repetición, de existir un cambio en \vec{w}^* , en particular cambian los puntos $\{u_i(\vec{w}^*, K^*)\}_{i \in [N]}$, por lo que se considera un criterio de parada razonable, que entre dos iteraciones consecutivas, el factor de escala \bar{K} no cambie significativamente (10^{-6}).

Para tener una idea del funcionamiento de la rutina en términos del error de aproximación, considerando una instancia difícil de resolver dado que la mayoría de los elementos $\{u_i\}_{i \in [N]}$ quedan en el intervalo $[-7, 1]$, que es obtenida con retornos distribuidos como $\vec{r} \sim \mathcal{N}(\vec{0}, I)$ de cuatro instrumentos, con aversión de $\theta = 1$, samples de tamaño 10.000, considerando 20 cortes y 100 realizaciones, se obtienen los errores de aproximación por etapa, que se muestran en el gráfico 3.6.

De forma explicativa, el algoritmo iterativo de resolución, puede ser expresado de forma resumida en el diagrama 3.7. En este diagrama se aprecian los *estados* inicial y final con los cuadros de color rojos, *procesos* con los cuadros naranjos y las *verificaciones* los cuadros verde.

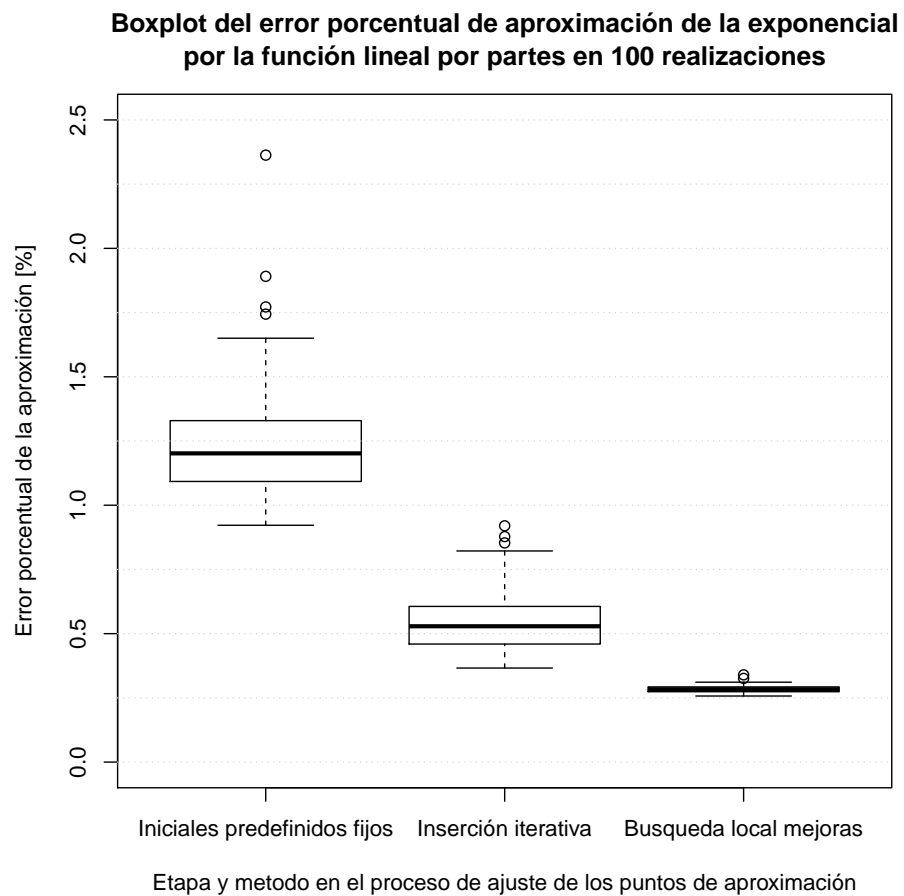


Figura 3.6: Boxplot con el error porcentual de aproximación de la función exponencial, evaluada en los puntos $\{u_i\}_{i \in [N]}$ con 10.000 realizaciones, considerando una instancia de prueba y 100 repeticiones iid.

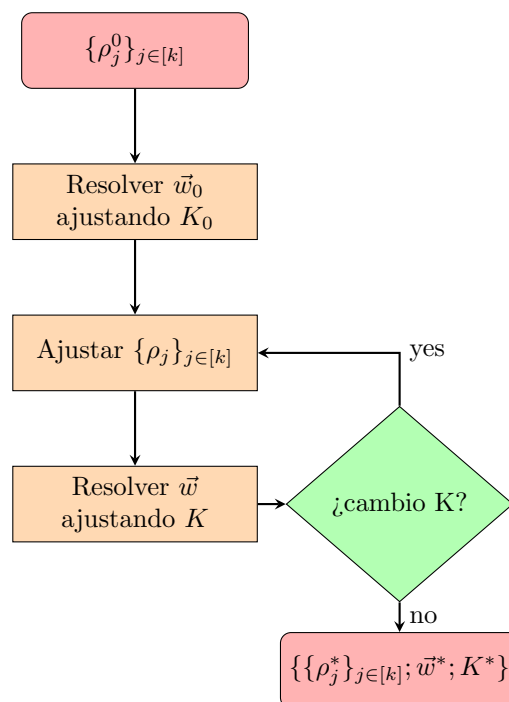


Figura 3.7: Diagrama resumen del algoritmo de aproximación, método de ajuste de los puntos de soportes y resolución.

Capítulo 4

Experimentos computacionales

En este capítulo se realizan los experimentos computacionales, con el fin de dar respuesta a las preguntas exploradas. Inicialmente se revisan aspectos generales de la implementación, para posteriormente entrar de lleno en los experimentos.

4.1. Condiciones generales de la implementación

La implementación de los principales modelos, programa de control y rutinas de soporte del solver, se realizó en lenguaje de programación C con la utilización de la *Callable Library* de CPLEX 12.0. Todos los experimentos del primer grupo fueron ejecutados de forma secuencial, restringiendo a CPLEX a un único *thread* por ejecución, garantizando la independencia entre las corridas y la correcta medición de los tiempos de ejecución.

Adicionalmente para el análisis de resultados, generación y control de las instancias, especialmente en el experimento multitemporal, se utilizó el ambiente y lenguaje de programación [70] (versión 3.1.0), implementando una aplicación en R, que hace uso del solver en C, resolviendo las instancias secuencialmente por hilos paralelos independientes.

Las ejecuciones de estos experimentos se realizaron en el *cluster científico Vulcan* del Instituto de Sistemas Complejos de Ingeniería (ISCI) en el Departamento de Ingeniería Industrial de la Universidad de Chile. El cluster se encuentra compuesto por diez nodos de computo. Para el primer grupo de experimentos, se usaron ocho máquinas con dos quad-core Intel® Xeon® E5620 de 2.4GHz, 48 GB de RAM, arquitectura x86_64, sistema operativo Linux 2.6.32, con las siguientes tecnologías deshabilitadas por BIOS: Intel® Turbo Boost, Intel® Hyper-Threading e Intel® Virtualization (VT-x), para garantizar una adecuada medición de los tiempos de CPU, baja interferencia entre ejecuciones y el mejor desempeño para un thread-core. Para el segundo grupo de experimentos, se utilizó como coordinador de ejecución multi-hilo la aplicación desarrollada en R, la ejecución de este grupo de experimentos, se realizó en una máquina especial del cluster, que posee ocho procesadores octo-core AMD Opteron® Processor 6272, 132GB de RAM, arquitectura x86_64, sistema operativo Linux 2.6.32.

Uno de los elementos críticos en la implementación es el generador de números aleatorios, para esto se utilizó como base el propuesto en [52]. Es conocido este generador pseudo aleatorio por poseer propiedades deseables y necesarias, tales como: período superior a los tradicionales generadores, garantías de independencia y homogeneidad entre los números generados. Adicionalmente permitir de forma sencilla y directa modificar el estado del generador, eliminando la necesidad de guardar realizaciones particulares en memoria principal, o realizar costosas operaciones en CPU para llevar al generador a un estado particular, facultando el manejo de grandes muestras a bajo costo en CPU y RAM, elemento que fracasó en la primera implementación en Java con más de 30GB de RAM por repetición, lo que contrasta con los no más de 50MB por repetición con la implementación posterior en C. Para mayor desglose y detalle de las propiedades se le recomienda al lector remitirse a [55, p. 3-5].

Dentro de las extensiones realizadas en la implementación de [55], se incorpora la correlación entre los números generados y un generador particular, para normales mixtas correlacionadas. Para la correlación entre instrumentos, se utilizó la factorización de Cholesky de la matriz de correlación, partiendo de vectores de muestras independientes para generar números correlacionados, para mayor detalle se recomienda revisar [33].

El generador que considera la mezcla de normales multivariadas, se inspira en considerar el mundo como una cadena Markoviana, con estados normales multivariados, donde la realización futura, proviene de considerar las posibles transiciones con sus respectivos procesos subyacentes.

La implementación total, sin considerar el *Crawler*¹ de descarga de datos históricos, y considerando las rutinas de análisis e implementación parcial inicial en Java, consta de 5912 líneas en C, 8507 líneas en R, 253 en Shell script y 3104 en Java.

4.2. Visión general de los experimentos realizados

El conjunto experimental se divide en dos grupos principales, a su vez estos se subdividen dependiendo de la hipótesis o elemento a experimentar.

El primer grupo lo llamaremos de *experimentos monotemporales*. Estos experimentos buscan comprender el comportamiento de las diferentes medidas de riesgo, de forma individual y comparativa ante variaciones en sus parámetros, y distribuciones de las funciones generadoras de retornos. En estos experimentos, se analizan aspectos tales como: tiempo de resolución, tasa de convergencia, distribución de retornos del portafolio óptimo, perfiles de riesgo de la solución, estabilidad y composición del portafolio óptimo, ante efectos de correlación y colas gordas, para finalmente hacer una comparativa con la clásica teoría de portafolios de Markowitz.

El segundo grupo de experimentos, lo compone un experimento, explora la componente multitemporal del problema de portafolio, donde es necesario tener la solución del período anterior, para continuar con el siguiente. Lo anterior se trata a través de una metodología de *rolling forward* entre los diferentes períodos, incorporando dentro de las restricciones la composición del portafolio

¹Aplicación implementada para la descarga de datos históricos desde la Superintendencia de Pensiones (SAFP), esta información fue utilizada post procesamiento, para calibrar las distribuciones históricas con la banda de precios de la SAFS.

del período anterior, imitando el comportamiento de un agente racional administrador de fondos de pensión en la industria, que toma una administración activa de la cartera. A este experimento se le llamará *experimento multitemporal*, en él se incorporan restricciones adicionales, emulando imperfecciones del mercado, como es el caso de la iliquidez con restricciones sobre movimientos máximos.

A continuación se resume y detalla la estructura en que se encuentran divididos los experimentos, así como se explica de forma general su propósito:

1. Monotemporales:

- a) **Convergencia estocástica y tiempos de resolución:** El objetivo de este experimento, es determinar de forma robusta, el número de repeticiones y tamaños muestrales (in-sample y out-sample) a utilizar, considerando en este análisis los diferentes gap estocásticos, como medidas de convergencia, y los tiempos de resoluciones. Fijados los parámetros a ser utilizados, de forma transversal, en la resolución de los demás experimentos que serán abordados en este trabajo. Se analiza la sensibilidad de los parámetros de convergencia, todo lo anterior, para garantizar la significancia estadística de los resultados de los demás experimentos.
 - b) **Efecto de la correlación entre instrumentos:** Un efecto a despejar es el comportamiento de las diferentes medidas, ante instancias con correlación en los retornos generados. Se analiza esta situación, en un ambiente completamente acotado, de un proceso normal multivariado, con instrumentos de valor esperado cero, y varianza idéntica e igual a uno. La estructura considera un instrumento independiente, y dos grupos, correlacionados de forma contraria. En este experimento se observa: la cobertura generada por instrumentos contra cíclicos, y la coherencia de las medidas al existir instrumentos *semejantes*.
 - c) **Efecto de colas gordas:** En este experimento, se evalúa el efecto de distribuciones con colas *gordas*, en su sentido tomando como referencia la distribución normal. Estos procesos son generados a partir de la mezcla de normales. Se evalúa el caso con distribuciones simétricas y asimétricas. En estos experimentos, se deja un ejemplo, donde ante distribuciones asimétricas, al trabajar sólo con los primeros dos momentos, se llega a conclusiones diferentes.
 - d) **Contraste con el enfoque clásico de Markowitz:** En el contexto de distribuciones asimétricas, se compara el clásico resultado de Markowitz, con los portafolios obtenidos con las medidas de riesgo implementadas. En este experimento, se propone una instancia de retornos, donde los resultados de Markowitz y las medidas en el espacio retorno-varianza no coinciden, en particular, se muestran soluciones de las medidas CVaR, ECVaR y Entrópica, sub-óptimas en retorno esperado a Markowitz para iguales varianzas, y una solución de Markowitz sub-óptima en el sentido de CVaR y entrópico.
2. **Experimento multitemporal:** La multitemporalidad en el problema de portafolios, resulta un problema interesante y desafiante de resolver. Se debe tener presente, por un lado, que el problema global no puede ser resuelto como una serie separable de problemas de optimización, y por otro lado, que cada instancia a resolver es en si un problema con incertidumbre, y por ello estocástico.

La aleatoriedad en el problema de portafolios ya ha sido ampliamente cubierta en el presente trabajo, por lo que nos centraremos en explicar la dificultad de dependencia intertemporal. Es claro observar, que la resolución secuencial del problema, dividiendo en periodos el mismo, genera dependencia entre los mismos, por un lado, la solución del periodo anterior, entrega

la composición inicial del portafolio en el siguiente periodo, y a su vez esta, no puede ser alterada de forma radical con restricciones propias del mercado, que llamaremos *restricciones de liquidez*. Estas restricciones actúan, y son interpretables, como restricciones de no anticipatividad, generando una estructura de dependencia, que queda al descubierto al aplicar las clásicas técnicas de descomposición, sobre la dimensión temporal.

En este trabajo, para abordar la multitemporalidad, se resuelve el problema considerando una estrategia *rolling forward*, es decir, se resuelve iterativamente cada periodo, partiendo desde el comienzo, utilizando la solución del periodo anterior en la resolución del siguiente. Cada trayectoria generada, actúa como un escenario del problema estocástico a resolver.

En este experimento, a diferencia de los anteriores, se considera la utilización de datos históricos del sistema de pensiones Chileno, el periodo de resolución considerado es de diez años, entre el 2003 y el 2013. En el experimento, la estimación de los procesos generadores, se condiciona a una ventana móvil de observaciones, dependiente del día de resolución. En cada periodo resuelto, se utiliza la observación realizada para medir el desempeño de la composición escogida. El marco general anterior, permite evaluar la utilización directa de los modelos implementados, en particular, permite comparar su comportamiento con los actuales participantes del mercado, con la atenuante de estar comparando la única historia conocida. Se consideran las restricciones del régimen de inversión.

Como se mencionó anteriormente, adicional a las restricciones del régimen de inversión, se incorporan restricciones que emulan el efecto de iliquidez de mercado, restringiendo la variación máxima en composición en el rebalanceamiento periodo a periodo. Con las restricciones de liquidez introducidas, se espera apreciar la diferencia en la resolución *rolling forward* como aproximación, versus un problema multiperiodo global, donde es evidente la necesidad de una estrategia de programación dinámica estocástica. Desde un punto de vista económico, las restricciones de iliquidez, modelan el fenómeno que las AFPs en Chile, no son tomadoras de precios.

La comparación con la observación histórica, como método de *backtesting*, es poco robusto, dado que utiliza la única historia observada para comparar, no obstante, con el propósito experimental, de medir la fricción generada por las restricciones de liquidez, como restricciones de no anticipatividad, cumple su objetivo.

En experimentos anteriores, quedó de manifiesto la relevancia de considerar procesos no simétricos, donde las medidas de riesgo, versus el enfoque clásico, entregan soluciones diferentes. En este contexto, el experimento considera una mezcla de dos normales multivariadas, una de ellas en estado crisis, y la otra en estado no crisis. Ambas normales son estimadas con datos históricos, realizando la separación de las bases de datos de entrenamiento, según la clasificación definida por el *Business Cycle Dating Committee* de la *National Bureau of Economic Research*². En este experimento se evalúa la pertinencia práctica de los resultados teóricos obtenidos anteriormente.

En cada experimento se describe su diseño experimental, distribución generadora de retornos, parámetros y demás. Adicionalmente son abordados en cada experimento, los resultados obtenidos, su análisis y conclusión referente a la pregunta que origina el experimento.

Para el caso de las medidas de riesgo con parámetro de aversión, se ocupará el concepto de medida-aversión para referirse a la dupla medida de riesgo y aversión específicas.

²Fuente: <http://www.nber.org/cycles.html>

4.3. Experimentos monotemporales

En los experimentos de correlación y colas gordas, se habla transversalmente de 35 combinaciones medida-aversión, estas son: valor esperado, $CVaR_\alpha$ con $\alpha \in \{0.01, 0.05, 0.1, 0.5, 0.8, 0.95\}$, $ECVaR_{(\alpha, \epsilon)}$ con los valores de α antes explorados y $\epsilon \in \{0.1, 0.5, 0.9\}$ y finalmente la medida entrópica con $\theta \in \{0.01, 0.05, 0.1, 0.5, 1, 5, 10, 20, 30, 40\}$.

4.3.1. Análisis transversal de convergencia y tiempos de ejecución

Un aspecto transversal a todas las diferentes instancias, y crítico para que los resultados tengan validez estadística, es la convergencia estocástica de los mismos. Como se discutió anteriormente en 2.4, en el marco conceptual del trabajo, dada la metodología de resolución adoptada, la convergencia estocástica hace referencia al tamaño relativo del intervalo de confianza de SAA, ya sea este de la medida de riesgo de la variable $w \cdot x$ original o trasladada por temas numéricos, al encontrarse la misma muy cerca de cero.

Aún cuando estos temas fueron revisados con detalle en el marco conceptual, en 2.4, se insiste en los mismos, para asegurar la comprensión del significado de los resultados que se exponen. Se entenderá por gap estocástico, al tamaño del intervalo de confianza del valor de la función objetivo generada vía SAA. Es decir, dadas k repeticiones del problema de optimización equivalente cierto, con muestras independientes e idénticamente distribuidos (iid) de tamaño n , con sus respectivas soluciones $\{v_r^*, x_r^*\}_{r \in [k]}$, y una muestra adicional de tamaño m , iid a los anteriores, utilizados para computar la cota superior, llamaremos **gap estocástico absoluto** al tamaño del intervalo de confianza, es decir:

$$\left(\frac{1}{m} \sum_{j \in [m]} f(w_j \cdot x^*) + \sigma_{UB}(f, x^*, \{w_j\}_{j \in [m]})Z_\alpha \right) - \left(\frac{1}{k} \sum_{r \in [k]} v_r^* - \sigma_{LB}(\{v_r^*\}_{r \in [k]})Z_\alpha \right) \quad (4.1)$$

donde x^* corresponde a la mejor solución de las k repeticiones, según la muestra independiente de tamaño m , utilizado únicamente para computar la cota superior, $f(\cdot)$ corresponde a la función dentro de la función objetivo del problema, σ_{UB} el operador desviación estándar de la función objetivo evaluada en el independiente de tamaño m con la solución x^* , y σ_{LB} el operador desviación estándar de las soluciones obtenidas en cada repetición.

En la expresión 4.1, el primer término de la sustracción, corresponde a la cota superior, y el segundo término a la cota inferior.

Este valor de gap absoluto, si bien sirve para cuantificar la convergencia, en particular en el ajuste de los parámetros de SAA (n, k, m), no sirve para comparar con otras instancias de medida de riesgo y aversión, dado que un “buen o mal” valor dependerá de la magnitud de la función objetivo, por esta razón es natural entrar a relativizarlo de forma porcentual. En este caso, se escoge relativizar el gap absoluto al valor de la cota inferior, por ser problemas de minimización siendo más exigente, llamando a este último, **gap estocástico relativo**.

Lo anterior no resuelve por completo la generación de un indicador de convergencia transversal a todas las medidas y aversiones. El problema surge cuando se busca optimizar una instancia, en que el valor óptimo medida-aversión se encuentra muy cerca de cero o es en efecto cero, es

decir: $\rho(\theta, \vec{w} \cdot \vec{x}) \in V(0, \epsilon)$. Lo anterior genera problemas numéricos, que hacen perder el sentido al indicador, dado que la velocidad de convergencia del gap absoluto puede ser menor (o mayor) a la velocidad de convergencia de la cota inferior.

Como alternativa a lo anterior, explotando la propiedad de traslación de las medidas de riesgo, se introduce el concepto de **gap relativo desplazado (o normalizado)**, refiriéndose al tamaño relativo del intervalo de confianza, desplazando en L la variable que la medida de riesgo mide. El valor de desplazamiento L es obtenido imponiendo que la cota inferior sea cien, se hace esto, buscando generar una interpretación equivalente a la del porcentaje, es decir: L en $\rho(\vec{w} \cdot \vec{x}) = \rho(\vec{w} \cdot (\vec{x} + L)) + L$ tal que $(\frac{1}{k} \sum_{r \in [k]} v_r^* - \sigma_{LB} Z_\alpha) + L = 100$. Una alternativa a lo anterior, habría sido imponer que la cota superior trasladada sea cien, no obstante, dado que el gap relativo se calcula con la cota inferior, se escoge este último.

El nuevo indicador de convergencia debe ser utilizado con cuidado, dado que presenta la desventaja de perder sentido al ser utilizado para comparar entre diferentes medidas de riesgo. De este modo, si bien este indicador es utilizado para comparar y evaluar la convergencia entre variantes de una misma medida, para comparaciones entre diferentes medidas sólo es valido utilizar el gap relativo sin traslación y el gap absoluto, siempre teniendo el cuidado de comparar magnitudes comparables, en el sentido de su significado en cuanto a aversión. La imposibilidad de comparar con el gap relativo trasladado, surge de la diferencia de sentido en la unidad de riesgo, que cada medida cuantifica.

Para todos los análisis siguientes, a menos que se diga lo contrario, se considera un nivel de confianza del 99.86 %, lo que considerando una distribución normal, nos entrega un Z_α entorno a 3, el que será utilizado vía SAA (por muestras iid) para obtener garantías asintóticas.

A las combinaciones de instancia (generadores de retornos de los instrumentos), medida y aversión, las llamaremos **configuración**. Nos referiremos al conjunto de configuraciones, como todas las combinaciones de ellos, en particular se consideran todas las configuraciones utilizadas en todos los experimentos monotemporales.

Como resultado en el ajuste de SAA, se determinó que 100 repeticiones, cada una con tamaño *in-sample* de 10.000 y tamaño *out-sample* de 400.000 para la cota superior, es suficiente transversal al conjunto de configuraciones. Los detalles de este resultado serán revisados a continuación.

En la figura 4.1, se observa para los parámetros anteriormente escogidos de SAA, el gap relativo sin y con traslación de las diferentes configuraciones (eje X), ordenando las configuraciones por gap relativo no trasladado. En el gráfico del lado izquierdo se aprecia el gap relativo, conservando el orden en el eje x de las configuraciones del gráfico de la izquierda. Se observa en el gráfico de la derecha, el gap relativo normalizado para las diferentes instancias. Acorde a lo esperado, y agregando la información de la magnitud de la constante de traslación de cada configuración en la figura 4.2, se observa como las configuraciones, con valor de la medida de riesgo cercana a cero (constante de traslación más cercano a 100) tienden a tener un gap relativo sin normalizar excesivo que carece de sentido.

Se concluye en cuanto a convergencia, del gráfico derecho de la figura 4.1, que con los parámetros de SAA escogidos, se logra un gap relativo normalizado entorno al 0.003, lo que a un nivel de confianza del 99.86 %, permite considerar que las configuraciones resueltas proporcionan soluciones estadísticamente significativas.

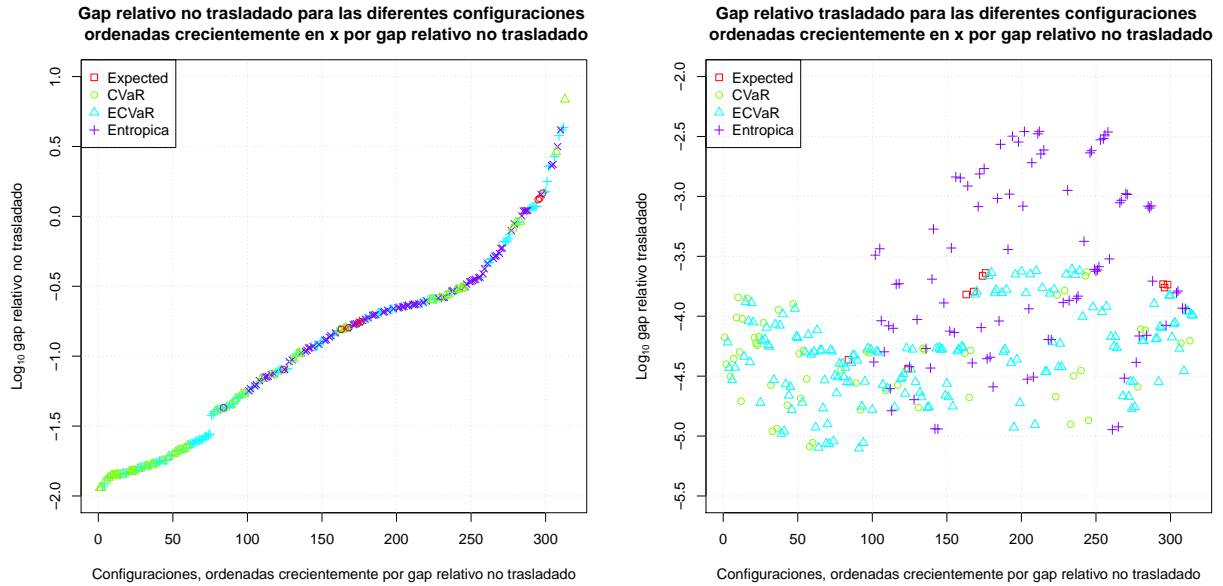


Figura 4.1: Gap relativo sin (izq.) y con (der.) normalización para las diferentes configuraciones, con los parámetros escogidos para SAA

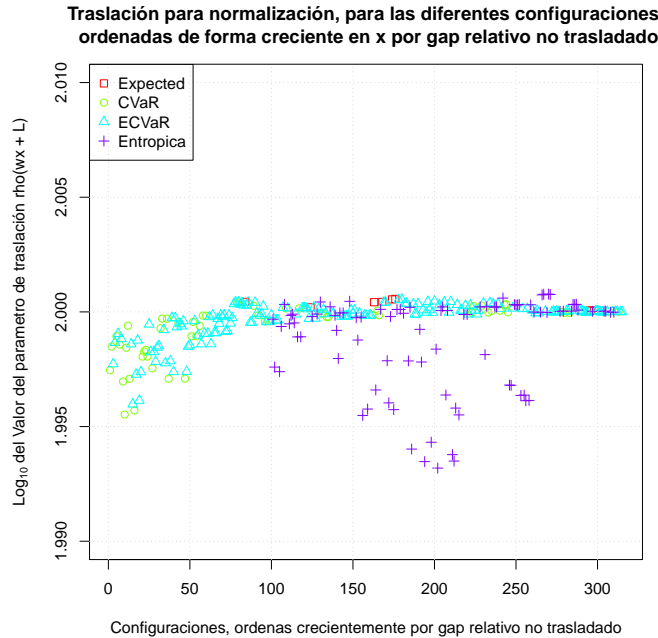


Figura 4.2: Valor de la constante de traslación L en $\rho(\vec{w} \cdot (\vec{x} + L))$, para las diferentes configuraciones, ordenadas crecientemente en el eje x por gap relativo no trasladado.

En cuanto a los tiempos promedios de resolución, en la figura 4.3 se observa predominantemente los tres grupos de problemas, las configuraciones involucradas con el valor esperado (en rojo), aquellas con medidas de distorsión (CVaR (verde) y ECVaR (celeste)) y aquellas configuraciones con la medida de riesgo entrópica (morado). Ambos gráficos de la figura 4.3, expresan la diferencia en ordenes de magnitud en el tiempo de resolución condicional a gap relativos normalizados equivalentes. Condicional a implementaciones “competitivas”, esto puede ser entendido como una mayor dificultad en el cómputo para alcanzar resultados estadísticamente significativos. Del gráfico

derecho, se concluye que los experimentos con menor tiempo de resolución es el valor esperado, consiguiendo una heterogeneidad de resultados en lo referente a convergencia, seguido por las medidas espectrales (CVaR y ECVaR) y finalmente las configuraciones con mayores tiempos de resolución, y peor gap relativo trasladado, la medida de riesgo entrópica.

Si bien el gráfico de la izquierda, con el simple ordenamiento creciente en x por tiempo promedio de resolución, deja de manifiesto estos tres grupos, el gráfico de la derecha nos entrega un orden de magnitud de las diferencias en los tiempos y su distribución.

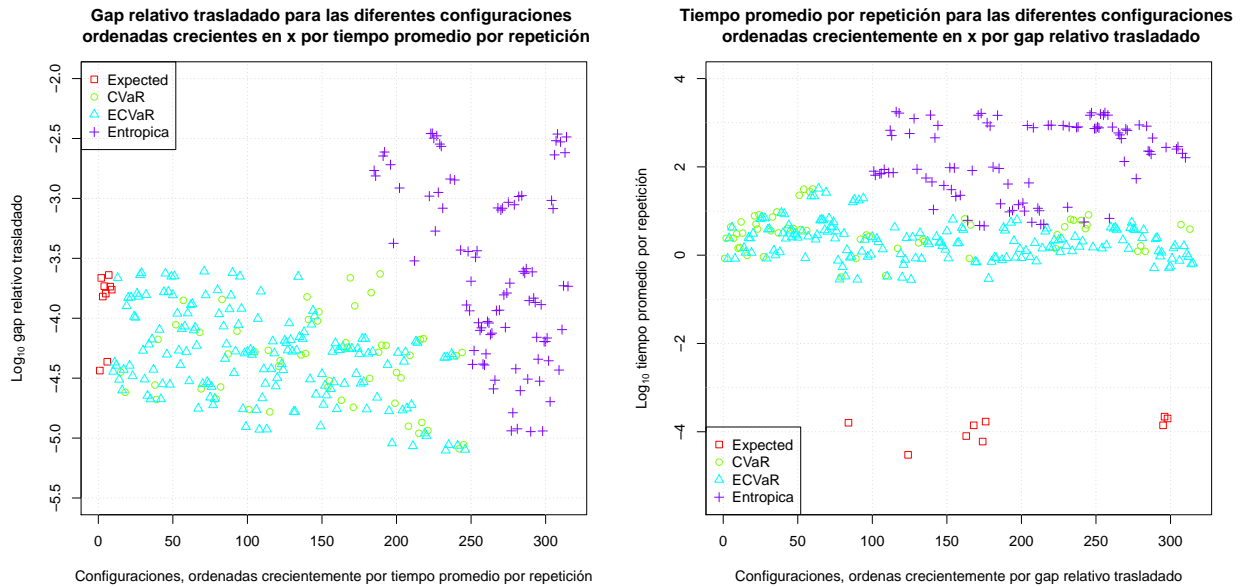


Figura 4.3: Gap relativo trasladado v.s. ordenamiento creciente por tiempo de las configuraciones (izq.), Tiempo promedio por repeticiones v.s. ordenamiento creciente de configuraciones por gap relativo trasladado (der.)

El resultado obtenido en los tiempos de resolución por medida, donde se observa un ordenamiento por medida, es esperable. La interpretación de este resultado radica en el número de procesos necesarios en la resolución de cada tipo de medida. La primera medida, requiere de un número reducido de operaciones, y la resolución de un problema de optimización sencillo, las segundas dos implementaciones, requieren la carga y resolución de problemas de optimización más grandes, no obstante, la resolución de la medida entrópica requiere de la resolución iterativa y ajuste de la grilla de aproximación, empleando un número mayor de ciclos de cómputo.

Un elemento que se consideró relevante incorporar, fue la distribución de valores de la constante de traslación en los óptimos, para las diferentes configuraciones, podemos observar esto en la figura 4.2. Este valor, transversal a las diferentes configuraciones, nos permite comparar los valores óptimos de las diferentes configuraciones, de una forma transversal. Predominantemente se observan valores entorno a 100, lo que es coherente a múltiples instancias con valores de la medida-aversión cercanos a cero. Se aprecia una predominancia, por parte de las observaciones que se alejan de 100, con valores inferiores a esta cota, especialmente de la medida de riesgo entrópica y de ciertas configuraciones de ECVaR, todas estas configuraciones están obteniendo soluciones muy superiores a cero, por lo que su constante de traslación es menor.

Como validez de la robustez de los parámetros de SAA escogidos, se realiza un análisis de sensibilidad de los mismos, separando el análisis en el efecto de cada uno de ellos, es decir: número

de repeticiones del problema equivalente cierto a resolver (k), tamaño *out-sample* para el cálculo de la cota superior (m), y finalmente el tamaño *in-sample* a utilizar en cada repetición del problema equivalente cierto (n).

Número de repeticiones (k)

Es conocido en la literatura que el número de repeticiones a resolver del problema equivalente cierto, con el supuesto de un tamaño *in-sample* representativo, por un lado reduce la varianza de la cota inferior, y por otro, a través de nuevos candidatos, reduce el expected gap (distancia entre la esperanza de la cota inferior y superior). Ambos efectos serán medidos y considerados en el análisis de sensibilidad.

Se considera que un rango razonable para el análisis de sensibilidad sobre este parámetro, se encuentra entre uno y cien repeticiones. Reforzando esta idea, las configuraciones con mayores tiempos de resolución, demoraron entorno a una hora por repetición, acotando la resolución de una configuración en el peor caso en aproximadamente cien horas, tiempo que se considera razonable.

Observando la evolución de los gap relativos, trasladado (der.) y no trasladado (izq.) en la figura 4.4, como se anticipó, surgen problemas numérico en el gap relativo no trasladado, lo que genera inestabilidades en el, problema que es corregido con la traslación en L de la medida, a través del gap relativo trasladado.

Como se aprecia del gráfico derecho de la figura 4.4, el gap relativo trasladado tiende a mostrar un comportamiento estable entorno a las 60 repeticiones, teniendo un efecto completamente marginal con los aumentos posteriores.

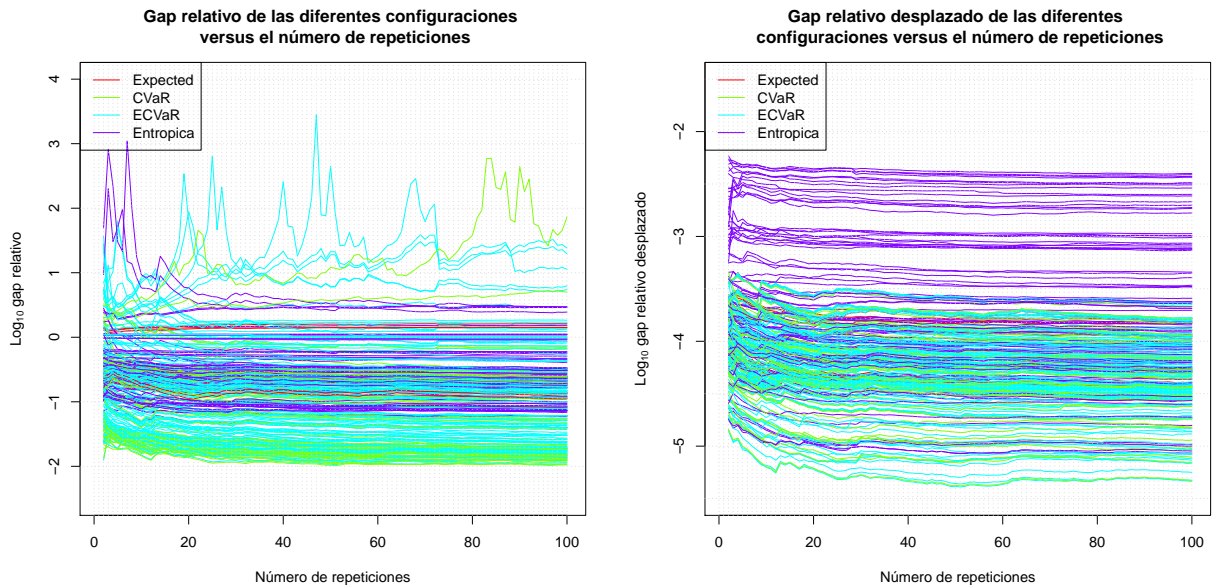


Figura 4.4: Evolución del gap relativo sin traslación (izq.) y con traslación (der.) versus el número de repeticiones, para las diferentes configuraciones.

Como fue expuesto anteriormente, el número de repeticiones afecta el conjunto de posibles candidatos óptimos, potencialmente alterando la cota superior. Este efecto lo podemos apreciar en la

figura 4.5, donde cada observación significa el cambio en el candidato óptimo para la cota superior, y el orden de magnitud, producto de una repetición adicional. Acorde a lo esperado, se observa un menor número de actualizaciones, y de menor magnitud, a medida que aumenta el número de repeticiones. Es importante destacar, que no se producen actualizaciones del valor esperado, siendo la solución obtenida en la primera repetición la utilizada. Contrario al resultado previo, observando sólo el número de repeticiones y los gap relativos, observamos que se produce un gran número de actualizaciones de la cota superior posterior a las 60 repeticiones.

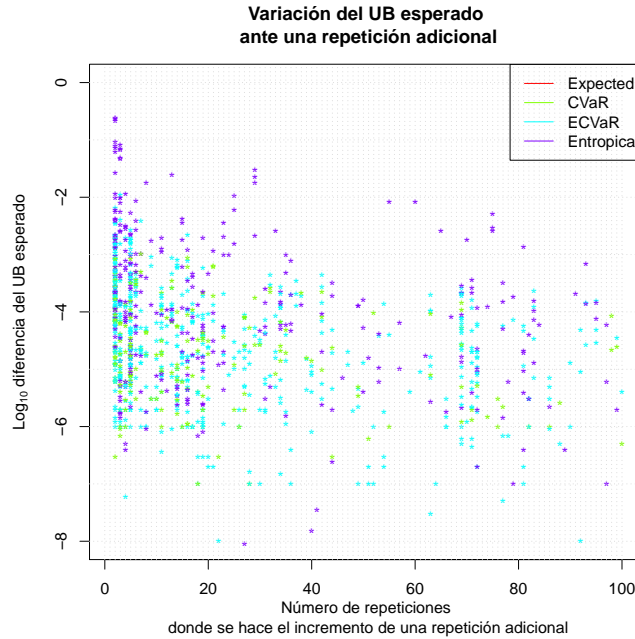


Figura 4.5: Cambio y magnitud de los mismos en la cota superior, al considerar una repetición adicional

De los resultados expuestos anteriormente, se concluye de forma transversal a las diferentes configuraciones que entorno a ochenta repeticiones es suficiente, no obstante, por los tiempos razonables de cómputo obtenidos, y a fin de estar por sobre lo esperado ante cualquier configuración, se decide mantener un número de cien repeticiones.

Tamaño out-sample (m)

Este sample es utilizado para evaluar las soluciones candidatas de las repeticiones, escogiendo el candidato con mejor valor para generar la cota superior. Implícitamente lo que se busca, es aproximar numéricamente la integral que esta involucrada en el problema, evaluando cada una de las soluciones candidatas.

De lo anterior, es claro que este parámetro afecta directamente la cota superior, ya sea a través de la elección de un candidato diferente, o a través de la convergencia por reducción de la varianza. Para medir estos efectos, coherente a lo ya desarrollado, y utilizando las medida de convergencia definidas, en la figura 4.6 podemos apreciar la evolución de los gap relativos, no normalizado (izq.) y normalizado (der.), versus el tamaño *out-sample*.

Podemos apreciar de la figura 4.6 que con tamaños sobre 100.000 se alcanza un grado de

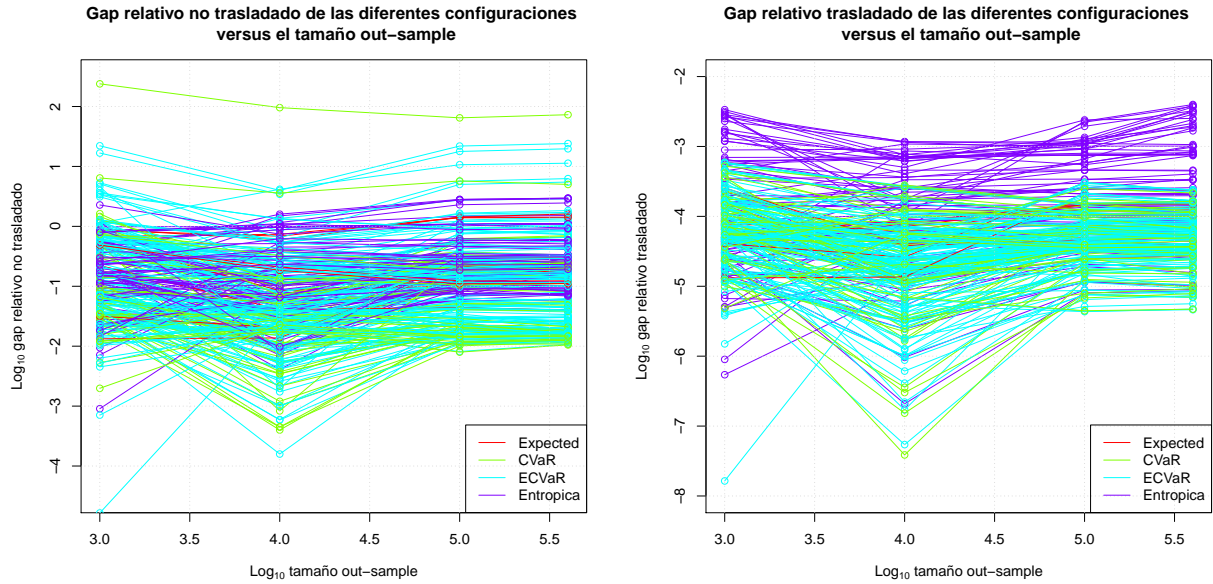


Figura 4.6: Evolución del gap relativo no normalizado (izq.) y normalizado (der.) versus el tamaño *out-sample*

estabilidad. Adicional a esta información, es relevante incorporar información relativa al tiempo de cálculo de la cota superior, en la figura 4.7 podemos apreciar los diferentes tiempos versus el el tamaño de la muestra, de este gráfico se observa que el peor tiempo de cálculo de la cota superior no supera los 10 segundos, no siendo una real restricción.

En la figura 4.7, apreciamos el crecimiento de los tiempos promedios de cálculo de la cota superior, versus diferentes tamaños. Se aprecian básicamente cuatro curvas. Las dos primeras, con un crecimiento sub lineal en escala log-log, agrupan las configuraciones con medidas de riesgo espectrales e instancias generadas: desde mezcla de normales correlacionadas (superiores celeste y verde) y normales independientes (celestes y verde inferiores); por otro lado las otras dos curvas, agrupan las configuraciones con medidas de riesgo entrópica y valor esperado, con crecimiento lineal en escala log-log, diferenciándose nuevamente por el tipo de instancia (mezcla de normales correlacionadas, superiores; normales independientes, inferiores).

De los resultados anterior, se concluye que un valor de 400.000 para este parámetro, es razonable y suficiente, teniendo como factores de elección el reducido tiempo de cómputo, y su contribución en la convergencia.

Tamaño in-sample (n)

El tamaño *in-sample*, como su nombre lo dice, se considera crítico para la obtención de soluciones representativas a la aleatoriedad del problema. Este parámetro de SAA, afecta directamente la estabilidad de las soluciones de las repeticiones y con esto, el efecto que pueda tener en la convergencia el número de repeticiones y el tamaño *out-sample*.

En la determinación de un valor razonable de este parámetro, intervienen dos efectos, por un lado el efecto en la convergencia estocástica del problema, y por otro lado, el tiempo promedio de

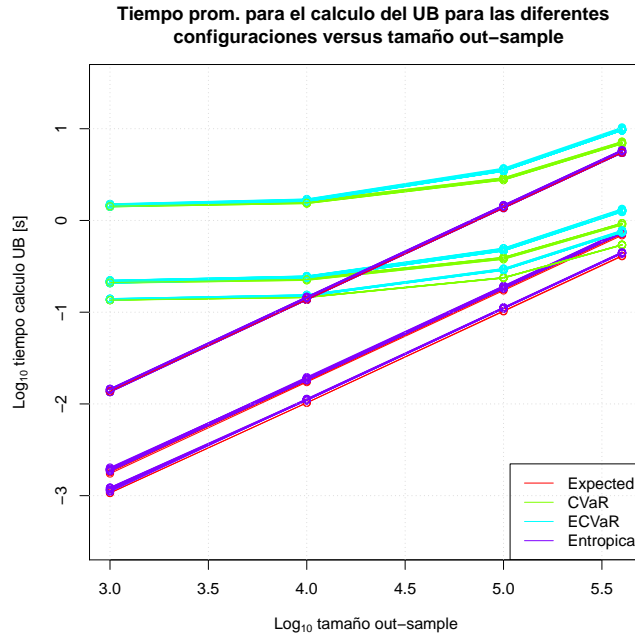


Figura 4.7: Evolución del tiempo promedio de cálculo de la cota superior, versus diferentes tamaños *out-samples*

resolución por repetición. Como es conocido, instancias con mayor número de realizaciones, tenderán a demorar más en su resolución, ya sea por la adición de restricciones o variables (lo que dependerá de cada medida de riesgo y su implementación).

Con el propósito de definir un valor razonable, a las diferentes configuraciones, tanto en tiempo promedio por repetición como en convergencia, se analizará transversal a las diferentes configuraciones, el efecto del tamaño de la muestra, a través de la evolución de los gap relativos, complementado lo anterior, con el tiempo promedio por repetición.

En la figura 4.8, podemos apreciar los gap relativos versus el tamaño *in-sample*, de donde observamos el efecto marginal, siendo razonable no más de 10.000 samples.

En la figura 4.9, podemos apreciar el crecimiento de los tiempos de cómputo, versus el tamaño de la muestra, observamos que este crecimiento es sub exponencial en general. Adicionalmente de este gráfico, observamos que la configuración con mayor tiempo de resolución, considerando el tiempo de cálculo de la cota superior despreciable, se demora entorno a una hora por repetición.

Considerando tiempos de cómputo razonables para una configuración, y el efecto marginal del aumento del sample en el gap relativo trasladado, como medida de la convergencia, se concluye que un tamaño *in-sample* de 10.000 es suficiente.

4.3.2. Efecto de la correlación entre instrumentos

Se busca identificar y cuantificar el comportamiento de las medidas de riesgo estudiadas, al enfrentar correlación entre los retornos de los instrumentos. Para simplificar la noción de correlación

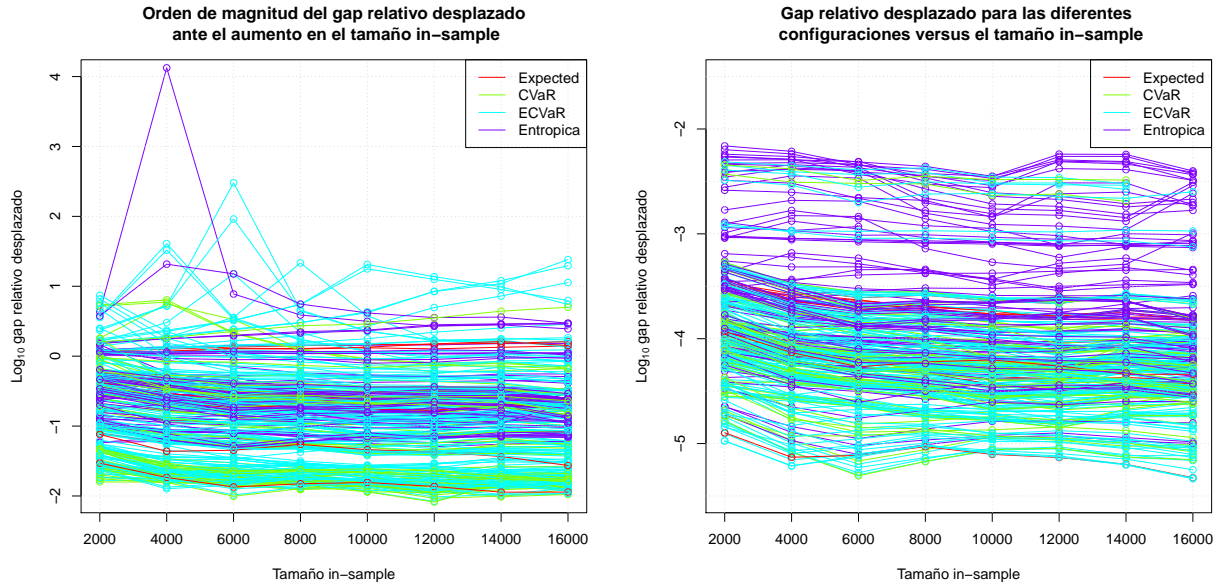


Figura 4.8: Evolución del gap relativo no trasladado (izq.) y trasladado (der.), versus el tamaño *in-sample* para las diferentes configuraciones

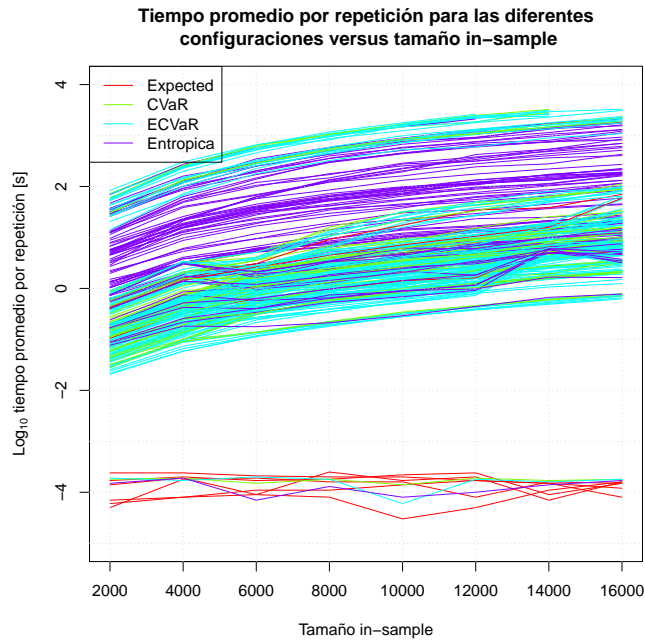


Figura 4.9: Tiempo promedio por repetición, versus el tamaño *in-sample*

estudiada, se trabaja en el contexto de procesos normales multivariados, donde la estructura de correlación queda completamente determinada por la matriz de correlaciones Σ , la que será tratada más adelante con cuidado.

Para controlar por factores adicionales no relevantes, se asume valor esperado cero e igual varianza para los diferentes instrumentos. Con lo anterior, el único elemento de experimentación son las correlaciones entre instrumentos, y la estructura que generan.

Es claro que para las diferentes medidas estudiadas, las soluciones que observaremos se distribuirán entre la entregada por el valor esperado, el que es “*ciego*” a la estructura de correlaciones, buscando sólo maximizar el retorno; y la entregada por la minimización del peor caso, donde se explotará al máximo la estructura de correlaciones para generar cobertura y diversificación sobre la distribución de retornos del portafolio, minimizando la noción de riesgo que se minimice. Los conceptos de cobertura y diversificación serán los fenómenos claves a ser identificados, y es necesario definir una noción clara de ellos.

La cobertura, se refiere al fenómeno conjunto de la distribución, donde en una misma realización, los peores casos de un instrumento siempre se compensan con buenos valores de otro. Este fenómeno se puede extender no sólo a un par, sino a un conjunto de activos, con los conceptos usualmente utilizados de *cíclico* y *contra cíclicos* o refugio, mitigando en el efecto agregando, casos extremos individuales. El *trade-off* natural que induce esto en el contexto de distribuciones normales multivariadas, y por ello simétricas, es la pérdida de muy buenos casos, al introducir cobertura a los peores casos, mediante instrumentos contraciclico que rentan poco, fenómeno que esperamos observar en los resultados de los experimentos.

El concepto de diversificación, tal como se dijo en 2.2, proviene de la coherencia de la medida, vía subaditividad o convexidad, esto nos dice que la diversificación puede reducir el riesgo, y por ello en ningún caso, aumentar el riesgo total. Este comportamiento de diversificación lo deberíamos observar ante la presencia de dos activos idénticos, donde la solución de cualquier medida de riesgo coherente, deberá preferir tener una fracción equitativa de ambos, al tener a alguno de ellos. Adicionalmente, este efecto debería ser observado, en la elección del activo control, el que es independiente a todos los demás activos.

Es importante distinguir y separar el efecto de la correlación, con el de colas gordas, que será tratado más adelante. Es conocida la expresión “*el efecto de las correlaciones entre instrumentos, generó colas gordas en los retornos de los portafolios*”, lo que constituye una expresión absolutamente falsa. La combinación convexa ($\mathbf{1} \cdot x = 1$, con $x \geq 0$) de variables aleatorias normales, genera una normal de media $x^t \mu$, y varianza $x^t \Omega x$ (con μ vector de valores esperados y Ω matriz de varianzas-covarianzas), lo que por definición no posee colas gordas. Con lo anterior es importante destacar el delicado rol que tiene la correlación en la diversificación del riesgo, basta analizar el caso de dos instrumentos idénticos, normales estándar ($\mathcal{N}(0, 1)$), analizando los casos extremos del portafolio con $\rho \in \{-1, 0, 1\}$ y composiciones $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0,5 \\ 0,5 \end{pmatrix} \right\}$, donde se obtienen como resultantes: una normal estándar, una normal de esperanza cero y varianza un medio y una constante igual a cero, eliminando la aleatoriedad; los resultados se pueden ver en la figura 4.10, donde vía simulación, se obtienen las diferentes distribuciones.

Diseño experimental

El experimento se encuentra conformado por dos configuraciones. Se resolverán las diferentes medidas-aversión de interés, en dos ambientes generadores de retornos, uno de ellos de control, con una instancia sin correlaciones $\mathcal{N}(\vec{0}, I\sigma^2)$, y el otro con una instancia experimental $\mathcal{N}(\vec{0}, \Sigma\sigma^2)$, con estructura de correlación diseñada para emular las dinámicas de diversificación y cobertura antes mencionadas.

Es usual identificar en el mercado grupos de instrumentos con comportamientos semejantes,

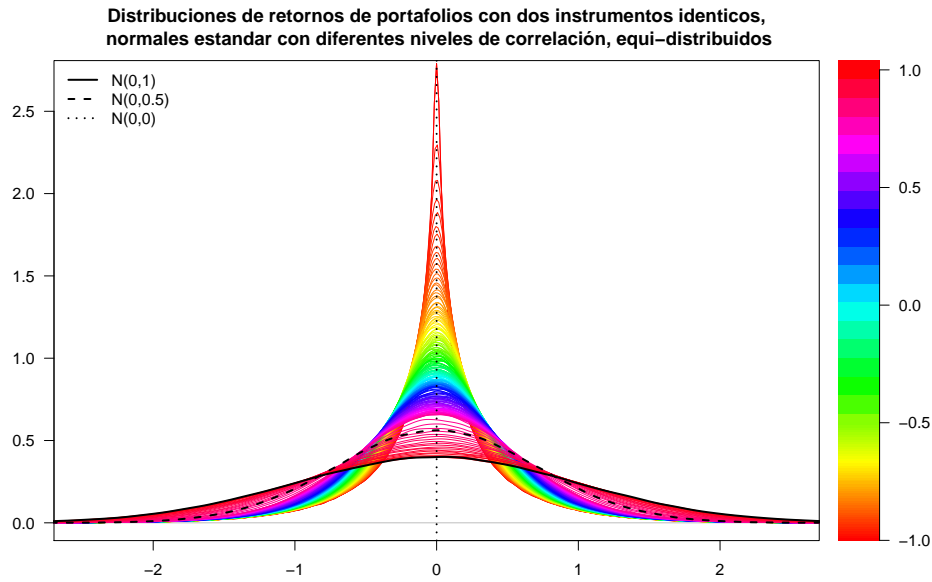


Figura 4.10: Simulación con dos variables normales estándar, modificando la correlación para portafolios equi-distribuidos

usualmente atribuido a variables comunes subyacentes no observables. A su vez, entre los grupos se generan dinámicas contracíclicas. Para facilitar el experimento, se definen dos grupos de instrumentos, cada uno de ellos, caracterizado con alta correlación positiva dentro del grupo, y de forma contraria, alta correlación negativa entre instrumentos de grupos distintos.

Lo anterior, permite un alto grado de generalidad, no obstante, para acotar los experimentos, se gradúa los niveles de correlación intra grupo en 0,5 y 0,75, e inter grupo en $-0,5$ y $-0,75$. Adicionalmente, se agrega un instrumento de control, independiente a todos los anteriores. Con lo anterior, se determina un mínimo de siete instrumentos.

La correlación extra grupo, queda definida por las nueve relaciones entre elementos de los dos grupos. Por consistencia, se consideran las correlaciones negativas, y al igual que en el caso intra grupo, se considera un elemento con mayor correlaciones negativa por relación. Se considerando un caso de las posibles combinaciones por permutaciones de lo anterior.

La matriz de correlaciones y varianzas-covarianzas (con $\sigma^2 = 1$), considerando siete instrumentos, con dos grupos ($g_1 = \{2, 3, 4\}$ y $g_2 = \{5, 6, 7\}$), y el primer instrumento independiente, es la siguiente:

$$\Sigma = \Omega = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0,5 & 0,75 & -0,75 & -0,5 & -0,5 \\ & & 1 & 0,5 & -0,5 & -0,5 & -0,75 \\ & & & 1 & -0,5 & -0,75 & -0,5 \\ & & & & 1 & 0,5 & 0,75 \\ & & & & & 1 & 0,5 \\ & & & & & & 1 \end{pmatrix}$$

Como se dijo anteriormente, es esperable que las soluciones, se encuentren entre la entregada al minimizar el peor caso (con máxima aversión), y la solución entregada al maximizar el valor esperado (neutro al riesgo).

Para la instancia de control $\mathcal{N}(\vec{0}, I)$, en el caso neutro al riesgo, es claro que el conjunto solución de esta instancia es todo el poliedro $\mathcal{P} = \{x \in [0, 1]^{|I|} \text{ tq : } \mathbf{1} \cdot x = 1\}$, y dada la metodología utilizada para su resolución, en cada instancia se tomará un instrumento como la solución, aquel que por temas numéricos, dado un *sample* de tamaño N obtenga $\max_{i \in |I|} -\hat{\mu}_i^N$, aún cuando en el límite todos tiene valor igual a cero, adicionalmente, dada esta solución, el portafolio óptimo tendrá distribución de retornos $\mathcal{N}(0, 1)$. En el caso de minimización del peor caso, la solución óptima será $x_i = \frac{1}{|I|} \forall i \in I$, lo que se obtiene de observar que dada una composición \vec{x} , el proceso conjunto $y_x = \vec{x} \cdot \vec{w}$ es una normal $\mathcal{N}(0, \|\vec{x}\|_2^2)$.

Las combinaciones medida-aversión consideradas son las 35 declaradas anteriormente, no obstante, sólo se reportan aquellas combinaciones que aportan información relevante.

Para analizar los resultados, se confrontan las distribuciones de retornos de los portafolios solución, para pares medida-aversión idénticos en ambos experimentos, utilizando para esto el test estadístico Kolmogorov-Smirnov entre ambas distribuciones. Es necesario comparar directamente la distribución del portafolio óptimo, dado que dos soluciones distintas en composición, en realidad pueden generar la misma distribución del portafolio, siendo en realidad otra solución del conjunto de soluciones, para ver lo anterior, basta considerar el caso de instrumentos normales independientes e idénticamente distribuidos, maximizando el valor esperado.

Adicionalmente, para contribuir con el análisis, se presenta la densidad de probabilidad, el perfil de distorsión y el perfil entrópico de los portafolios óptimos.

Resultados y análisis

Contrario a lo espero, en el experimento de control no sólo se obtienen dos portafolios óptimos, como se puede apreciar de la tabla resumen, y de la figura 4.11. Adicional a los dos portafolios límites: obtenidos de la minimización del peor caso y el caso neutro al riesgo, se obtienen seis combinaciones adicionales en el rango de medida-aversión explorado, estos portafolios son soluciones intermedias.

Para el caso de CVaR, se aprecia como con sólo dejar de ver el 5 % mejor de la distribución, se obtiene la solución de minimización del peor caso, situación semejante para el caso de la medida de riesgo entrópica con niveles de aversión sobre 0.05. Para el caso de ECVaR, al considerar combinaciones convexas de valor esperado y CVaR, donde el peso de CVaR sea mayor a 50 %, considerando α 's de 90 %, 50 % y 10 % se obtiene la solución de minimización del peor caso. Para otras combinaciones, se obtiene un portafolio que se encuentra entre ambas soluciones extremas.

Para tener una visión general de los resultados, a continuación se muestra un cuadro resumen, con estadística de las distribuciones de los portafolios solución de las principales configuraciones (medida-aversión), reportando: su valor esperado, desviación estándar muestral, estadísticos D y p -value del test Kolmogorov-Smirnov (en el cruce de pares medida-aversión similares, con y sin correlación).

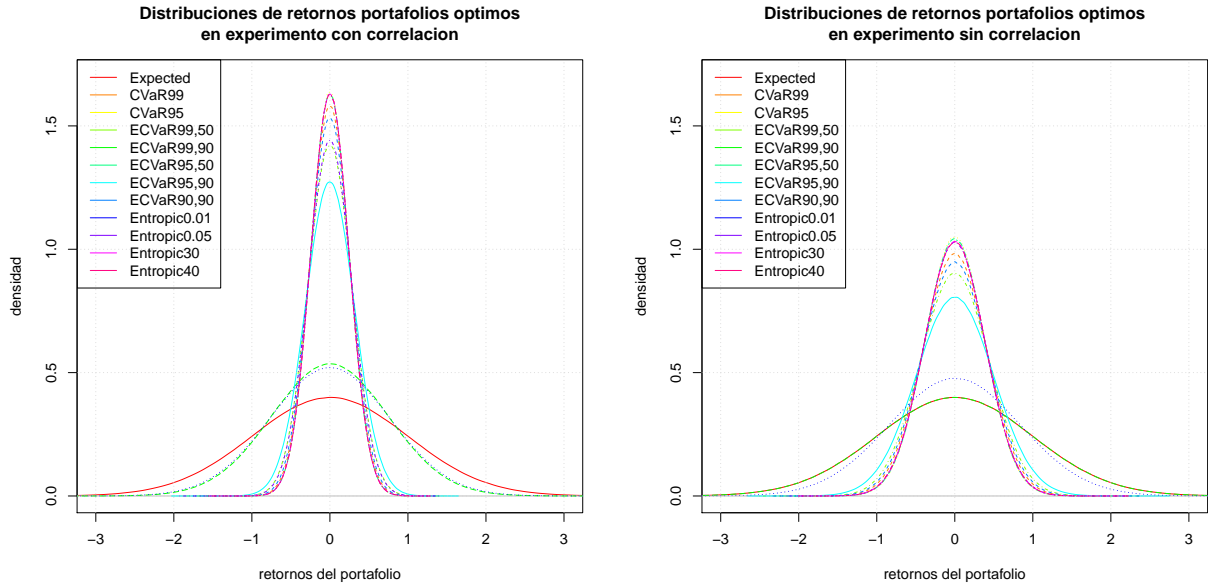


Figura 4.11: Densidades de probabilidad de los retornos, para los portafolios óptimos de las diferentes combinaciones medida-aversión, en la instancia con (izq.) y sin correlación (der.)

La estructura de correlaciones es explotada, generando cobertura entre instrumentos, obteniendo distribuciones óptimas con dispersiones inferiores a las obtenidas en la instancia de control. En la instancia de control, la ausencia de correlaciones se traduce en ausencia de coberturas, por lo que se aprecia reducción del riesgo sólo a través de la diversificación entre instrumentos.

Podemos observar que la única combinación medida-aversión que conserva la solución, en la instancia sin y con correlación, es el valor esperado. La solución de minimización del peor caso, en ambas soluciones difiere, donde a través del estadístico D , se observa que este desplazamiento es común. Valores donde el estadístico alcanza sus segundos mejores valores, hace referencia a combinaciones medida-aversión, que su significado es semejante al valores esperado.

Como se discutió anteriormente, coherente a lo obtenido en la tabla resumen, se aprecia en la figura 4.11, que los portafolios óptimos en la instancia con correlación explotan estas coberturas, reduciendo la dispersión de los posibles resultados. Es interesante observar, que dada la instancia simétrica escogida, todas las distribuciones óptimas que muestran una cobertura a potenciales pérdidas mayores, pierden exposición a casos favorables, controlando que para todos los posibles casos el valor esperado es nulo.

Como se observó de la figura 4.11 con las densidades, los perfiles espectrales, en la figura 4.13, muestran una reducción en el riesgo para iguales aversiones, entre la instancia con y sin correlaciones, resultado esperable al observar que en las densidades la cola izquierda de pérdidas se concentra más hacia el cero en presencia de correlaciones.

En el caso de los perfiles entrópicos, se observa coherente a los resultados anteriores, una reducción general en el riesgo para las diferentes configuraciones al considerar correlación. En el caso con correlación, en los perfiles entrópicos se observa un comportamiento irregular al pasar de niveles de aversión entre 10 y 15, pasando de crecimiento lineal a un crecimiento logarítmico en el riesgo al crecer en aversión, donde para el caso de las soluciones que incorporan diversificación (más

Medida-Aversión	μ_{corr}	μ_{ind}	σ_{corr}	σ_{ind}	$p-value$	D
Expected	0	0	1	1	0.12	0
CVaR99	0	0	0.25	0.41	0	0.11
CVaR95	0	0	0.24	0.38	0	0.11
ECVaR99,50	0	0	0.28	0.44	0	0.11
ECVaR99,90	0	0	0.75	1	0	0.07
ECVaR95,50	0	0	0.25	0.38	0	0.11
ECVaR95,90	0	0	0.31	0.5	0	0.11
ECVaR90,90	0	0	0.26	0.42	0	0.11
Entropic0.01	0	0	0.77	0.84	0	0.02
Entropic0.05	0	0	0.28	0.39	0	0.08
Entropic30	0	0	0.24	0.39	0	0.11
Entropic40	0	0	0.25	0.39	0	0.11

Cuadro 4.1: Tabla resumen de las distribuciones de los portafolios óptimos distintos, en los experimentos de correlación.

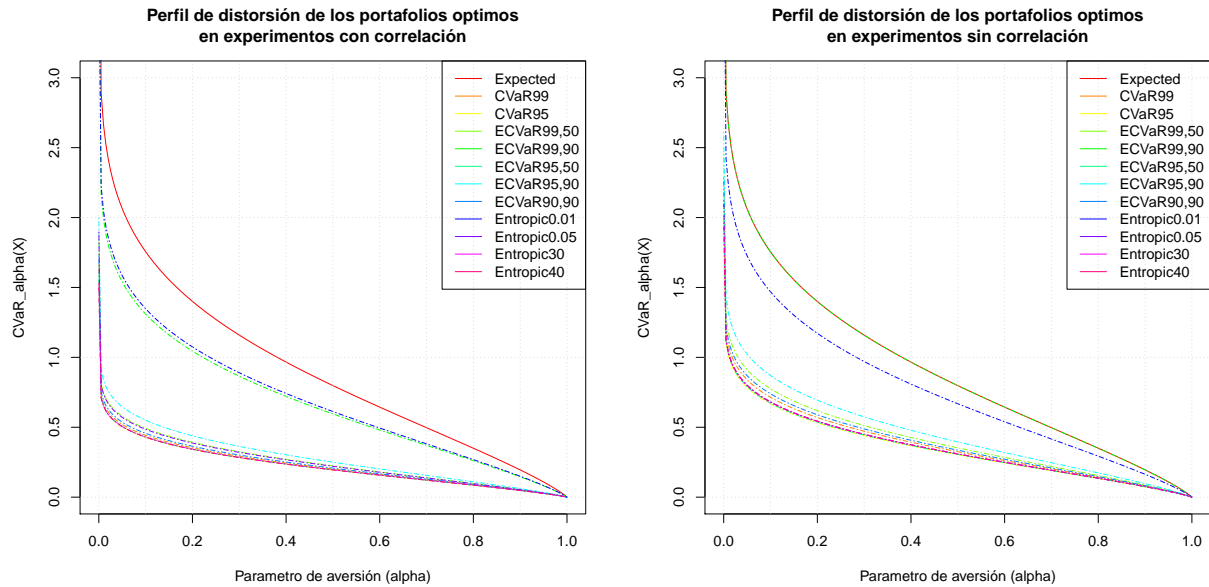


Figura 4.12: Perfiles espectrales de los diferentes portafolios óptimos en las instancias con y sin correlación.

cercanas a soluciones de la minimización del peor caso) se aprecia un quiebre en la tercera derivada del perfil de riesgo, con respecto a su nivel de aversión. Este comportamiento resulta especialmente interesante.

4.3.3. Efecto de colas gordas

Este grupo de experimentos, busca determinar y conocer el comportamiento de las medidas estudiadas ante la presencia de procesos de retornos con las llamadas *colas gordas*. Entenderemos de la noción clásica, que una distribución posee una *cola gorda*, cuando en alguna de sus colas de distribución, concentra mayor probabilidad que la distribución normal de referencia.

Perfiles entrópicos de los portafolios óptimos

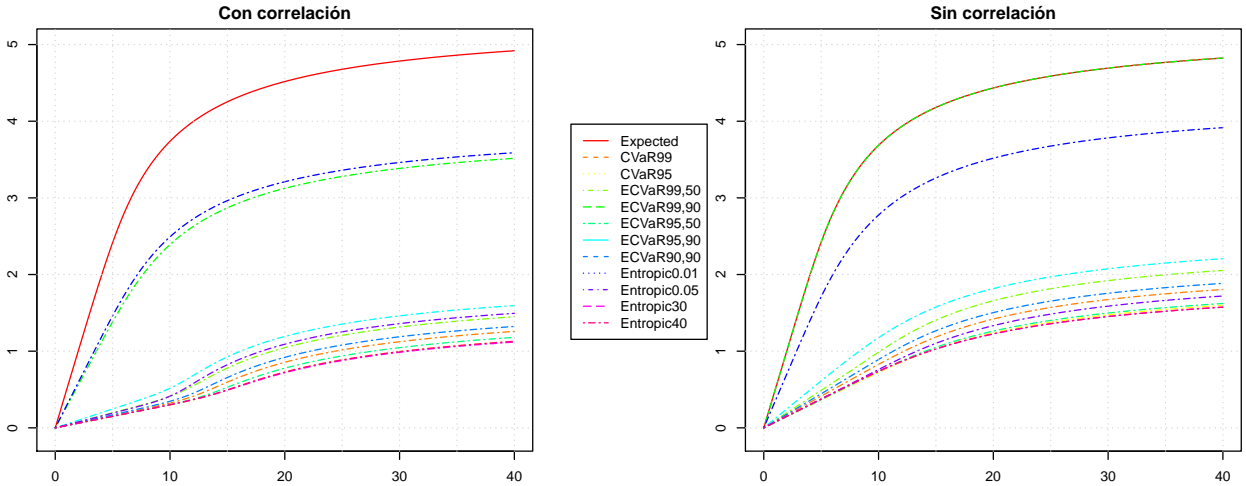


Figura 4.13: Perfiles entrópicos de los diferentes portafolios óptimos en las instancias con y sin correlación.

Las distribuciones de cola gorda recientemente han tomado un especial realce, este interés se ha generado por hallazgos empíricos, que han surgido dada la actual facilidad en la captura y recolección y almacenamiento de datos. Usualmente la identificación, caracterización y explicación de estos casos extremos, desempeña un rol crítico en la industria. Por nombrar algunos ejemplos: modelos de riesgo y estabilidad de sistemas, detección de operaciones anómalas, seguridad informática, diagnóstico asistido de pacientes y procesamiento de imágenes en sistemas semi asistidos, entre otros.

Un marco conceptual relativamente general, que a su vez, permite incorporar sentido en su construcción, hace referencia a las *mixture distributions*, en particular asumiendo distribuciones marginales normales a ser mezcladas, dando origen a las *mixture normal*. Estas distribuciones, en su versión de mezcla discreta, son obtenidas de suponer que un proceso escogerá aleatoriamente entre $|S|$ posibles estados, con probabilidades $p_s \forall s \in S$, donde cada estado se comportará como una distribución normal.

La distribución de probabilidad acumulada de un proceso normal de mezcla, suponiendo una mezcla finita de $|S|$ normales $\mathcal{N}(\vec{\mu}_s, \Omega_s)$, con distribuciones de probabilidad acumulada $P(x)$ será:

$$P(x) = \sum_{s \in S} p_s P_s(x) \quad (4.2)$$

Llevando este modelo a un contexto financiero, se supone una economía que puede encontrarse en diferentes estados, donde en cada estado ($s \in S$) se observa una distribución normal multivariada $N(\vec{\mu}_s, \Sigma_s)$ de los retornos de los instrumentos, y es conocida la matriz de probabilidades de transición a los posibles estados futuros j , dado que estoy en el estado i , denotando lo anterior como $\Pi_{i,j}$.

Es importante agregar, que este contexto permite trabajar con distribuciones simétricas y asimétricas, donde se espera tener resultados más interesantes en el caso asimétrico. Todos los conceptos anteriores, se consideran en su sentido clásico. En particular, se hace uso de los estadísticos: Kurtosis como medida de concentración central y del peso de las colas, y Skewness como medida de simetría de la distribución.

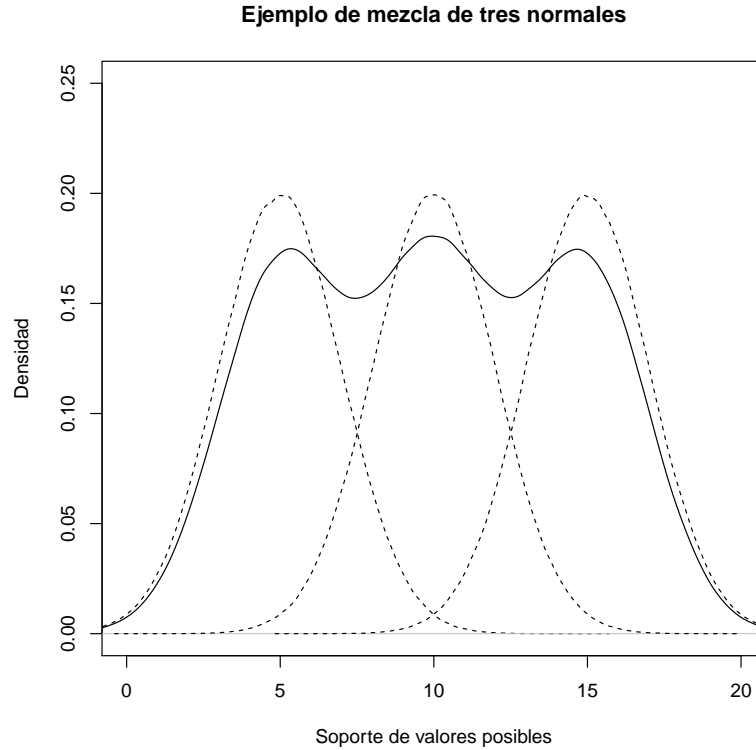


Figura 4.14: Ejemplo de la distribución mezcla de tres normales $\sigma = 2$ y $(\mu_1, \mu_2, \mu_3) = (5, 10, 15)$

Diseño experimental

Como primera aproximación se considera una instancia con dos estados, como se muestra en la figura 4.15, emulando una economía que tiene dos posibles estados, normal de expansión y de crisis. Con esto dependiendo del estado en que se encuentra la economía, se genera la distribución de posibles retornos futuros por instrumento, como una mezcla de ambas distribuciones, proporcional a las probabilidades de transición de estados.

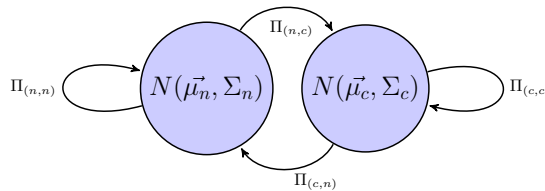


Figura 4.15: Ilustración de instancia con sólo dos estados: normal (n) y crisis (c)

Es importante observar que dado el marco general escogido, para generar distribuciones simétricas de cola gorda, bastará con un instrumento en que en ambos estados de la naturaleza tenga igual retorno esperado, pero niveles de volatilidad disímiles. Por otro lado, para generar instancias con distribuciones asimétricas, es suficiente con instrumentos con volatilidad no nula, con valores esperados de rentabilidad diferentes en ambos estados.

Se construye la siguiente configuración para la instancia con colas asimétricas, conservando un

instrumento de control, que se distribuye normal con igual valor esperado y varianza. Se consideran procesos sin correlación, dado que la misma fue tratada con anterioridad.

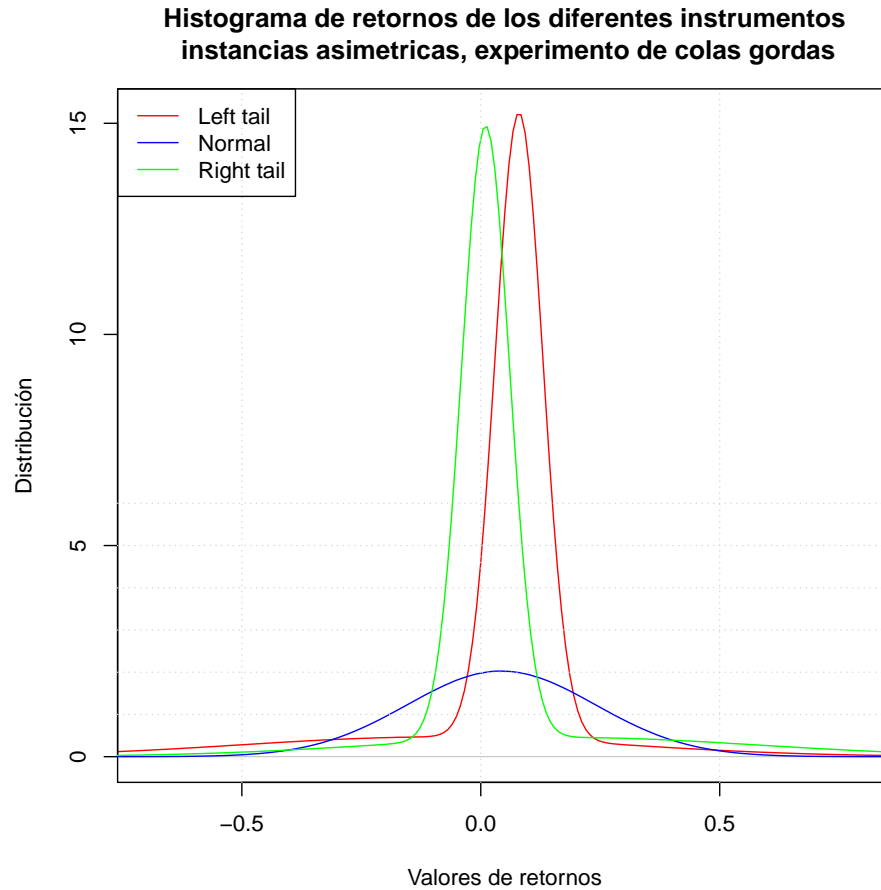


Figura 4.16: Densidades de los diferentes instrumentos utilizados en la instancia de colas gordas asimétricas

	μ	σ	Kurtosis	Skewness	$q5 - \mu$	$q95 - \mu$
Left tail	0.04	0.20	12.68	-1.83	-0.41	0.16
Normal	0.04	0.20	3.00	0.00	-0.32	0.32
Right tail	0.04	0.20	12.75	1.75	-0.16	0.41

Cuadro 4.2: Tabla resumen de los principales estadísticos de los instrumentos utilizados en cola gorda asimétrica

En la configuración anterior, adicional al instrumento de control con retornos normales, se introducen dos instrumentos de colas gordas asimétricas, concentrando en uno de los casos una cola derecha gorda, y en el otro caso una cola izquierda. Los tres instrumentos son generados y ajustados, con tal que posean los mismos primeros dos momentos (esperanza y varianza).

Es importante acotar que ambos instrumentos de colas gordas asimétricas, no poseen igual kurtosis y skewness dado que los instrumentos poseen valor esperado distinto de cero.

De forma equivalente, para el caso de colas gordas simétricas, se escoge una instancia con dos instrumentos, donde uno de ellos en ambos estados se comporta idéntico, generando una mezcla

normal, y el segundo de ellos posee diferentes varianzas entre los estados, con igual valor esperado, generando las colas gordas en la mezcla.

De forma idéntica a lo desarrollado para el caso asimétrico, el instrumento de control normal, tiene igual valor esperado y varianza que el instrumento de colas gordas simétricas.

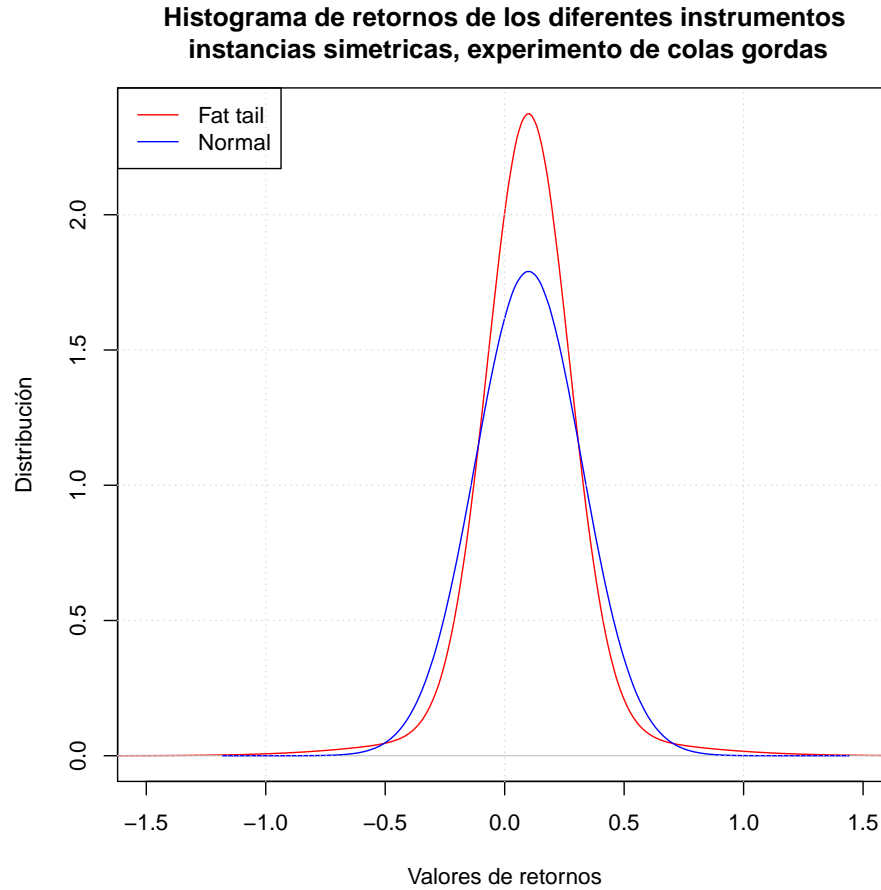


Figura 4.17: Densidades de los diferentes instrumentos utilizados en la instancia de colas gordas simétricas

	μ	σ	Kurtosis	Skewness	$q5 - \mu$	$q95 - \mu$
Normal	0.10	0.22	7.81	0.00	-0.33	0.33
Fat tail	0.10	0.22	3.00	0.00	-0.37	0.37

Cuadro 4.3: Tabla resumen de los principales estadísticos de los instrumentos utilizados en cola gorda simétrica

Resultados y análisis

Coherente a lo esperado, los diferentes portafolios óptimos poseen el mismo valor esperado, no obstante, sus densidades y perfiles de riesgo son distintos.

Tanto en la instancia asimétrica como simétrica, los portafolios óptimos nuevamente se distribuyen entre el obtenido con el valor esperado, neutro al riesgo, donde en todas las instancias

la solución óptima será escoger y concentrar toda la inversión en alguno de los instrumentos; y el portafolio obtenido de minimizar el peor caso posible, con máxima aversión.

Distribuciones asimétricas Es interesante destacar que ante la presencia de colas gordas asimétricas, los portafolios óptimos presentan distribuciones asimétricas, donde el comportamiento de una medida que actúa sobre alguna de las colas, y no sobre la distribución completa, no necesariamente tendrá el mismo efecto que otra más general, como es el caso de CVaR y la medida de riesgo entrópica.

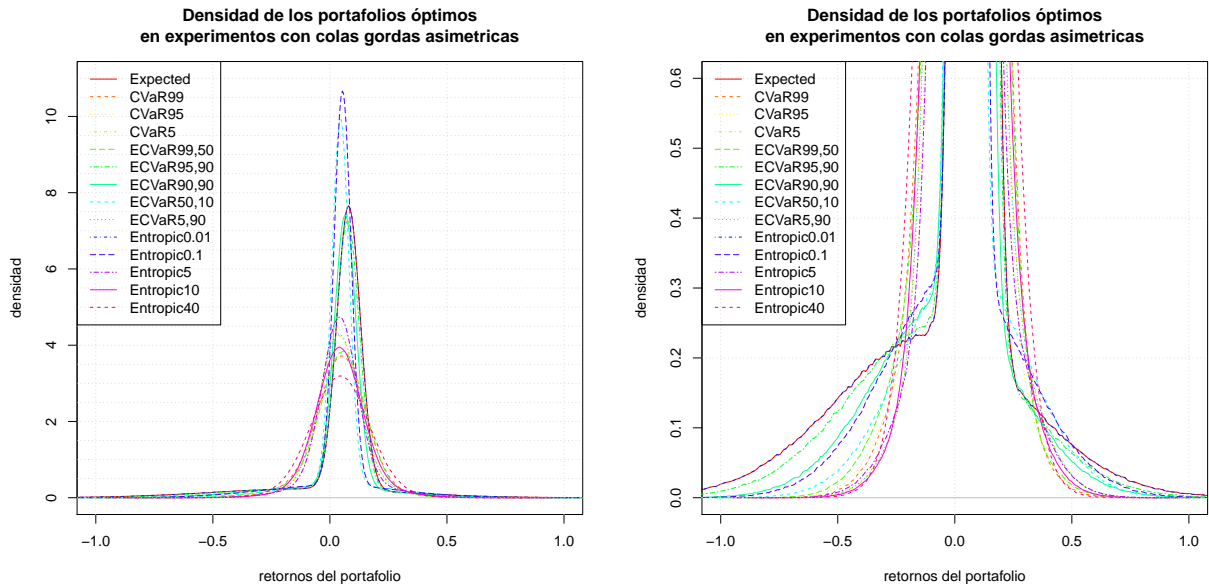


Figura 4.18: Densidades de los diferentes portafolios óptimos en la instancia con colas gordas asimétricas, general (izq.) y con acercamiento a las colas (der.)

A modo de resumen, se presenta la siguiente tabla con la estadística de los principales portafolios óptimos de las configuraciones medida-aversión analizados, considerando la instancia de colas gordas asimétricas.

Se puede apreciar que la solución neutra al riesgo, es alcanzada con: el valor esperado, ECVaR con 99% de valor esperado y un α de 0.9, y por la medida entrópica con θ de 0.01. En todos estos casos, la solución es escoger cualquiera de los instrumentos, en particular aquel que en determinada realización entrega un mayor valor, lo que resulta muy razonable y esperable.

La solución de minimizar el peor valor, se obtiene con CVaR α 0.95 y 0.05, resulta curioso que con α de 0.8 no suceda esto. Con la medida entrópica, se alcanza con nivel de aversión (θ) de 0.5, 1 y 5. Con niveles de aversión sobre 10, la dispersión del portafolio aumenta, no obstante su kurtosis se reduce.

Se presenta una diversidad de soluciones entre las dos expuestas anteriormente, algunas de ellas muy cercanas en varianza, no obstante, al observar sus distribuciones se aprecian sus diferencias, las mismas quedan en evidencia al analizar los quintiles, la kurtosis y el skewness.

Resulta interesante destacar, que al revisar grupos de soluciones *semejantes* en varianza,

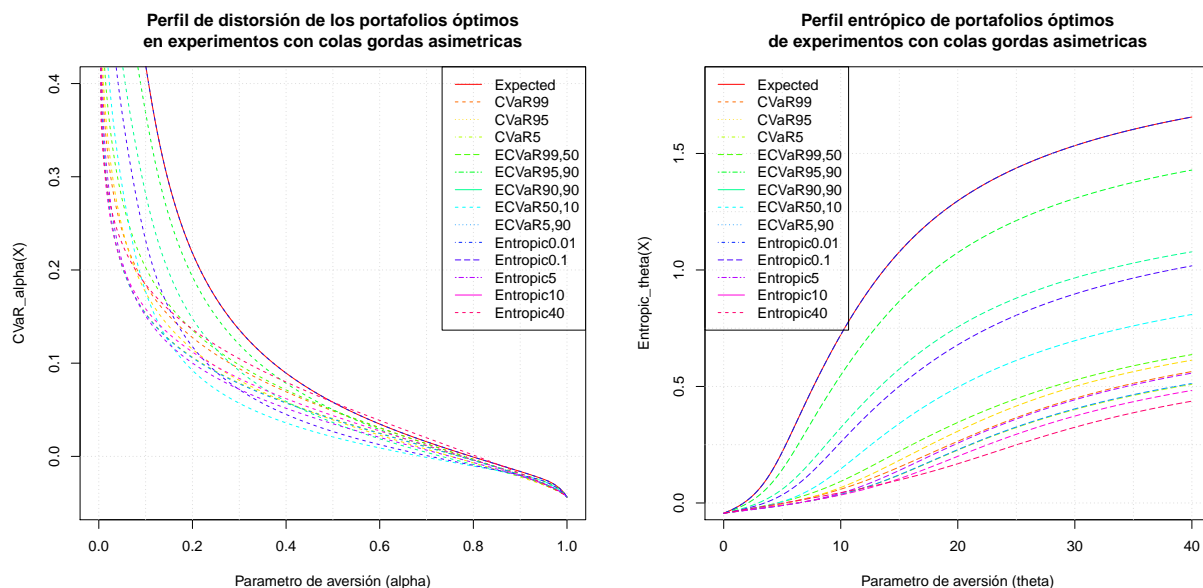


Figura 4.19: Perfil espectral (izq.) y entrópico (der.) de los portafolios óptimos en la instancia de colas gordas asimétricas.

	med.avr	mean	sigma	kurtosis	skewness	q5.Mean	q95.Mean
1	Expected	0.04	0.2	12.67	-1.83	-0.41	0.16
2	CVaR99	0.04	0.12	4.56	-0.3	-0.2	0.19
3	CVaR95	0.04	0.11	6.27	-0.42	-0.18	0.17
4	CVaR5	0.04	0.11	5.67	0.27	-0.17	0.18
5	ECVaR99,50	0.04	0.13	5.51	-0.55	-0.21	0.18
6	ECVaR95,90	0.04	0.18	12.38	-1.79	-0.37	0.15
7	ECVaR90,90	0.04	0.15	12.17	-1.5	-0.29	0.14
8	ECVaR50,10	0.04	0.12	12.86	-0.1	-0.18	0.17
9	ECVaR5,90	0.04	0.11	5.75	0.26	-0.17	0.18
10	Entropic0.01	0.04	0.2	12.67	-1.83	-0.41	0.16
11	Entropic0.1	0.04	0.14	13.21	-0.8	-0.24	0.16
12	Entropic5	0.04	0.11	6.71	0.16	-0.16	0.17
13	Entropic10	0.04	0.12	4.71	0.13	-0.18	0.18
14	Entropic40	0.04	0.13	3.3	-0.02	-0.21	0.21

Cuadro 4.4: Tabla resumen de los principales estadísticos de los portafolios óptimos en la instancia de colas gordas asimétricas

como es el caso de las soluciones de mínima varianza, se observan diferentes composiciones, las que en términos de varianza son indistinguibles. Es así como en el caso de CVaR (con α de 0.95) la solución óptima es invertir entorno al 40% en los activos normal y de cola izquierda gorda, con el 20% restante en el activo de cola derecha gorda. Por el contrario, con idéntica mínima varianza, se encuentra la solución de la medida de riesgo entrópica con θ de 5, con 27% en el activo de cola izquierda gorda, 36% en el activo normal y 37% en el activo con cola derecha gorda.

Para facilitar la comprensión visual del caso identificado, se muestra en 4.20, el boxplot de las diferentes soluciones por repetición de los portafolios. Es importante señalar, para evitar confusiones,

que esta gráfica de distribución y concentración de las soluciones, no significa necesariamente que la mejor solución este en ese sector, por esta misma razón es que anteriormente se reporto la composición óptima de los respectivos portafolios.

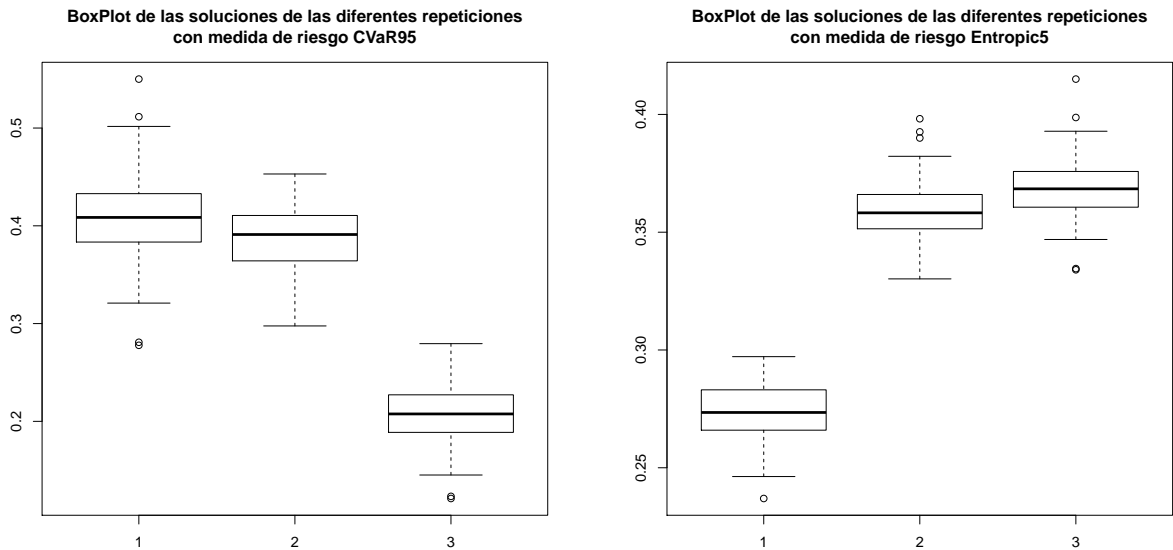


Figura 4.20: Boxplot de la composición en las diferentes repeticiones para CVaR95 y Entrópica(5).

Al analizar cuidadosamente el resultado anterior, del portafolio CVaR95 y la medida riesgo entrópica con θ de 5, observamos que aún cuando en términos de varianza ambos portafolios son idénticos, las densidades y demás medidas son diferentes, situación que puede ser apreciada en las figuras 4.21 y 4.22.

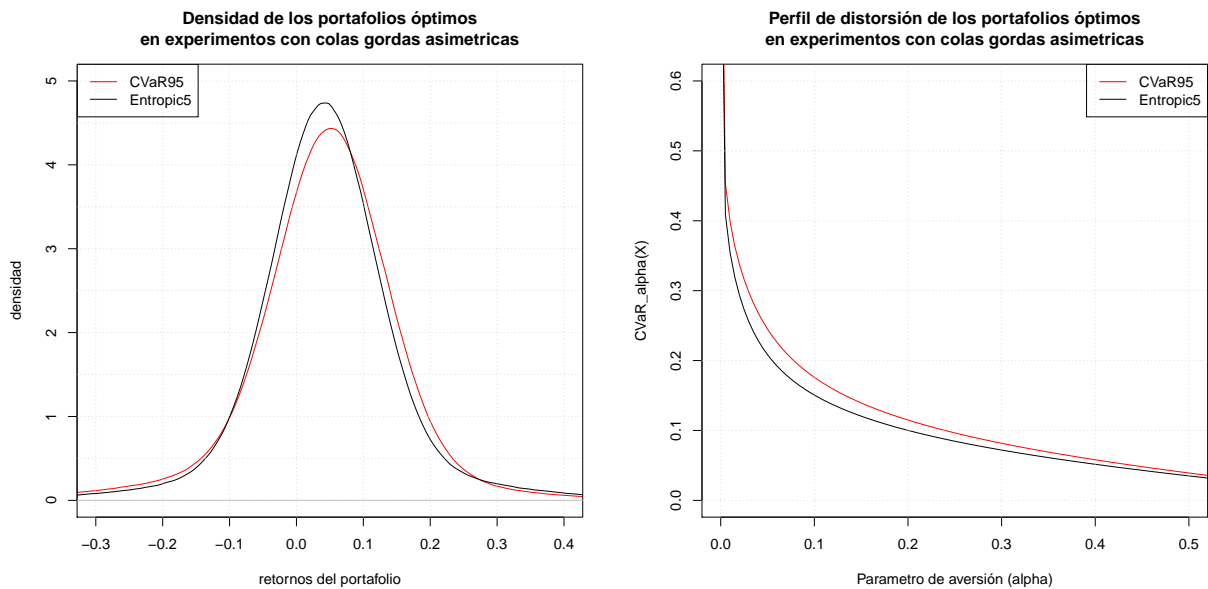


Figura 4.21: Densidades y perfiles espectrales de los portafolios óptimos con CVaR95 y Entrópica (5)

El resultado anterior es una solida evidencia, de las omisiones que puede generar trabajar sólo con enfoques basados en los primeros dos momentos de las distribuciones.

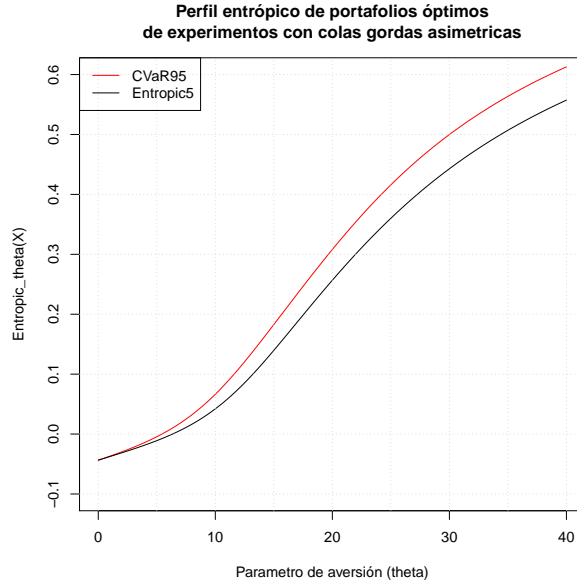


Figura 4.22: Perfiles entrópicos de los portafolios óptimos con CVaR95 y Entrópica (5)

Distribuciones simétricas Al analizar los resultados obtenidos con la instancia de colas gordas con distribuciones simétricas, observamos que la cantidad de portafolios óptimos, en la misma cantidad de combinaciones medida-aversión explorada es menor. Se observa nuevamente una distribución de portafolios entre los dos casos extremos de valor esperado, y minimización del peor caso. Adicional a los dos portafolios óptimos extremos, los únicos diferentes son los obtenidos por la medida de riesgo entrópica con θ de 10 y 20, estos dos portafolios se caracterizan por una mayor dispersión, pero niveles de kurtosis menores.

Simétrico a lo desarrollado para las colas gordas asimétricas, a continuación se presenta una tabla resumen con los principales estadísticos de los portafolios óptimos, densidades y perfiles.

	med.avr	mean	sigma	kurtosis	skewness	q5.mu	q95.mu
1	Expected	0.1	0.22	3	0	-0.37	0.37
2	CVaR99	0.1	0.16	3.24	0	-0.27	0.27
3	CVaR5	0.1	0.16	3.78	0	-0.26	0.26
4	ECVaR99,50	0.1	0.16	3.26	0	-0.27	0.27
5	ECVaR90,90	0.1	0.16	3.96	0	-0.25	0.25
6	ECVaR95,10	0.1	0.16	3.79	0	-0.26	0.26
7	ECVaR5,90	0.1	0.16	3.76	0	-0.26	0.26
8	Entropic0.01	0.1	0.22	3	0	-0.37	0.37
9	Entropic0.05	0.1	0.16	3.59	0	-0.26	0.26
10	Entropic5	0.1	0.16	3.74	0	-0.26	0.26
11	Entropic10	0.1	0.17	3.23	0	-0.27	0.27
12	Entropic20	0.1	0.18	3.04	0	-0.29	0.29

Cuadro 4.5: Tabla resumen de los principales estadísticos de los portafolios óptimos en la instancia de colas gordas simétricas

De la tabla 4.5, como se dijo anteriormente, en términos de varianza se identifican cuatro portafolios óptimos distintos, no obstante, al considerar la kurtosis, observamos que son diferentes.

Al evaluar los perfiles entrópicos de los diferentes portafolios, en la figura 4.24, se observa que efectivamente los portafolios seleccionados son diferentes en términos de perfiles de riesgo entrópico, aún cuando en términos de varianza puedan ser semejantes.

En línea con lo anterior, es importante destacar que el perfil espectral que se muestran en la figura 4.24, muestra resultados coherentes a los obtenidos al analizar puramente con un enfoque basado en la varianza, identificando sólo cuatro posibles portafolios, uno de ellos completamente diferentes, neutro al riesgo, con riesgo espectral mayor que todos los demás, y tres portafolios relativamente semejantes.

Podemos observar de la figura 4.23 las diferentes distribuciones de los portafolios óptimos en la instancia simétrica. Nuevamente esta simetría se traduce en que la cobertura sobre casos desfavorables, genere una perdida sobre casos favorables.

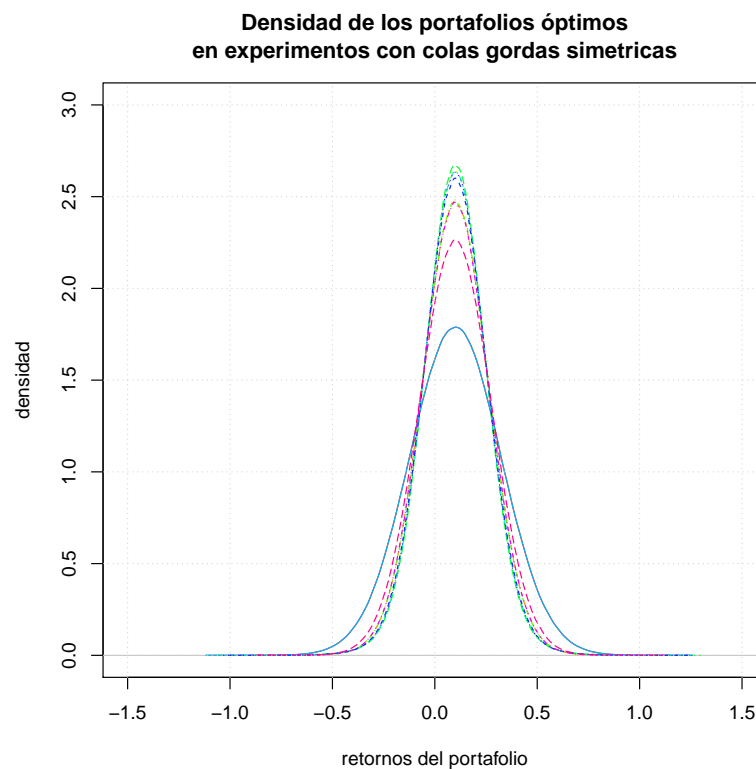


Figura 4.23: Densidades de los diferentes portafolios óptimos en la instancia con colas gordas simétricas

Entre los resultados interesantes y completamente inesperados, se aprecia el comportamiento lineal de los portafolios neutros al riesgo en el perfil entrópico (figura 4.24), que de forma tardía con una aversión recién entorno a 20, se produce el quiebre en su pendiente con incrementos decrecientes (comportamiento logarítmico).

4.3.4. Las Medidas de Riesgo y el enfoque clásico de Markowitz

Una pregunta que surge de forma natural es ¿cómo se compara lo desarrollado, con la clásica solución de minimización de varianza de Markowitz?. Con los experimentos anteriores, es claro que

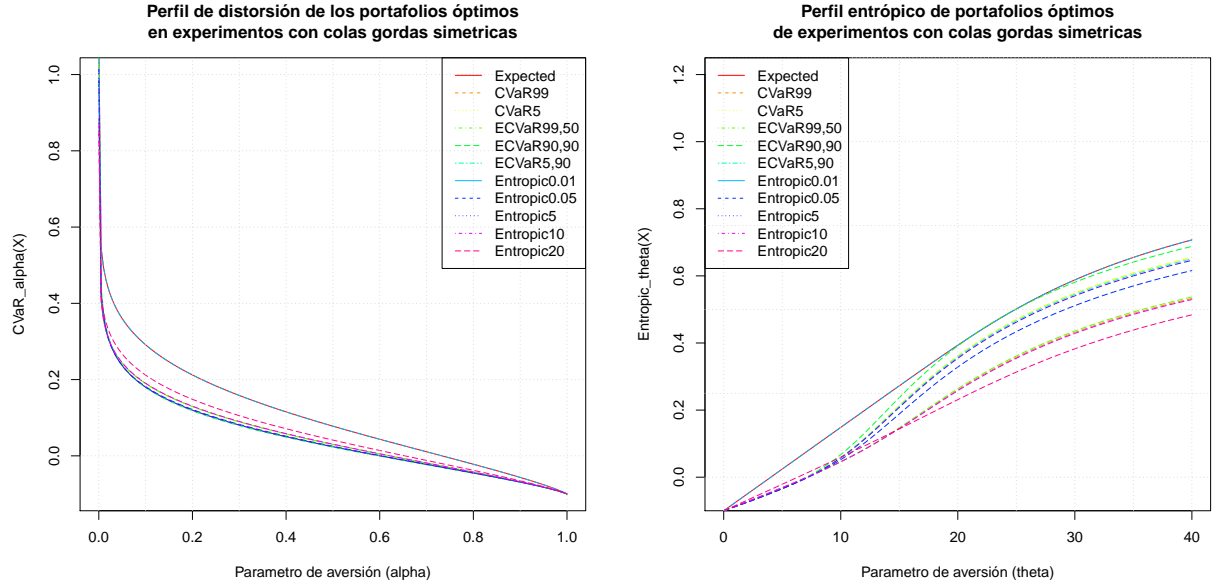


Figura 4.24: Perfil espectral (izq.) y entrópico (der.) de los portafolios óptimos en la instancia de colas gordas simétricas.

ante instancias simétricas, en particular normales, los diferentes enfoques obtienen soluciones idénticas, *mapeables* en términos de factores de aversión. Frente a instancias con colas gordas asimétricas, como quedo en evidencia, portafolios óptimos con igual retorno esperado y varianza, obtienen perfiles de riesgo entrópico y espectral diferentes, en particular, se observaron composiciones completamente diferentes. El presente experimento, evalúa el desempeño de los diferentes portafolios implementados, agregando en la comparación, el clásico enfoque de minimización de varianza con positividad en las soluciones, introducido en [58].

A fin de realizar una instancia sencillas, pero suficientes para exponer las diferencias de los portafolios, se considera una instancia artificial, mezcla de normales multivariadas, con resultante asimétrica. Como método de comprobación, replicando lo desarrollado en [58] esta vez con positividad, se obtienen las fronteras eficientes en el espacio varianza-retorno para los diferentes modelos. Posteriormente se evalúan portafolios específico, que dejan al descubierto y ejemplifican las diferencias en las medidas, en particular, se muestra la no linealidad en varianza y retorno, de las instancias ante las variaciones en los parámetros de aversión. Se muestran soluciones de Markowitz, que no pueden ser replicadas por las medidas con algún nivel de aversión.

Un elemento crítico al entrar a trabajar con lo desarrollado por Markowitz, constituye la obtención de la matriz de covarianzas Ω . Este elemento es obtenido con su estimador insesgado, la matriz de *covarianzas estimada* $\hat{\Omega}$ desde un sample representativo de la distribución conjunta, obtenida de la mezcla de normales. Es conocida la sensibilidad de esta estimación a observaciones atípicas, elemento que se tendrá presente en su estimación.

Tanto la varianza como medida de dispersión, y la covarianza como medida de correlación, asumen simetría en las distribuciones para que sus valores tengan sentido, este elemento será explotado en el diseño experimental. Como se dijo anteriormente, con la instancia asimétrica, con igual valor esperado, se espera encontrar una composición óptima en varianza por Markowitz, dominada desde el punto de vista entrópico y espectral; y por otro lado, soluciones óptimas de los otros portafolios, sub-óptimas en el sentido de Markowitz.

Diseño experimental

Se considera una instancia con dos instrumentos, facilitando su visualización, con asimetría en las distribuciones marginales, vía mezcla de dos estados normales. En el diseño del proceso generador de retornos de este experimento, se busca maximizar la disociación entre el proceso real asimétrico, y el simétrico supuesto estimado con $\hat{\Omega}$. [58].

Por simplicidad, en cada uno de los estados normales, se mantiene la independencia entre los instrumentos. En el proceso conjunto, se consideran instrumentos con igual valor esperado y varianza.

Con las condiciones antes enunciadas, se propone la siguiente configuración para el proceso generador que se muestra a continuación:

El sistema se compone de dos estados, de encontrarse en el estado de no crisis (nc), presenta retornos positivos con valor esperado uno. El primer instrumento presenta la menor dispersión, con varianza uno. No obstante, el segundo instrumento tiene una varianza de nueve.

$$r_{nc} \sim \mathcal{N}\left(\mu_{nc} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \Omega_{nc} = \begin{pmatrix} 1 & 0 \\ 0 & 3^2 \end{pmatrix}\right)$$

Por el contrario, de encontrarse el sistema en estado de crisis (c), los retornos esperados de ambos instrumentos son negativos, invirtiendo el orden de dispersión de los instrumentos:

$$r_c \sim \mathcal{N}\left(\mu_c = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \Omega_c = \begin{pmatrix} 3^2 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

Considerando que ambos estados suceden equiprobablemente, se genera la distribución de probabilidad conjunta, que se observa en la figura 4.25. En la figura 4.25, se aprecia la superposición de ambas distribuciones en valores extremos.

Estimando $\hat{\Omega}$ y $\hat{\mu}$ de los datos, y asumiendo normalidad, se obtiene el siguiente proceso:

$$r_{est} \sim \mathcal{N}\left(\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega = \begin{pmatrix} 6 & 1 \\ 1 & 6 \end{pmatrix}\right)$$

Tener instrumentos de igual valor esperado, y la restricción de positividad, teóricamente implica en el caso de Markowitz, que la frontera eficiente retorno-varianza se colapsa a un punto, que a su vez, es el punto de mínima varianza. En la práctica, dada la utilización de realizaciones de las variables aleatorias, se espera obtener fronteras eficientes en el clásico sentido de Markowitz, con un rango entre el portafolio de máximo y mínimo retorno pequeño, y dependiente del número de realizaciones consideradas en la estimación de $(\hat{\mu}, \hat{\Omega})$.

Distribución real desde el proceso de mezcla de normales

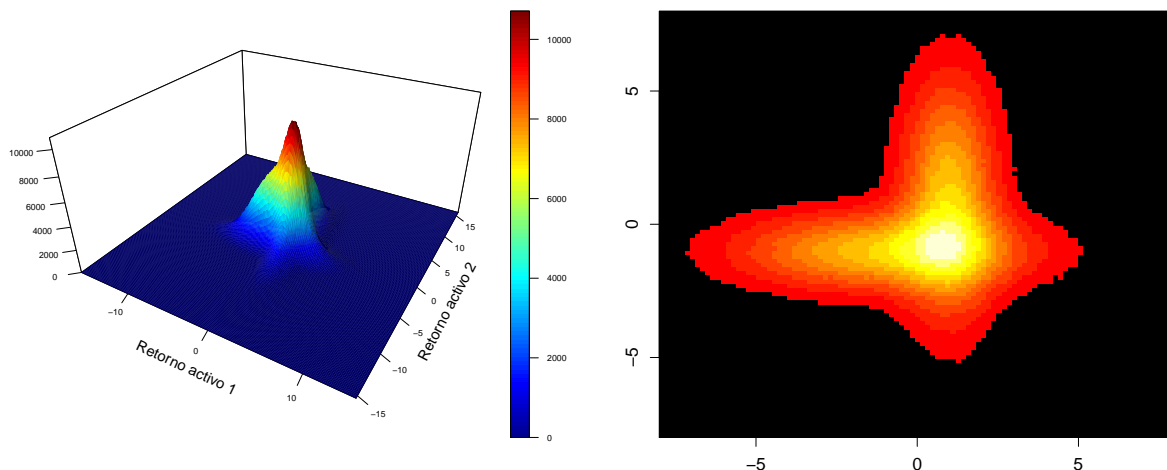


Figura 4.25: Histogramas de la distribución de mezcla generada con dos instrumentos.

La distribución de probabilidad que observamos en la figura 4.26, corresponde a asumir que el proceso conjunto anterior (r_{est}), corresponde a un proceso simétrico normal. En la figura 4.27, se contrapone con el histograma de la distribución real desde la mezcla de estados normales, dejando en evidencia de forma visual, el error asumido.

Distribución estimada como normal multivariada, para Markowitz

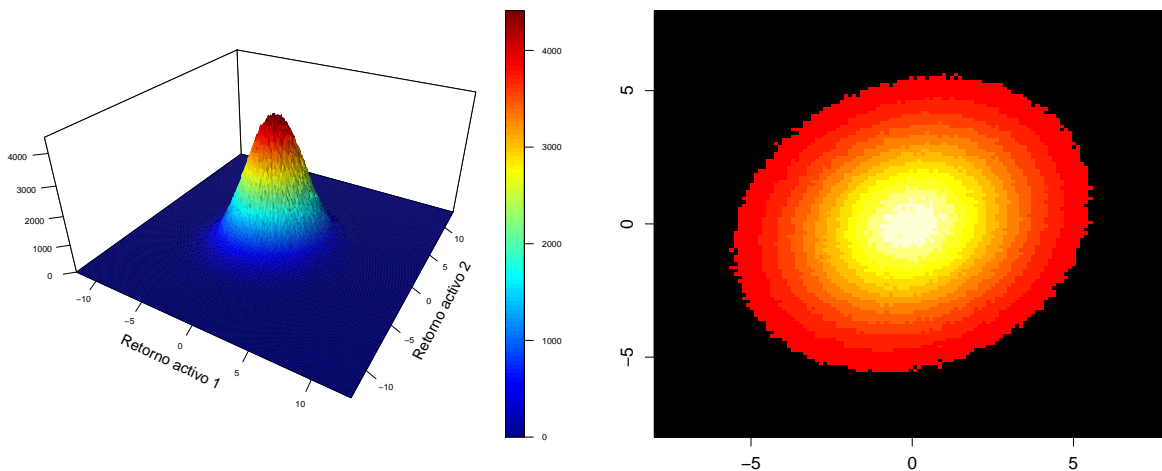


Figura 4.26: Histogramas de la distribución estimada.

A su vez, las distribuciones marginales de los instrumentos, como se pretendía, son asimétricos lo que se aprecia en la figura 4.28, en particular, la concentración y cola con mayor dispersión son opuestas, siendo una la rotación de la otra, entorno al promedio.

Posterior a la obtención de la frontera eficiente, en términos de varianza-retorno, de todos los

Histogramas 2D de la distribución real y estimada

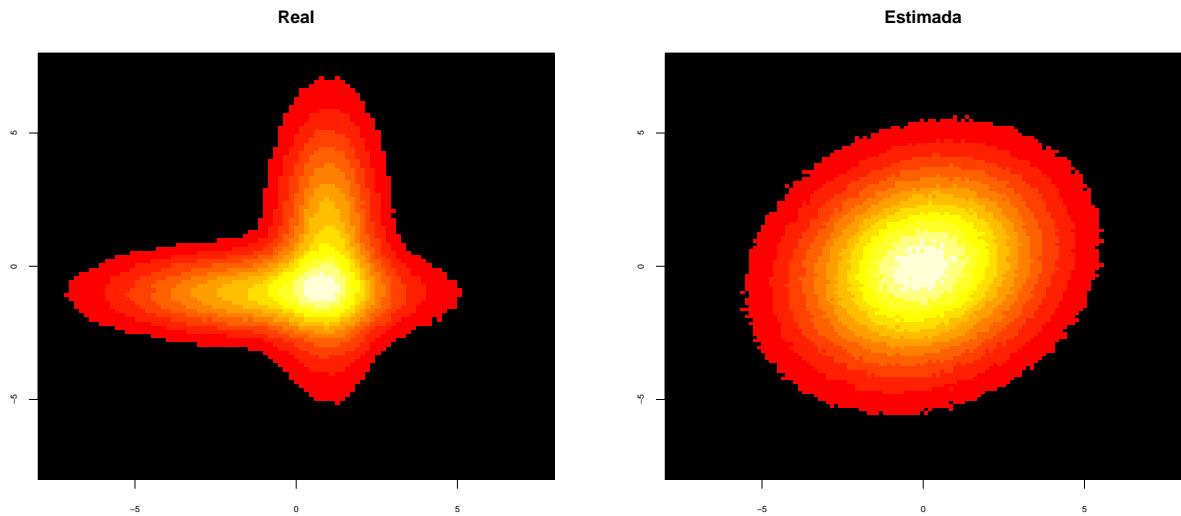


Figura 4.27: Comparación de los histogramas de ambas distribuciones

Distribuciones individuales

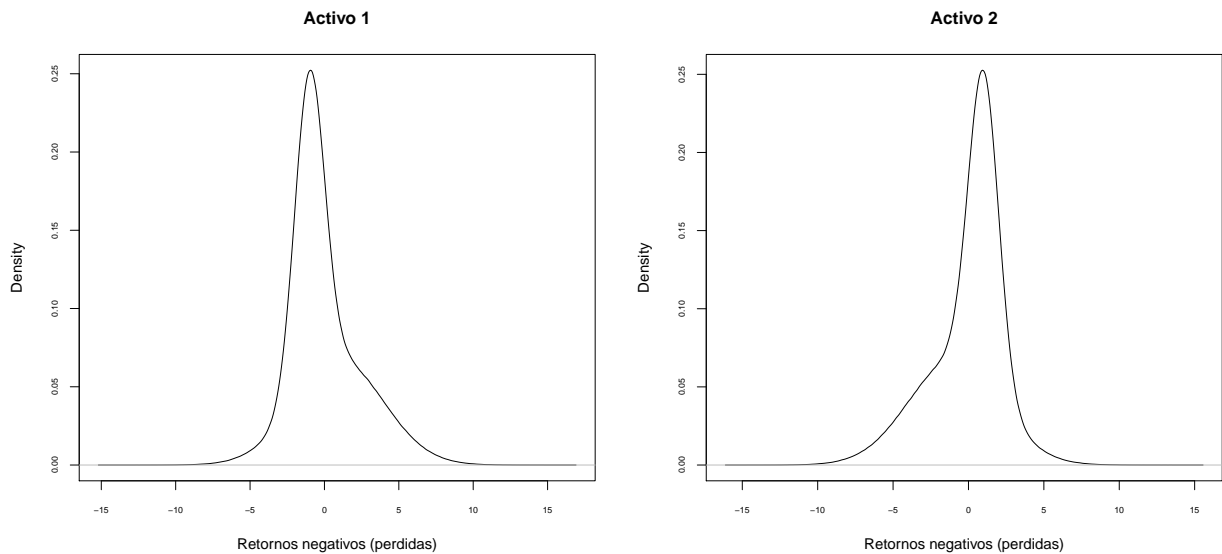


Figura 4.28: Distribuciones marginales de los instrumentos

modelos y rangos de aversión, incluyendo Markowitz con positividad, se analizan puntos específicos de la frontera, en particular, se analiza la *equivalencia* de estas soluciones, entre los diferentes portafolios.

De los portafolios particulares a analizar, se obtienen los perfiles entrópicos y espectrales, en particular, se evalúa la dominancia de soluciones iguales para Markowitz (retorno-varianza), con diferente riesgo entrópico y/o espectral.

Resultados y análisis

Como primer resultado, es interesante observar que las soluciones extremas de minimización del peor caso, para CVaR y Markowitz no son iguales, en particular, al minimizar la varianza, la solución óptima será tomar igual monto de inversión en ambos instrumentos, reduciendo la dispersión. Por el contrario, al considerar un CVaR extremo, es decir que considere sólo el peor 1 %, la solución se encontrará concentrada a la distribución con menor cola de pérdidas, tomando mayor posición sobre el segundo activo. Lo anterior también puede ser interpretado, como una solución de mínima varianza de Markowitz, que es dominada en riesgo CVaR, por la solución de CVaR al 1 %.

Al continuar haciendo el símil de Markowitz, con las otras medidas, es interesante observar que en esta instancia, ambos instrumentos tienen igual varianza, y por ello son indistinguibles para Markowitz, no obstante, son completamente distintos en términos de CVaR, un caso que ejemplifica lo anterior, son los portafolios extremos de concentrar todo en cada uno de los activos. Ambos portafolios tendrán igual varianza y valor esperado, no obstante, el CVaR considerando sólo el 10 % peor de la distribución es respectivamente 4.37 y 3.54, siendo dominada la primera solución, por la segunda.

Se obtiene la frontera eficiente varianza-retorno de todos los diferentes modelos, incluyendo en el análisis el clásico portafolio de Markowitz con positividad. Se espera como resultado de control, que todo punto de la frontera eficiente de Markowitz con positividad debe obtener la varianza mínima, de lo contrario Markowitz no sería óptimo en varianza. Como se puede observar del gráfico 4.29, se conserva la frontera varianza-retorno para todos los portafolios.

Las fronteras eficientes varianza-retorno, que se muestran en la figura 4.29, son construidas resolviendo la misma medida de riesgo, con valores del parámetro de aversión equi espaciados. Es interesante observar de las soluciones, que la elasticidad varianza aversión, varía fuertemente medida a medida, y en particular se concentra en soluciones extremas, lo que apreciamos al observar la densidad de portafolios óptimos en las soluciones de cada punta de la frontera.

Al buscar portafolios equivalentes en distribución de retornos, entre distintas medidas, para algún par de parámetro de aversión, quedó al descubierto la relación no lineal, y no monótona en términos de la varianza y el parámetro de aversión, que se muestra en la figura 4.30. Este mismo fenómeno, no se observa en las instancias simétricas, menos aun en aquellas normales.

Lo sucedido en la figura 4.30, se interpreta adecuadamente observando la evolución de la composición del portafolio condicional a la aversión. En el caso de CVaR, con niveles muy bajos de α , como se discutió anteriormente, se elige aquel instrumento con cola de pérdidas menor, concentrando toda la inversión en este instrumento con un α entorno a 0.18. Posteriormente con un α entorno a 0.53 se obtiene el portafolio de mínima varianza, invirtiendo equitativamente en ambos instrumentos, posteriormente a medida que aumenta la aversión, la inversión en el primer instrumento se intensifica, aumentando nuevamente la varianza.

El mínimo en varianza, alcanzado por CVaR con α de 0.53, como se dijo anteriormente, corresponde al punto de mínima varianza de Markowitz, a su vez corresponde al punto, donde se produce el cambio en composición del portafolio, que genera la zona inferior de la frontera varianza-retorno, la llamada frontera ineficiente en términos de retorno-varianza, dado que para todo punto de esta frontera inferior, existe un portafolio igual en varianza, con mucho mejor valor esperado.

Fronteras Eficientes en el espacio retorno–varianza

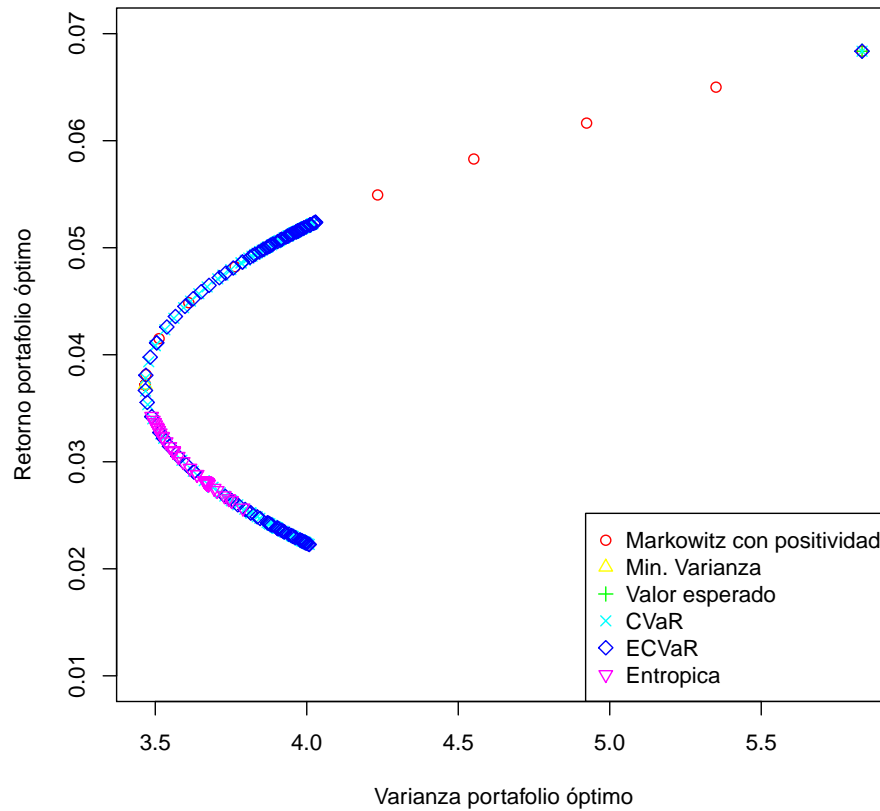


Figura 4.29: Frontera eficiente en espacio varianza-retorno, con portafolios equi-espaciados en parámetro de aversión

En el caso de la Medida de riesgo Entrópica, el comportamiento de la varianza condicional a la aversión, es aún más irregular. Se observa un cambio de comportamiento con θ entorno a 2.8.

Al interpretar los portafolios con parámetro de aversión sobre 2.8, el principal elemento a tener en mente, es la dispersión amplificada por θ de los retornos, que a su vez, se potencia por la exponencial, y en concreto, en el intervalo de aproximación de la exponencial, dejando cada vez más realizaciones fuera del intervalo de aproximación, considerando un menor número de realizaciones extremas u_i dentro del intervalo de aproximación. Los saltos que se observan, se producen cuando las realizaciones remanentes en el intervalo de aproximación, pasan a la aproximación lineal de otro punto de soporte.

Con aversiones bajo 2.8, se realizó un refinamiento de estos portafolios óptimos, obteniendo el gráfico 4.32. En él observamos que la distribución de la varianza, con respecto a la aversión, en este intervalo, es muy semejante a una distribución beta, con un máximo local con θ entorno a 0.72, donde posterior a este punto se produce un decaimiento exponencial. Interesante es observar una irregularidad a este comportamiento, que se inicia entorno a 2.55 y concluye entorno a 2.65, donde se produce un salto no lineal en la varianza, de valores entorno de 3.6 a valores entorno a 3.8.

Para verificar que la irregularidad observada anteriormente, no sea producto de un problema

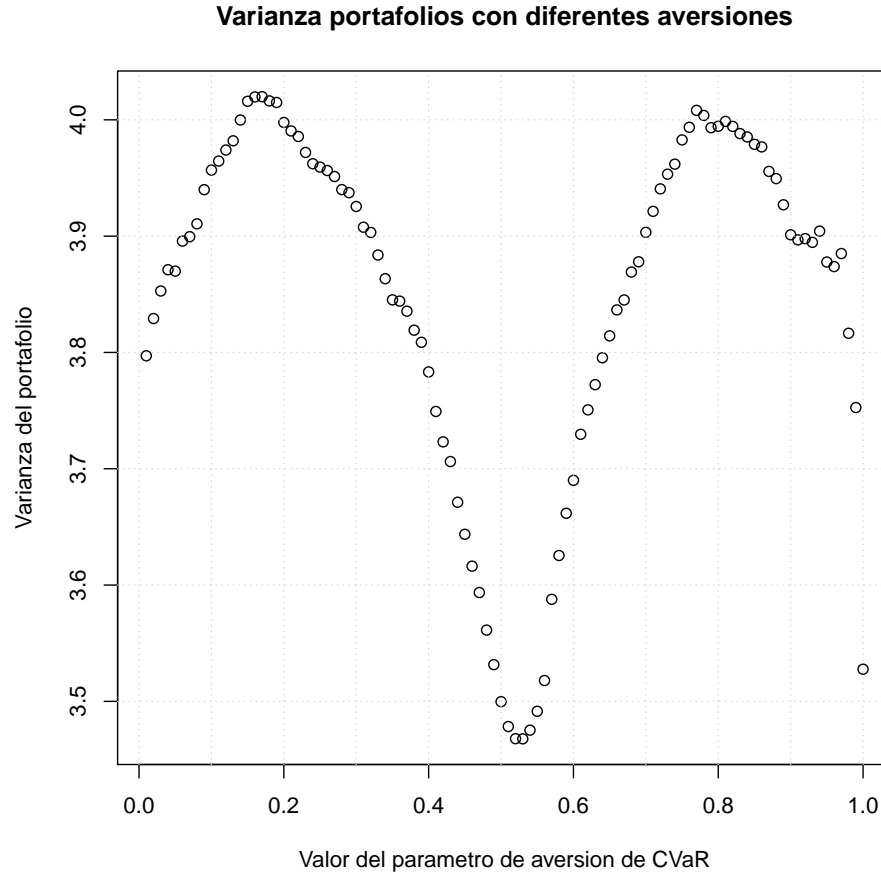


Figura 4.30: Curva de varianza de portafolios CVaR óptimos para diferentes valores de α

de convergencia de estos portafolios óptimos, se verifica la convergencia de ellos, obteniendo con parámetros SAA de 100 repeticiones, 20,000 in-sample, y 200,000 out-sample un gap relativo no trasladado promedio de 0.2 % entre los 11 portafolios con este comportamiento, para la confección del gráfico se tomó una grilla de tamaño 0.01 en el espaciamiento de θ . Se descarta que este comportamiento, sea producto de la aproximación de la exponencial, al pasar ciertas realizaciones de diferentes puntos de soportes de aproximación.

Al relacionar el resultado anterior de la medida entrópica, con la composición de los portafolios óptimos, obtenemos el gráfico 4.33. En el podemos observar, como se pasa de una solución predominantemente compuesta por el primer instrumento, a una compuesta en su mayoría por la segunda. Es interesante observar que el portafolio de mínima varianza, en el sentido Markowitz, se alcanza con un θ entorno a 1.5, adicionalmente se tiende asintóticamente a un portafolio con iguales pesos, esta vez aproximándose en sentido contrario en la composición.

Un aspecto relevante de acotar, es la presencia de portafolios dominados en el sentido Markowitz, es decir igual varianza y mejor retorno, en ciertos portafolios óptimos de la medida Entrópica. Esta misma situación se produce en el caso de portafolios óptimos de CVaR, en particular, en aquellos con α de aversión sobre 0.53. Esto genera un elemento adicional, a tener presente en la elección de la medida a usar, en instancias donde el valor esperado de los retornos es relevante.

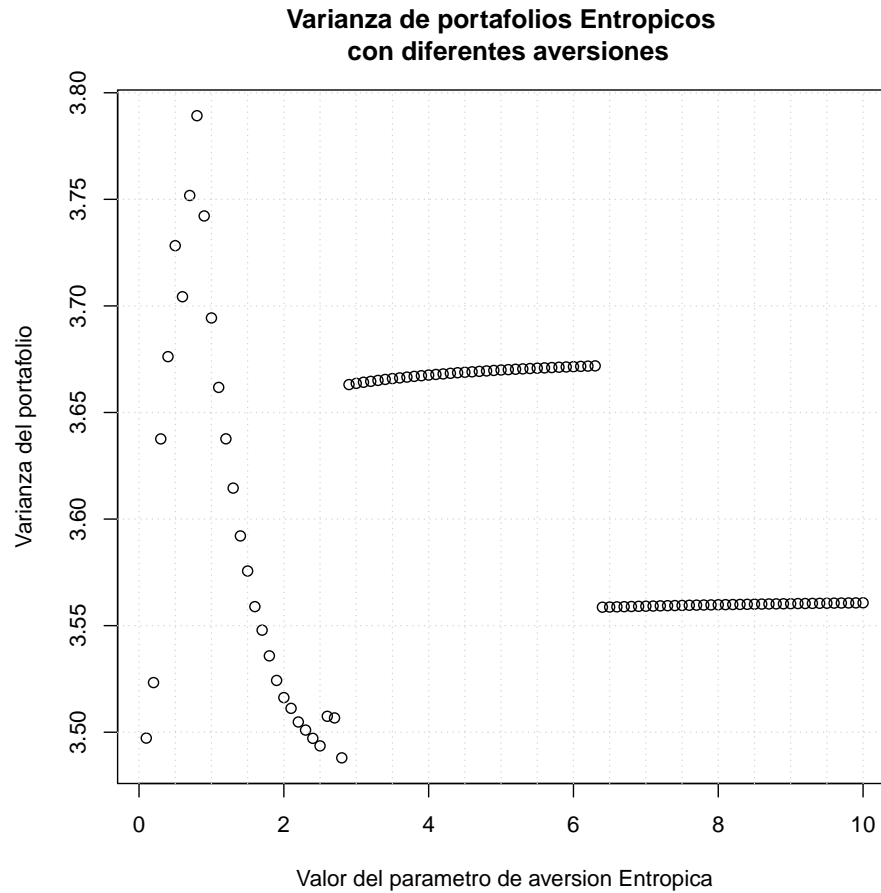


Figura 4.31: Curva de varianza de portafolios Entrópicos óptimos para diferentes valores de θ

Las curvas de las figuras 4.30 y 4.31, son de varianza de los portafolios óptimos CVaR y Entrópicos, para diferentes niveles de aversión respectivamente. Las curvas de ECVaR, por su naturaleza, son una traslación de CVaR por lo que se decidió omitir.

Al responder la siguiente pregunta de este experimento, referente al significado de cada riesgo, observamos que llevando al limite de máxima aversión cada medida, el portafolio generado por CVaR queda dominado por Markowitz en términos de varianza, no obstante, el portafolio de Markowitz obtiene un CVaR 0.1 % peor al óptimo de CVaR. Situación equivalente sucede con la medida de riesgo Entrópica. Lo anterior, deja de manifiesto que las soluciones en máxima aversión no son equivalentes, más aún coherente a lo esperado, cada modelo busca el óptimo bajo su significado particular de riesgo, que en esta instancia no son equivalente.

Para una fácil visualización de las dominancias antes descritas, en la siguiente tabla se resumen los resultados:

	Riesgo entrópico 40	CVaR al 0.1 %	Varianza	Skewness	Kurtosis
Portafolio Entrópico 40	6.18	5.51	3.58	0.30	2.89
Portafolio CVaR 0.1 %	6.78	5.71	3.71	0.44	3.05
Port. Markowitz min. varianza	6.25	5.92	3.47	0.00	2.76

Cuadro 4.6: Tabla resumen de los principales estadísticos de los portafolios óptimos extremos

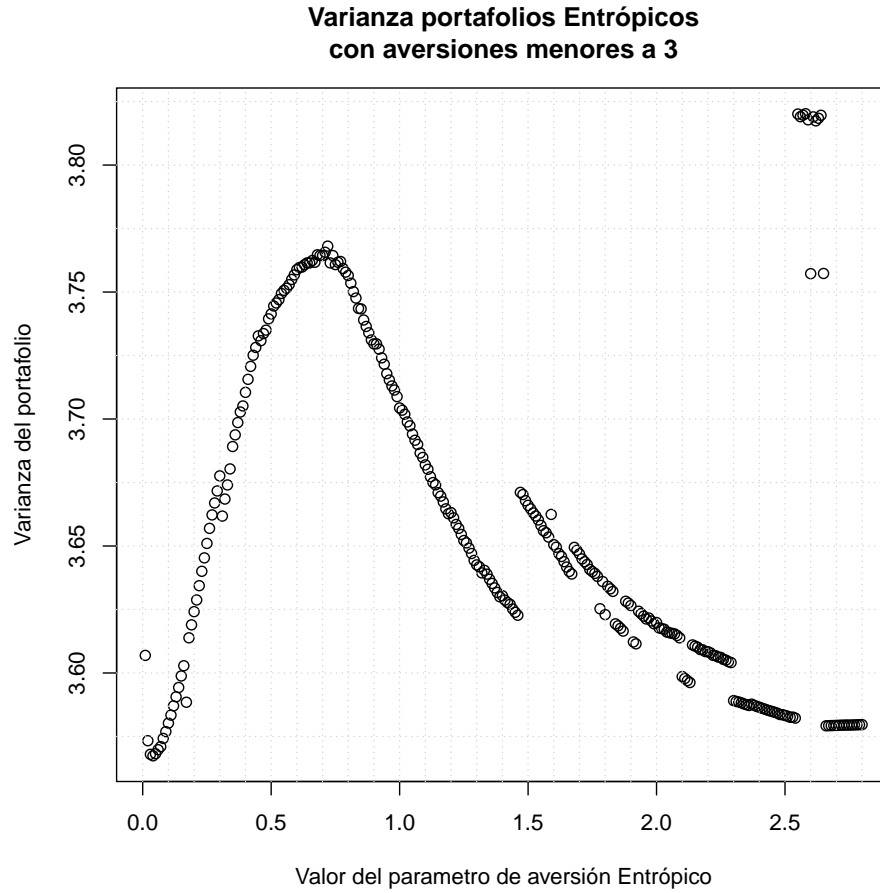


Figura 4.32: Curva de varianzas de portafolios Entrópicos óptimos, con aversión menor a 3

Adicionalmente, si observamos los histogramas de los portafolios óptimos a máxima aversión, en la figura 4.34, observamos como cada medida se traduce en su sentido práctica, es así como Markowitz genera el portafolio con mínima dispersión simétrica, CVaR minimiza el promedio de los peores resultados de la cola de pérdidas, y la medida Entrópica al modificar la medida de probabilidad de los peores casos, actúa sobre los momentos de orden superior de la variable aleatoria, lo que se traduce en una solución intermedia a las dos anteriores, pero con menor Skewness que el portafolio CVaR y Kurtosis más cercana a 3, lo que se puede verificar en la tabla 4.6.

4.4. Experimento Multitemporal

La resolución de un problema de inversión a largo plazo, que permite el rebalanceamiento del portafolio, en el contexto antes desarrollado de optimización con aversión al riesgo, puede ser expresado como:

$$\begin{aligned}
 & \min_{w_t \in \mathcal{P}_t} \rho_\theta(r|_T) \\
 & s.a. \\
 & w_t \in \mathcal{P}_t(\{w_{t'}\}_{t' \in \{0, \dots, t-1\}}) \quad \forall t \in T
 \end{aligned} \tag{4.3}$$

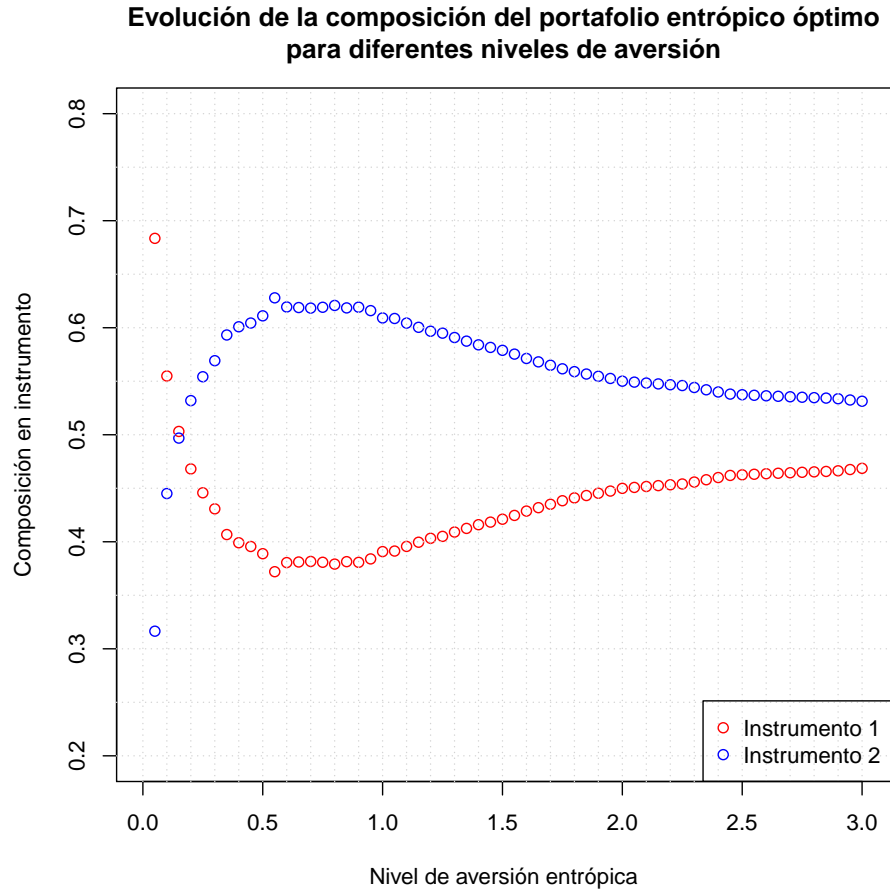


Figura 4.33: Evolución de composición de portafolios Entrópicos óptimos, con aversión menor a 3

Donde en el contexto de retornos logarítmicos, el retorno final en $|T|$ del portafolio, es igual a los retornos en cada periodo, es decir: $r_{|T|} := \sum_{t \in T, t > 0} r_{t-1,t}$, donde a su vez, $r_{t-1,t} := \vec{w}_{t-1} \cdot \vec{r}_t$.

La resolución de un problema intertemporal de portafolios, adicional a la componente de incertidumbre ya abordada, introduce la dificultad de una estructura intertemporal, observable en restricciones entre periodos, que vinculan la solución óptima del periodo anterior, con el problema en resolución del periodo, y la dificultad adicional de aleatoriedad que se revela en diferentes momentos, lo que se traduce en un número mayor de variables aleatorias. Las restricciones anteriores, considerando adicionalmente el crecimiento del capital dentro del periodo, hace que sean interpretables como restricciones de no anticipatividad, haciendo imposible su resolución como problemas separados.

En un contexto general, el problema 4.3, debería ser resuelto mediante programación dinámica estocástica, asegurando una solución global. En este experimento, se aborda mediante una estrategia secuencial, que llamaremos de *rolling forward*, donde en una trayectoria (realización) particular, se resuelve de forma local cada periodo, utilizando la resolución del periodo anterior w_{t-1} , en la resolución del problema vigente para t .

La relación entre periodos antes mencionada, se introduce mediante las restricciones de iliquidez, donde para un activo $i \in I$, se acota la variación entre periodos como $|w_i^t - w_i^{t-1}| \leq \delta(w_i^{t-1})$

Histograma de los tres portafolios optimos a maxima aversion

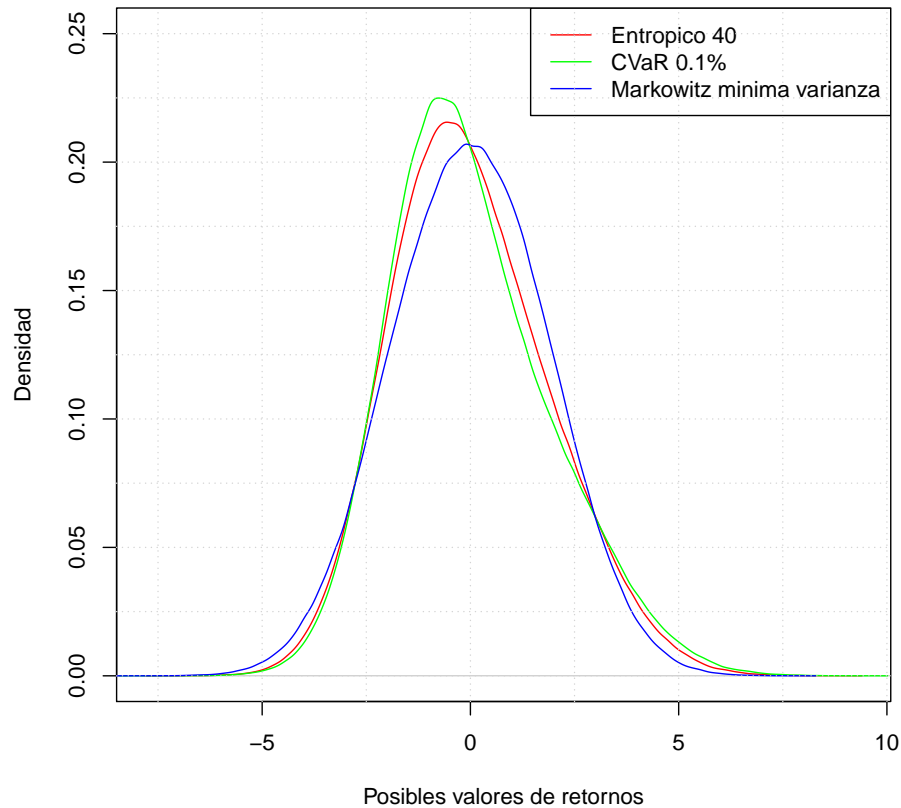


Figura 4.34: Histograma de los tres portafolios óptimos extremos

$\forall(t, i) \in T' \times I^3$, donde en este caso particular $\delta(w_i^{t-1})$ será una cota porcentual de w_i^{t-1} ($w_i^{t-1}\delta$), parámetro que en el problema en t , para una trayectoria dada, es un parámetro fijo.

Es claro observar, que esta solución secuencial, no necesariamente entregará una solución global al problema 4.3, no obstante, la lógica *rolling forward* de resolución, replica lo que sucede habitualmente con las AFPs, donde se toma una decisión de estructuración, llega el final del período, se actualizan los precios y sus creencias, se vuelve a rebalancear el portafolio - considerando restricciones de liquidez - y el proceso se vuelve a repetir iterativamente.

En estas instancias al problema de optimización definido inicialmente, se le agregarán las restricciones de mercado, que buscan emular la iliquidez del mercado, limitando las variaciones máximas de las posiciones. Al poliedro de soluciones factible inicial \mathcal{P} , se le agrega las restricciones $|w_i^t - w_i^{t-1}| \leq \delta w_i^{t-1}$ donde para $t \in T$, la posición en el instrumento i de w_i^{t-1} , es un dato fijo.

³ $T' := T - \{0\}$

4.4.1. Diseño experimental

Como fue mencionado anteriormente, se resolverá un conjunto de problemas $\{\mathcal{L}_t\}_{t \in T}$, de donde se obtendrá el conjunto de soluciones $\{w_t^*\}_{t \in T}$. Donde cada problema es de la forma:

$$\begin{aligned}
 \mathcal{L}(\vec{w}_{t-1}) &:= \min_{\vec{w}_t} \rho_\theta(\vec{w}_t \cdot \vec{r}_t) \\
 \text{s.a.} & \\
 \vec{w}_t &\in \mathcal{P} \\
 w_t^i &\in B(w_{t-1}^i, \delta) \quad \forall (i, t) \in \text{Ix}\{T - \{0\}\}
 \end{aligned} \tag{4.4}$$

Donde \mathcal{P} hace referencia al poliedro del problema de portafolios en dos tiempos, definido anteriormente en 2.1, que posee las restricciones de: positividad, inversión total y de régimen de inversión de las AFPs en Chile. La función ρ_θ , hace referencia a la medida de riesgo particular utilizada en cada portafolio. El parámetro θ , fija la aversión al riesgo de la medida, que se considera fijo durante todo el periodo de resolución, lo que supone que los agentes son consistentes intertemporalmente, y no cambian en aversión, y finalmente, la restricción adicional de iliquidez, que puede ser interpretada como que la nueva posición w_t^i , se encuentre en la bola de centro w_{t-1}^i y radio δ , en la instancia, se consideró por simplicidad, una fricción por iliquidez fija del 10%.

La instancia es generada con 30 instrumentos representativos de las inversiones de los fondos de pensión. Este conjunto de instrumentos esta compuesto por instrumentos de renta variable (RV) nacional, instrumentos de renta fija nacional (RF), Exchange-Traded Funds (ETFs)⁴ extranjeros, representantes de inversión en Estados Unidos, Europa y Asia.

Los instrumentos considerados en el panel de datos, fueron seleccionados de un original de 122 instrumentos que representaban el 80% de los instrumentos transados de los fondos, toda esta información fue obtenida mediante un Crawler Web diseñado y programado para obtener todos estos datos, desde la cinta de precios de la Superintendencia de Pensiones de Chile, de acceso público. La selección de los 30 instrumentos, de los 122 originales, se realizó con los criterios de: (1) Existencia del instrumento en la ventana completa de estimación y (2) Variación del precio, como indicador de transabilidad y liquidez, en al menos el 90% de las observaciones.

Para la mezcla de los dos estados (crisis y no crisis) normales multivariados, se considera una estimación de los parámetros $\{\{\mu_i^s, \sigma_i^s\}_{i \in I}; \{\rho_{i,j}^s\}_{(i,j) \in I^2}; \{\Pi_{s,s'}\}_{s' \in S}\}_{s \in S}$ que definen el proceso conjunto de los retornos. Para esta estimación, sobre el panel de datos antes descrito, se considera una ventana de tiempo móvil de 100 días, de donde se estima los procesos normales base de la mezcla. La ventana móvil, se mueve sobre el soporte de diez años desde el 2003 al 2013, donde se cuenta con las observaciones diarias.

En la estimación de los parámetros para los estados de crisis y expansión (no crisis) del ciclo económico, se realiza el proceso de estimación, separando el conjunto de datos diarios según el estado del ciclo económico al cual pertenece $s \in S$, para esto se utilizo la información proporcionada por el *Business Cycle Dating Committee* de la *National Bureau of Economic Research*⁵ como

⁴Instrumentos de inversión que son generados minimizando el tracking error con algún índice, a su vez, se les exige condiciones de liquidez y de representación de riesgo equivalente en exposiciones que no hacen referencia al riesgo de mercado, como es el riesgo crediticio.

⁵Fuente: <http://www.nber.org/cycles.html>

	$\mathbb{E}(r_i)$	σ_i	$\mathbb{E}(r_i s=c)$	σ_i^c	$\mathbb{E}(r_i s=nc)$	σ_i^{nc}	Categoría	NZ
SP500	0.012	1.271	0.064	1.57	0.001	1.2	ETF Internacional	97.2
STOXX50	-0.002	1.461	0.034	1.723	-0.009	1.401	ETF Internacional	91.4
Nikkei.225	0.018	1.532	0.024	1.658	0.016	1.505	ETF Internacional	91.5
Hang.Seng	0.023	1.568	0.084	1.863	0.01	1.5	ETF Internacional	90.5
IBOV	0.026	1.735	0.08	1.65	0.014	1.752	ETF Internacional	91.8
IGBVL	0.063	1.557	0.134	1.504	0.048	1.568	ETF Internacional	94.8
IPyC	0.052	1.262	0.062	1.373	0.05	1.238	ETF Internacional	94.8
BRP	0.031	0.25	0.026	0.264	0.032	0.247	RF Nacional	100
LHF	-0.014	0.25	-0.008	0.276	-0.016	0.245	RF Nacional	100
DPF	0	0.054	0.007	0.068	-0.002	0.05	RF Nacional	100
PRC	-0.029	0.284	-0.011	0.328	-0.033	0.274	RF Nacional	100
DEB	0.003	0.226	0.027	0.235	-0.002	0.224	RF Nacional	100
CERO	0.008	0.253	0.003	0.262	0.009	0.251	RF Nacional	100
BCS	-0.006	0.291	-0.005	0.322	-0.006	0.284	RF Nacional	100
BEF	0.018	0.32	0.032	0.224	0.016	0.337	RF Nacional	100
BSF	-0.006	0.31	0.007	0.338	-0.008	0.304	RF Nacional	100
BCU	-0.001	0.232	0.004	0.219	-0.002	0.234	RF Nacional	100
BCP	-0.003	0.189	0.003	0.216	-0.004	0.183	RF Nacional	99.9
COPEC	0.027	1.463	0.113	1.524	0.009	1.45	RV Nacional	93.7
ENDESA	0.051	1.284	0.051	1.023	0.05	1.332	RV Nacional	94.2
ENERSIS	0.047	1.496	0.089	1.24	0.038	1.544	RV Nacional	94.1
ENTEL	0.057	1.428	0.065	1.084	0.055	1.49	RV Nacional	93.3
COLBUN	0.037	1.386	0.081	1.169	0.028	1.427	RV Nacional	92.4
SQMB	0.095	1.999	0.091	1.76	0.096	2.046	RV Nacional	93.2
BSANTANDER	0.036	1.564	0.162	1.72	0.009	1.528	RV Nacional	91.3
FALABELLA	0.087	1.663	0.233	1.453	0.057	1.702	RV Nacional	92.8
CMPC	0.037	1.534	0.177	1.724	0.008	1.491	RV Nacional	90.9
CAP	0.123	2.153	0.212	1.911	0.104	2.2	RV Nacional	92.1
LAN	0.086	1.686	0.204	1.485	0.061	1.724	RV Nacional	92.2
CCU	0.062	1.434	0.123	1.553	0.05	1.408	RV Nacional	92.4

Cuadro 4.7: Estadística descriptiva de los retornos logarítmicos de los instrumentos

variable de clasificación. Esta fuente es aceptada tanto por la industria, como la literatura en este ámbito. En aquellos periodos, donde la ventana móvil de estimación, no contenía los 100 registros, se complementa con los datos históricos anteriores hasta completarlo.

De la estimación de los parámetros de la distribución generadora de retornos, a modo de resumen, se presenta el cuadro 4.7 con el detalle estadístico de cada uno de los instrumentos, considerando toda la ventana de datos, mostrando su primer y segundo momento sin separar los datos y los respectivos estados de crisis (c) y no crisis (nc) en la instancia mezcla, considerando todos los eventos clasificados en el panel de datos, adicionalmente, se muestra la categoría del instrumento y porcentaje de observaciones no cero (NZ).

Es interesante observar del cuadro 4.7, que los procesos normales multivariados al ser separados en el estado de crisis y no crisis, muestran promedios y varianzas, condicionales al estado, muy diferentes, más aún estas distribuciones, considerando la probabilidad empírica de crisis, en la ventana de 10 años, reúnen las condiciones necesarias para generar distribuciones de retornos como las generadas artificialmente en los experimentos monotemporales anteriores, dejando al descubierto, la presencia de procesos asimétricos en la instancia real.

Como objetivo del experimento, se pretende evaluar el desempeño de los portafolios propuestos, en comparación con los actuales participantes de mercado. Todo lo anterior, con el supuesto y en el contexto de la única historia conocida. Se considera este modelo siempre en la generación de retornos, y esta *backtesting* con una única historia, la historia realizada, para simplificar lo referente a modelos en los procesos generadores, lo que sale del alcance de este trabajo.

4.4.2. Resultados y análisis

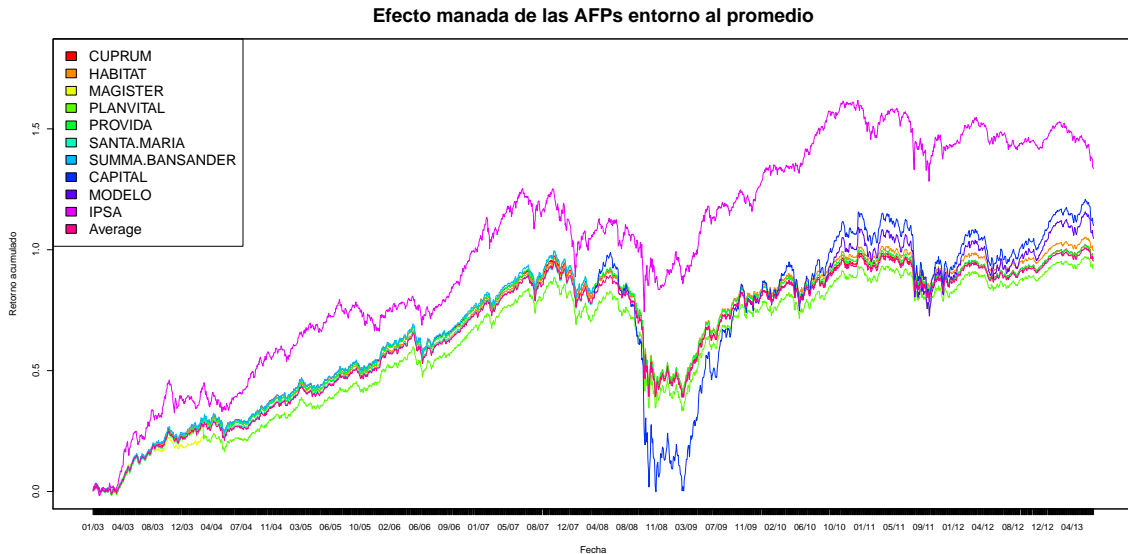


Figura 4.35: Gráfico del retorno acumulado, del valor cuota de los fondos A de cada AFP.

Coherente a lo descrito en la literatura del tema, en el gráfico 4.35, se observa a simple vista, el comportamiento de manada y concentración que produce tener como *benchmark*, el promedio de los demás participantes en el periodo anterior, concentrándose entorno a él. Se observa un fondo, con color azul, con una estrategia más riesgosa, y mayor tiempo de reacción, que ante la crisis económica de 2009, presentó una lenta reacción, siendo por lejos el participante con peor desempeño en la crisis. En el mismo gráfico, se introduce como elemento de referencia de rentabilidad, color morado, completamente independientes, el IPSA, principal indicador de mercado agregado en renta variable en Chile, este indicador aún cuando presenta un comportamiento en rentabilidad y regulación completamente distinto, nos permite tener una impresión de las condiciones de rentabilidad del mercado.

Si realizamos el análisis de la distribución de retornos, del proceso periodo a periodo, de los diferentes participantes, y en particular, del promedio de la industria, obtenemos el gráfico 4.36, de donde en conjunto con el gráfico 4.35, podemos concluir que para análisis posterior, es suficiente trabajar como referencia, sólo con el promedio de la industria, y no con cada participante de forma agregada. Esto permite simplificar los gráficos posteriores, con los resultados de cada modelo.

En el gráfico 4.36, adicionalmente se observa una distribución fuertemente no simétrica. En general las diferentes distribuciones se aprecian como semejantes, coherente a lo observado en el gráfico 4.35, la AFP *Summa Bansander* que posteriormente fue fusionada, presenta un comportamiento levemente diferente al del resto de los participantes del mercado. Adicionalmente, podemos observar

Distribución del retorno diario para las diferentes AFPs

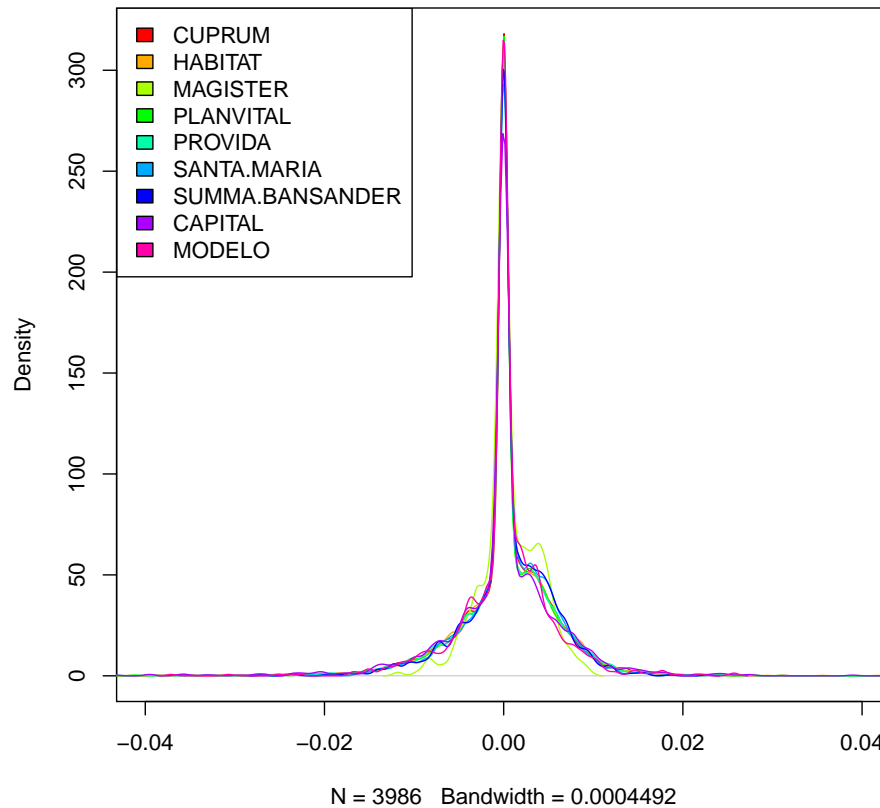


Figura 4.36: Distribuciones de retornos del proceso diario del fondo A, de los diferentes agentes, y del promedio anterior

agentes pro activos, con mejor gestión de la cola de pérdidas, y reactivos, con colas levemente más irregulares.

Se resuelve secuencialmente los problemas de portafolio, para cada medida, para simplificar el análisis, se toman los portafolios considerados más representativos de cada medida, se escoge el CVaR observando sólo el peor 1% pero de los retornos, con $\alpha=0.01$, ECVaR considerando CVaR observando el 75% de la distribución ($\alpha=0.75$), y considerando una componente de valor esperado del 20% ($\epsilon=0.2$), y finalmente, la medida de riesgo Entrópica con parámetro de aversión igual a 100 ($\theta=100$).

Se calculan los portafolios óptimos, de los problemas antes enunciados, obteniendo los resultados por cada periodo. Con estos resultados, ocupando la información histórica a cada periodo, se obtiene el gráfico 4.37, de retornos acumulados para cada portafolio óptimo, y el promedio de los demás participantes del mercado (AFP's), para igual fondo, en iguales periodos.

Se observa claramente la caída del sistema, producida por la *Crisis de hipotecas subprime* entorno a principios del 2009, y la reacción de los participantes y de los modelos. Es consistente observar, que aquellos modelos más adversos, presentan una caída menor, cuando en contra posición los otros modelos presentan una mayor sensibilidad, y tiempo de reacción, producto de las restricciones de

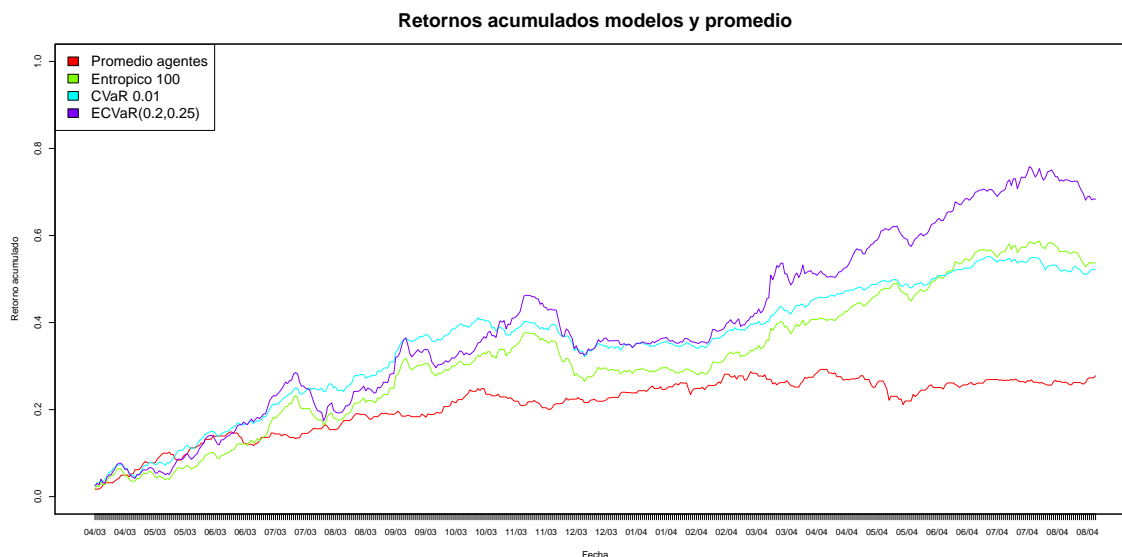


Figura 4.37: Gráfico de retornos acumulados, de los diferentes portafolios óptimos, y el promedio de los participantes.

liquidez, en el cambio de composición del portafolio, todos los efectos mencionados se ven reflejados también, de forma consistente en la distribución de los retornos, y en el perfil espectral. Es interesante observar el comportamiento contra cíclico de ECVaR y la medida Entrópica, que exagera el crecimiento en periodos donde CVaR y el promedio de la industria bajan, fenómeno que tiende a caer, cuando dicho crecimiento se frena y/o revierte.

Como fue mencionado anteriormente, introducir algún análisis adicional, sobre el desempeño de los demás participantes (AFPs), sin considerar ninguna histórica adicional a la conocida, más aún, sin tener algún modelo valido de su reacción frente a otras realizaciones, resulta imposible. Por otro lado, evaluar otros modelos generadores de retornos, con actualización de expectativas, o reacciones definidas a eventos conocidos en la histórica, que es el usual enfoque utilizado.

Es importante tener en cuenta los supuestos realizados con este análisis, más allá del sólo hecho de estar ocupando la única historia conocida, que es una realización particular. Se asume que cada agente, puede ser caracterizado, por su distribución de retornos de inversión diaria, la que se asume única y transversal a todo el periodo, donde el retorno acumulado a largo plazo, es obtenido de la suma de estos retornos individuales. Esto permite analizar la distribución generada, por las diferentes AFPs y los modelos implementados. Como supuestos adicionales, se considera valido el modelo de dos estados, con procesos subyacentes normales multivariados, con ventana de tiempo móvil para su ajuste sólo con datos observados al instante de resolución. Se escogió este modelo, por su simplicidad.

De los resultados de cada modelo, podemos apreciar como, dada la asimetría de las distribuciones generadoras de retornos, los diferentes modelos, difieren en el significado de sus nociones de riesgo, generando portafolios distintos, con trayectorias agregadas distintas.

Adicionalmente, para facilitar la comparación de los diferentes portafolios, en el cuadro resumen 4.8, se entregan los principales estadísticos descriptivos de los procesos.

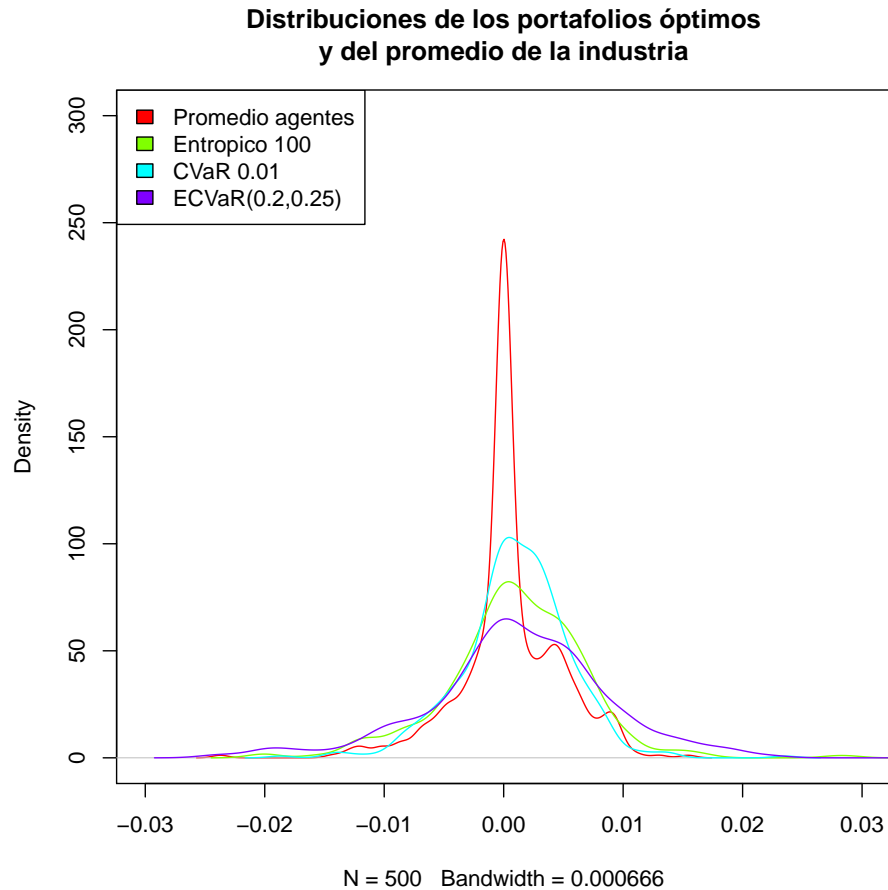


Figura 4.38: Distribuciones de los procesos de retornos, periodo a periodo, para los modelos implementados, y el promedio del sistema.

Del cuadro 4.8, es interesante observar, que en los diferentes riesgos, con excepción de la varianza, el peor desempeño lo obtiene el promedio de los demás participantes del mercado. Como es consistente de esperar, cada portafolio óptimo, es el mejor en su respectivo riesgo asociado. El portafolio con mejor Skewness, aspecto muy valorado en la industria, es el portafolio Entrópico, siendo el segundo mejor el promedio del desempeño de las AFPs en el periodo anterior.

Adicional al análisis descriptivo anterior, es interesante conocer el perfil de riesgo espectral, de los diferentes portafolios evaluados, lo que se puede apreciar en el gráfico 4.39.

Coherente a lo observado en las densidades en el gráfico 4.38, observamos que en el promedio de los agentes, se produce un quiebre entorno a $5E-3$, producto de una fuerte concentración de eventos en esa zona, para posteriormente, presentar una cola derecha, muy delgada.

	Riesgo entrop. 100	CVaR al 1%	Promedio	Varianza	Skewness	Kurtosis
Promedio agentes	3.03E-03	2.64E-02	2.60E-04	3.30E-05	-6.47E-01	1.82E+01
Entropico100	7.58E-04	1.82E-02	1.04E-03	3.37E-05	-6.99E-02	5.48E+00
CVaR 0.01	6.86E-05	1.46E-02	1.01E-03	2.03E-05	-2.02E-01	5.15E+00
ECVaR(0.2,0.25)	2.01E-03	2.24E-02	1.01E-03	2.03E-05	3.43E-01	7.18E+00

Cuadro 4.8: Tabla resumen de estadísticos descriptivos, del proceso de retornos periodo a periodo

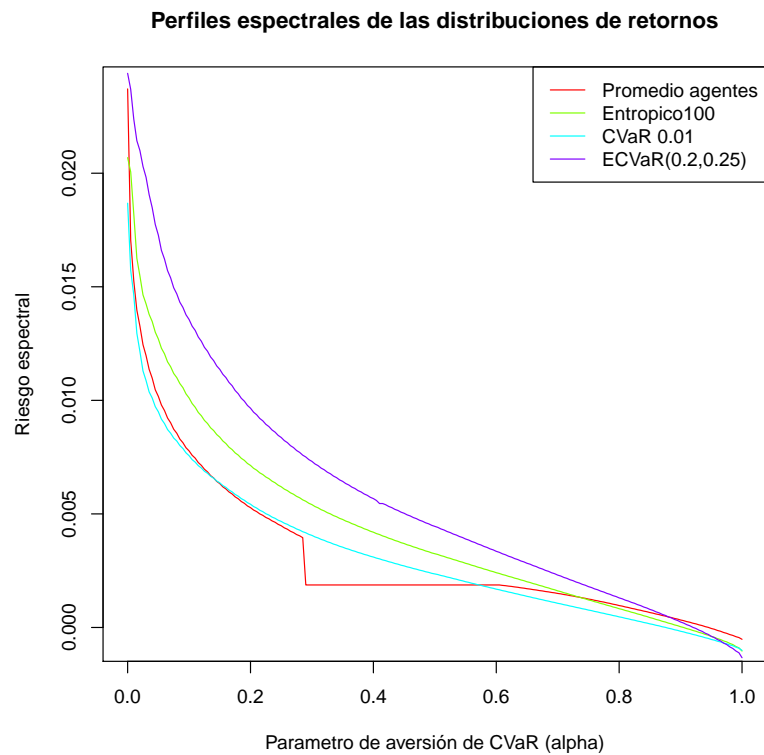


Figura 4.39: Perfil espectral de los diferentes portafolios, en el problema multiperiodo

Capítulo 5

Conclusiones y trabajo futuro

5.1. Conclusiones

Referente al objetivo metodológico y aplicado del trabajo

La metodología de resolución del problema de optimización de portafolios con la medida de riesgo entrópica, fue satisfactoria, y en general, la resolución iterativa, vía métodos de aproximación lineal de la vecindad óptima, en problemas no lineales, resultó ser un enfoque efectivo. Este enfoque es extensible, con iguales resultados esperados en otros problemas convexos, dada la pertinencia de cortes lineales.

Las soluciones de las medidas de riesgo implementadas, mostraron una clara diferenciación sólo en las instancias asimétricas generadas, en el caso de distribuciones con colas gordas simétricas, para encontrar soluciones iguales con diferentes medidas, se produce un desplazamiento en los parámetro de aversión involucrado.

Las propuestas de portafolios resueltos, pueden ser utilizadas satisfactoriamente como benchmark de rentabilidad por fondo, en particular, como se observó en el experimento multitemporal, las soluciones resultan ser competitivas con el promedio de los actuales fondos, capturando las mismas restricciones y fluctuaciones del mercado. Separando la componente de habilidad de gestión de un fondo, con las fluctuaciones propias del sistema, y entregando una real gestión del riesgo.

Para la implementación de la metodología propuesta, como benchmark de rentabilidad para el sistema, es crítico validar las entradas del sistema, considerando en éstas las restricciones y los procesos generadores de retornos. En particular, dado el objetivo con que el sistema fue implementado, resulta necesario que esta discusión sea guiada con un horizonte a largo plazo, coherente al objetivo del sistema.

La correlación entre instrumentos, genera una dinámica entre los fenómenos de diversificación y cobertura antes descritos, no obstante, no produce distribuciones asimétricas que diferencien las soluciones de las diferentes medidas. En este sentido, un aspecto no explorado, es el efecto de la correlación en una instancia de mezcla de normales, donde la correlación se produzca entre

las componentes normales de cada estado, este enfoque necesitaría un marco más general a lo desarrollado, pero permitiría abstraer a un modelo, donde los estados de crisis y expansión se encontrarían relacionados y latentes en cada instante.

Referente al Sistema de Pensiones en Chile.

El actual Sistema de Pensiones en Chile, enfrenta fuertes cuestionamientos públicos, algunos de ellos con argumentos y fundamentos contradictorios. En particular, en el presente trabajo se mostró una metodología implementada en un sistema, que de forma potencialmente automática, haría entrega de retornos acumulados semejantes a los alcanzados por las AFPs, con la misma información y restricciones de inversión. En este sentido, cabe preguntarse por la razón de la diferencia entre la comisión de administración, que entre partícipes difiere, obteniendo todos retornos acumulados a largo plazo semejantes (no así sus perfiles de riesgo).

El actual sistema de pensiones mixto implementado, se define como un sistema con dos pilares: uno privado de capitalización individual, donde cada individuo guarda sus ahorros, los que son invertidos para fondos de vejez; y de uno público, de carácter solidario que actúa como un seguro, garantizando una pensión mínima. En este sentido, la pensión mínima se encuentra regulada por este mecanismo, y por ende la discusión se debería enfocar en: (1) qué se entiende por una pensión mínima, (2) cómo financiar un sistema de esta naturaleza a largo plazo, y finalmente, (3) qué mecanismo de incentivos implementar, para que los individuos ahorren para su vejez, acorde al real perfil de ingresos y costo de vida en el país.

Finalmente, queda de manifiesto al analizar la composición de inversiones de las AFPs, que el sistema de pensiones en Chile, no sólo busca administrar los fondos de pensión, sino que ha sido un pilar fundamental en la liquidez y financiamiento del sistema económico nacional.

Referente a las técnicas empleadas

Mediante la metodología SAA, para la resolución del problema estocástico generado al introducir las medidas de riesgo, se logró tener certeza estadística de la significancia de las soluciones.

La metodología empleada para la generación de los retornos, permitió generar colas no simétricas y gordas, que diferenciaron el comportamiento de las medidas, no obstante, son necesarios experimentos adicionales para conocer en profundidad su comportamiento, en particular convergencia, ante distribuciones más generales. En esta misma línea, el generador de mezcla de normales implementado, cumplió con la implementación de instancias con distribuciones asimétricas, y de colas gordas, no obstante, limitar los diseños a sólo dos estados, imposibilitó la evaluación de instancias *heavy tails*, donde es necesario contar con mezclas de más de diez estados, o en su efecto, evaluar otros generadores.

La utilización de cortes lineales, para aproximar la vecindad convexa a la solución, complementando este enfoque con un método de escalamiento, vía descomposición de promedio ($X = \mu(X) + (X - \mu(X))$), por temas numéricos, y una heurística de ajuste *On-the-Fly* de los cortes de aproximación, resultó ser un enfoque suficiente y efectivo en el balance convergencia/tiempo de resolución del problema entrópico.

El real costo computacional involucrado en la resolución de problemas de optimización estocástica, por lo general, no se encuentra en encontrar candidatos a solución de realizaciones parti-

culares, sino en asegurar la convergencia y por ende significancia estadística de las soluciones. En este sentido, SAA entrega un marco de trabajo, y condiciones muy generales, no obstante, complementar esta generalidad con técnicas *ad hoc*, propias de particularidades del problema, permite reducir estos tiempo de cómputo a igual convergencia, lo que resulta crítico en aplicaciones.

Referente al enfoque de aversión al riesgo y las medidas de riesgo

El enfoque de optimización con aversión, vía medidas de riesgo en el problema de portafolios, generaliza lo desarrollado en [58], en particular, ante instancias no simétricas las diferencias en la noción de riesgo, se traduce en soluciones y perfiles de riesgo distintos. Al buscar concluir sobre que metodología o noción de riesgo utilizar, la respuesta dependerá fuertemente del objetivo buscado en cuanto al tipo de riesgo a controlar, y del tipo de distribuciones de entrada. En particular, ante instancias simétricas, un enfoque retorno-varianza es suficiente, considerando otras soluciones equivalentes, que serán obtenidas con mayor costo de cómputo.

Al pretender concluir en un significado más práctico del comportamiento de la medida de riesgo entrópica, y en particular, de los resultados obtenidos en las instancia no simétricas, viene a la mente recordar una expansión en serie de Taylor de la exponencial, donde se encuentra explicación a los momentos de orden superior ($\sum_{n=0}^{\infty} \frac{(-\theta)^n}{n!} \mathbb{E}(X^n)$), lo cual resulta imposible dejarlo fuera de una conclusión, por su simplicidad y belleza.

Referente a la aplicación en otros campos

Lo desarrollado en el contexto del problema de portafolios, fue aplicado en otros campos, en particular, durante la realización de este trabajo, la medida entrópica y la metodología desarrollada para su aproximación y resolución como problemas lineales (un problema por cada repetición), fue utilizada en conjunto con la librería OpenCV¹, para el procesamiento de imágenes en la detección de eventos, en particular, para procesar por componente (RGB), la distribución generada de la diferencia entre matrices de frames, como una extensión al *Entropy filter*. Este método resulto ser más robusto que la varianza (por la asimetría de las distribuciones), y mostró mejor desempeño que el CVaR, generando menos falsos positivos.

A modo de síntesis, la metodología desarrollada para la resolución de la medida de riesgo entrópica, puede ser aplicada a cualquier problema que: (1) Minimice $\rho_{ent}(F(\cdot))$, donde la función $F(\cdot)$ que da estructura al problema, sea lineal. (2) El poliedro sea generado por cortes lineales, y finalmente (3) las distribuciones generadoras cumplan las condiciones de convergencia de SAA.

¹<http://opencv.org>

5.2. Trabajos futuros

Del presente trabajo, se consideran trabajos futuros que extienden y complementan lo desarrollado los siguientes focos:

- Automatizar el proceso de: colección, procesamiento, optimización y reporte. A fin de entregar periódicamente, un antecedente de acceso público, vía portal web, que permita conocer un indicador de rentabilidad base, separando el efecto de gestión y habilidad de administración, de las fluctuaciones propias del mercado. Entregando un antecedente técnico a la discusión sobre la rentabilidad de los fondos.
- Modelar el problema de incentivos, y no de rentabilidad que enfrentan las AFPs, minimizando el tracking error al promedio mensual de los otros participantes, agregando restricciones de rentabilidad al modelo. Todo lo anterior, como un modelo de equilibrio general con los demás agentes, buscando emular el mercado en su conjunto. El enfoque anterior, es crítico vía modelos de equilibrio general, dado que es conocido que las AFPs en Chile, mueven el mercado. Este trabajo, permitiría evaluar en un ambiente general, las regulaciones vigentes en el área.
- Modelar y resolver el problema de ruta óptima, en el cambio de fondo, de los diferentes agentes. En la actualidad, la regla implementada como estándar en el sistema, carece de un sustento técnico sólido, y sólo se basa en edad, cuando otras dimensiones como monto, y tiempo para la jubilación son relevantes.
- Resolver el problema global multitemporal, con un enfoque de optimización dinámica estocástica, donde en particular, el solver desarrollado será utilizado en la resolución de los periodos, resolviendo el problema desde el final hacia el principio, encontrando la ruta óptima de inversión para cada medida. Esta extensión más que entregar un sistema para rutas futuras óptimas, permitiría de forma clara conocer el gap de optimalidad global, generado al resolver instancias locales.
- Evaluar las condiciones necesarias y suficientes, para que con las medidas utilizadas, utilizando una estrategia *rolling forward*, la solución global sea alcanzada. Este trabajo futuro, va en la misma línea que el anterior, entregando las condiciones para que ambos enfoques entreguen las mismas soluciones. Entendiendo que una resolución vía programación dinámica estocástica, implica un costo en computabilidad mucho mayor a un problema estocástico multitemporal como el abordado en este trabajo.
- Generar y evaluar instancias con *heavy tails* vía, múltiples estados en el proceso *mixed normal*, permitiendo conocer el comportamiento de las diferentes medidas. En particular el efecto de la medida de riesgo entrópica. Este trabajo extiende lo desarrollado con sólo distribuciones asimétricas, permitiendo introducir procesos generadores con *jumps*, con llegadas de un proceso de Poisson, lo que en la literatura es considerado un enfoque validado de modelos con casos extremos.
- Un aspecto del presente trabajo, que por alcance y tiempo no fue explorado, se encuentra en implementar una metodología, más elaborada en la actualización de los procesos generadores de retornos futuros por instrumentos, posterior a que es revelada una nueva realización del precio del instrumento. En esta línea, en Finanzas existen múltiples modelos que pueden ser considerados, que concentran lo desarrollado en más de 15 años, y un Premio Nobel en Economía, como son los modelos de factores que extienden los clásico enfoques de Capital

Asset Pricing Model (CAPM) y Fama-French. Al lector, se le recomienda leer en esta línea, el trabajo de síntesis [32], más que los trabajos originales y su evolución a la fecha.

- Un aspecto pendiente, de investigación aplicada, consiste en aplicar lo desarrollado en otros ámbitos de la ingeniería, como es el procesamiento de datos on-line, sistemas de alerta temprana, sistemas pre-diagnósticos en medicina, sistemas de decisión semi-asistida en la detección de patrones, entre otros.
- Un aspecto a extender del presente trabajo, en específico en el área algoritmo de la aproximación de la región factible, mediante cortes lineales, es el diseño, implementación y evaluación de una estrategia paralela, separando la vecindad óptima.
- El trabajo dejó de manifiesto, que el real potencial de las medidas de riesgo, se encuentra en las medidas asimétricas. En este contexto, se considera interesante extender lo desarrollado, con distribuciones más generales: Por un lado, distribuciones no elípticas, como lo desarrollado con el Entropic Value at Risk en [34]; Y por otro lado, con distribuciones con un sustento más económico, con funciones de copula, como es el trabajo de [27].

Bibliografía

- [1] Kjersti Aas. To log or not to log: The distribution of asset returns. *Norwegian Computing Center, Research Note*, 2004.
- [2] Carlo Acerbi. Risk aversion and coherent risk measures: a spectral representation theorem. *arXiv preprint cond-mat/0107190*, 2001.
- [3] Carlo Acerbi. Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518, 2002.
- [4] Carlo Acerbi and Dirk Tasche. Expected shortfall: a natural coherent alternative to value at risk. *Economic notes*, 31(2):379–388, 2002.
- [5] Alexandre Adam, Mohamed Houkari, and Jean-Paul Laurent. Spectral risk measures and portfolio selection. *Journal of Banking & Finance*, 32(9):1870–1882, 2008.
- [6] Yakov Amihud and Haim Mendelson. Asset pricing and the bid-ask spread. *Journal of financial Economics*, 17(2):223–249, 1986.
- [7] Philippe Artzner. Application of coherent risk measures to capital requirements in insurance. *North American Actuarial Journal*, 3(2):11–25, 1999.
- [8] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.
- [9] Louis Bachelier. *Théorie de la spéculation*. Gauthier-Villars, 1900.
- [10] Egon Balas and Andrew Ho. *Set covering algorithms using cutting planes, heuristics, and subgradient optimization: a computational study*. Springer, 1980.
- [11] Alejandro Balbás, José Garrido, and Silvia Mayoral. Properties of distortion risk measures. *Methodology and Computing in Applied Probability*, 11(3):385–399, 2009.
- [12] G. Barbarosoglu and Y. Arda. A two-stage stochastic programming framework for transportation planning in disaster response. *Journal of the Operational Research Society*, 55(1):43–53, 2004.
- [13] Dimitris Bertsimas and David B Brown. Constructing uncertainty sets for robust linear optimization. *Operations research*, 57(6):1483–1495, 2009.
- [14] Ludwig B Chincarini and Daehwan Kim. *Quantitative equity portfolio management: an active approach to portfolio construction and management*. McGraw-Hill, 2006.
- [15] Gordon L Clark and Roger Urwin. Best-practice pension fund governance. *Journal of Asset Management*, 9(1):2–21, 2008.
- [16] Roberto Cominetti and Alfredo Torrico. Additive consistency of risk measures and its application to risk-averse routing in networks. *arXiv preprint arXiv:1312.4193*, 2013.
- [17] Gregory Connor and Robert A Korajczyk. Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of financial economics*, 15(3):373–394, 1986.
- [18] George B. Dantzig. Programming in a linear structure. Comptroller, USAF Washington D.C., 1948.

- [19] George B Dantzig. *Linear Programming and Extensions*. Princeton landmarks in mathematics and physics, Princeton University Press, Princeton, 1963.
- [20] Dieter Denneberg. Distorted probabilities and insurance premiums. *Methods of Operations Research*, 63(3), 1990.
- [21] Vittas Dimitri. Regulatory controversies of private pension funds. *Economic Development Institute of the World Bank: Pension Systems—From Crisis to Reform*. Washington, DC: World Bank, 1996.
- [22] Peter F Drucker. *The pension fund revolution*. Transaction Publishers, 1976.
- [23] Peter F Drucker. *The unseen revolution: How pension fund socialism came to America*. Elsevier, 2013.
- [24] Edwin J Elton and Martin J Gruber. Portfolio theory when investment relatives are lognormally distributed. *The Journal of Finance*, 29(4):1265–1273, 1974.
- [25] Edwin J Elton and Martin J Gruber. Modern portfolio theory, 1950 to date. *Journal of Banking & Finance*, 21(11):1743–1759, 1997.
- [26] Edwin J Elton and Martin Jay Gruber. *Finance as a dynamic process*. Prentice Hall, 1974.
- [27] Paul Embrechts, Alexander McNeil, and Daniel Straumann. Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond*, pages 176–223, 2002.
- [28] Daniel Espinoza and Eduardo Moreno. A primal-dual aggregation algorithm for minimizing conditional value-at-risk in linear programs. *Computational Optimization and Applications*, 59(3):617–638, 2014.
- [29] Eugene F Fama, Lawrence Fisher, Michael C Jensen, and Richard Roll. The adjustment of stock prices to new information. *International economic review*, 10(1):1–21, 1969.
- [30] Eugene F Fama and Kenneth R French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.
- [31] Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.
- [32] Eugene F Fama and Kenneth R French. The capital asset pricing model: Theory and evidence. *The Journal of Economic Perspectives*, 18(3):25–46, 2004.
- [33] Haw-ren Fang and Dianne P O’leary. Modified cholesky algorithms: a catalog with new approaches. *Mathematical Programming*, 115(2):319–349, 2008.
- [34] Hassan Omid Firouzi and Andrew Luong. Optimal portfolio problem using entropic value at risk: When the underlying distribution is non-elliptical. *arXiv preprint arXiv:1406.7040*, 2014.
- [35] Hans Föllmer and Thomas Knispel. Entropic risk measures: Coherence vs. convexity, model ambiguity and robust large deviations. *Stochastics and Dynamics*, 11(02n03):333–351, 2011.
- [36] Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and stochastics*, 6(4):429–447, 2002.

- [37] Hans Föllmer and Alexander Schied. Convex and coherent risk measures. *preprint, Humboldt University*, 2008.
- [38] Hans Föllmer and Alexander Schied. *Stochastic finance: an introduction in discrete time*. 2011.
- [39] Virginie Gabrel, Cécile Murat, and Aurélie Thiele. Recent advances in robust optimization: An overview. *European Journal of Operational Research*, 235(3):471–483, 2014.
- [40] Richard C Grinold and Ronald N Kahn. *Active portfolio management*, volume 2. McGraw-Hill New York, 2000.
- [41] Gzyl Henryk and Mayoral Silvia. On a relationship between distorted and spectral risk measures. 2006.
- [42] Frank L Hitchcock. The distribution of a product from several sources to numerous localities. *J. Math. phys*, 20(2):224–230, 1941.
- [43] Werner Hürlimann. Distortion risk measures and economic capital. *North American Actuarial Journal*, 8(1):86–95, 2004.
- [44] Koji Inui and Masaaki Kijima. On the significance of expected shortfall as a coherent risk measure. *Journal of banking & finance*, 29(4):853–864, 2005.
- [45] Rabih Jabr, Izudin Dzafic, Bikash C Pal, et al. Robust optimization of storage investment on transmission networks. *Power Systems, IEEE Transactions on*, 30(1):531–539, 2015.
- [46] Leonid V Kantorovich. The mathematical method of production planning and organization. *Management Science*, 6:363–422, 1939.
- [47] Leonid G Khachiyan. Polynomial algorithms in linear programming. *USSR Computational Mathematics and Mathematical Physics*, 20(1):53–72, 1980.
- [48] Anton J Kleywegt, Alexander Shapiro, and Tito Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.
- [49] Zbigniew Kominek. Regulatory induced herding? evidence from polish pension funds. *Economic Change and Restructuring*, 45(1-2):97–119, 2012.
- [50] Alan Kraus and Robert H Litzenberger. Skewness preference and the valuation of risk assets*. *The Journal of Finance*, 31(4):1085–1100, 1976.
- [51] Shigeo Kusuoka. On law invariant coherent risk measures. In *Advances in mathematical economics*, pages 83–95. Springer, 2001.
- [52] Pierre L’Ecuyer. Good parameters and implementations for combined multiple recursive random number generators. *Operations Research*, 47(1):159–164, 1999.
- [53] Jeff Linderoth, Alexander Shapiro, and Stephen Wright. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research*, 142(1):215–241, 2006.
- [54] John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, 47:13–37, 1965.
- [55] Pierre L’Ecuyer. *Random number generation*. Springer, 2012.

- [56] Sébastien Maillard, Thierry Roncalli, and Jerome Teiletche. On the properties of equally-weighted risk contributions portfolios. *Available at SSRN 1271972*, 2008.
- [57] Burton G Malkiel. The efficient market hypothesis and its critics. *Journal of economic perspectives*, pages 59–82, 2003.
- [58] Harry Markowitz. Portfolio selection*. *The journal of finance*, 7(1):77–91, 1952.
- [59] Harry M Markowitz. *Portfolio selection: efficient diversification of investments*, volume 16. Yale university press, 1968.
- [60] Babu M Mehtre, Mohan S Kankanhalli, and Wing Foon Lee. Shape measures for content based image retrieval: a comparison. *Information Processing & Management*, 33(3):319–337, 1997.
- [61] Attilio Meucci. Quant nugget 2: Linear vs. compounded returns—common pitfalls in portfolio management. *GARP Risk Professional Magazine*, pages pp. 49–51, 2010.
- [62] Naomi Miller and Andrzej Ruszczyński. Risk-averse two-stage stochastic linear programming: modeling and decomposition. *Operations research*, 59(1):125–132, 2011.
- [63] Franco Modigliani and Leah Modigliani. Risk-adjusted performance. *The Journal of Portfolio Management*, 23(2):45–54, 1997.
- [64] Jan Mossin. Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society*, pages 768–783, 1966.
- [65] Katharina Müller. The political economy of pension reform in central-eastern europe. *Cheltenham Northampton*, 1999.
- [66] Karthik Natarajan, Dessislava Pachamanova, and Melvyn Sim. Constructing risk measures from uncertainty sets. *Operations Research*, 57(5):1129–1141, 2009.
- [67] Tatsushi Nishi, Yuichiro Hiranaka, and Masahiro Inuiguchi. Lagrangian relaxation with cut generation for hybrid flowshop scheduling problems to minimize the total weighted tardiness. *Computers & Operations Research*, 37(1):189–198, 2010.
- [68] Manfred Padberg and Giovanni Rinaldi. A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems. *SIAM review*, 33(1):60–100, 1991.
- [69] Frau Dipl-Math Irina Penner. *Dynamic convex risk measures: time consistency, prudence, and sustainability*. PhD thesis, Humboldt-Universität zu Berlin, 2007.
- [70] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.
- [71] R Tyrrell Rockafellar, Stan Uryasev, and Michael Zabaranin. Generalized deviations in risk analysis. *Finance and Stochastics*, 10(1):51–74, 2006.
- [72] R Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.
- [73] R Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471, 2002.

- [74] LCG Rogers and Liang Zhang. Understanding asset returns. *Mathematics and Financial Economics*, 2011.
- [75] Richard Roll and Stephen A Ross. An empirical investigation of the arbitrage pricing theory. *Journal of finance*, pages 1073–1103, 1980.
- [76] Brian M Rom and Kathleen W Ferguson. Post-modern portfolio theory comes of age. *The Journal of Investing*, 3(3):11–17, 1994.
- [77] Birgit Rudloff, Jörn Sass, and Ralf Wunderlich. Entropic risk constraints for utility maximization. *et al., Festschrift in celebration of Prof. Dr. Wilfried Grecksch's 60th birthday. Aachen: Shaker Verlag. Berichte aus der Mathematik*, pages 149–180, 2008.
- [78] Tjendera Santoso, Shabbir Ahmed, Marc Goetschalckx, and Alexander Shapiro. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115, 2005.
- [79] Ekaterina N Sereda, Efim M Bronshtein, Svetozar T Rachev, Frank J Fabozzi, Wei Sun, and Stoyan V Stoyanov. Distortion risk measures in portfolio optimization. In *Handbook of Portfolio Construction*, pages 649–673. Springer, 2010.
- [80] Alexander Shapiro. On kusuoka representation of law invariant risk measures. *Mathematics of Operations Research*, 38(1):142–152, 2013.
- [81] Alexander Shapiro. Sample average approximation. In *Encyclopedia of Operations Research and Management Science*, pages 1350–1355. Springer, 2013.
- [82] Alexander Shapiro, Darinka Dentcheva, et al. *Lectures on stochastic programming: modeling and theory*, volume 16. SIAM, 2014.
- [83] Alexander Shapiro and Andrzej Ruszczyński. Lectures on stochastic programming. *Preprint, Georgia Tech*, 2008.
- [84] William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk*. *The journal of finance*, 19(3):425–442, 1964.
- [85] Andrei Shleifer. *Inefficient markets: An introduction to behavioral finance*. Oxford university press, 2000.
- [86] Frank A Sortino and Lee N Price. Performance measurement in a downside risk framework. *the Journal of Investing*, 3(3):59–64, 1994.
- [87] Frank A Sortino and Stephen Satchell. *Managing downside risk in financial markets*. Butterworth-Heinemann, 2001.
- [88] Frank A Sortino and Robert Van Der Meer. Downside risk. *The Journal of Portfolio Management*, 17(4):27–31, 1991.
- [89] B Stein, Pedro Miranda, Rodolfo Risco, et al. Herding in chile: The case of equity trading in the chilean pension fund market. 18(1):pp. 23–44, 2011.
- [90] Edward J Sullivan. A brief history of the capital asset pricing model. *APUBEF Proceedings*, pages 207–210, 2006.

- [91] James Tobin. Liquidity preference as behavior towards risk. *The review of economic studies*, pages 65–86, 1958.
- [92] Michael J Todd. The many facets of linear programming. *Mathematical Programming*, 91(3):417–436, 2002.
- [93] Jack L Treynor. Market value, time, and risk. Technical report, 1961.
- [94] Jack L Treynor. Toward a theory of market value of risky assets. *Unpublished manuscript*, 6, 1961.
- [95] Andreas Tsanakas. Dynamic capital allocation with distortion risk measures. *Insurance: Mathematics and Economics*, 35(2):223–243, 2004.
- [96] Bram Verweij, Shabbir Ahmed, Anton J Kleywegt, George Nemhauser, and Alexander Shapiro. The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications*, 24(2-3):289–333, 2003.
- [97] Shaun S Wang, Virginia R Young, and Harry H Panjer. Axiomatic characterization of insurance prices. *Insurance: Mathematics and economics*, 21(2):173–183, 1997.
- [98] Julia L Wirch and Mary R Hardy. Distortion risk measures. coherence and stochastic dominance. In *International Congress on Insurance: Mathematics and Economics*, pages 15–17, 2001.
- [99] Julia Lynn Wirch and Mary R Hardy. A synthesis of risk measures for capital adequacy. *Insurance: mathematics and Economics*, 25(3):337–347, 1999.

Capítulo 6

Apéndices

6.1. Resumen extendido

En Chile el sistema de pensiones, es un mecanismo de previsión y protección social, que a través del ahorro e inversión, genera el capital para la futura pensión de vejez o potencial pensión de invalidez. Adicionalmente, el sistema ha generado externalidades positivas, por sus regulaciones en composición de inversión, lo que se traduce en un gran volumen de capital que ha dado liquidez al mercado, y en parte, ha facilitado el desarrollo económico.

En línea con lo anterior, la administración de las inversiones a largo plazo, en pos de la rentabilidad de los fondos de los afiliados, y en particular, la administración del riesgo que estas inversiones implican, es un aspecto que abre grandes oportunidades de mejora y perfeccionamiento, migrando de la imitación de estrategias actualmente existente en la industria, a una lógica de administración de riesgo-retorno.

El problema de inversiones es un área extremadamente explorada y desarrollada, que presenta diversos enfoques y ramificaciones, en su sistematización, es considerado como trabajo original de la llamada Teoría Moderna de Portafolios, el trabajo de [58], donde se define un problema de optimización que minimiza la varianza, como medida de dispersión, y por ende *proxy* de riesgo, sujeta a restricciones que definen un portafolio de inversión, y a una rentabilidad esperada. El enfoque anterior, al ser evaluado en el contexto del cumplimiento de axiomas básicos esperables, de la cuantificación del riesgo en condiciones generales, es insuficiente, basta con pensar en eventos aleatorios asimétricos, donde la simple dispersión no dice mucho. Un ejemplo sencillo de lo anterior, es observar una variable que presenta fuertes caídas, pero usualmente tiene un crecimiento suave constante.

Como respuesta a lo anterior, se desarrolló un marco sistemático y riguroso, de condiciones esperable de una llamada *Medida de Riesgo*, en esta área han surgido diversos desarrollos, originando diferentes familias de medidas, cada una de ellas con un origen particular, que se traduce en una definición diferente del *riesgo*.

De lo expuesto anteriormente, resulta natural proponer utilizar esta sistematización, en el

problema de optimización de las inversiones en Chile, en particular, como propuesta de un índice base de referencia de rentabilidad. Este nuevo índice, se complementaría con el actual promedio de los demás agentes, permitiendo evaluar el desempeño, con un enfoque de optimización enfrentando iguales restricciones, información y condiciones de mercado, cuantificando la habilidad de los agentes, y no condiciones erráticas de mercado, o su habilidad de imitación.

Metodológicamente se exploran dos familias de medidas: las *medidas espectrales*, con sus representantes: Valor Esperado, *Conditional Value at Risk* (CVaR) y una combinación convexa de ambas; y la medida Entrópica. CVaR se incorpora como generalización convexa, y por ello optimizable del conocido, y utilizado en la industria *Value at Risk* (VaR), conservando su fácil interpretación.

La resolución del problema de optimización que implementa la Medida de Riesgo Entrópica, constituye el principal aporte metodológico. Para su resolución, al igual que para las demás medidas, se empleó la metodología de resolución de problemas estocásticos *Sample Average Approximation* (SAA), obteniendo de él garantías estadísticas de convergencia estocástica. Adicionalmente, para la medida Entrópica fue necesario desarrollar una metodología de aproximación y ajuste de cortes locales, que permitió resolver en instancias de prueba, con un *gap* estocástico relativo menor al 0.1 %.

La medida de riesgo Entrópica es la única medida que cumple consistencia en la aditividad, es decir el problema original puede ser fragmentado en sub problemas aditivos, con igual solución original. En el contexto de inversiones garantiza la consistencia de resolver una inversión de diez etapas, como una secuencia iterativa de diez problemas de una etapa.

Como resultado del trabajo, se desarrolló una metodología aplicable y replicable para la generación de una referencia de inversión (*benchmark*). Metodológicamente se desarrolló un algoritmo vía generación y ajuste de cortes en la aproximación de la vecindad de la región óptima, para la adecuada resolución de la medida de riesgo Entrópica.

Con los modelos desarrollados, se realizaron diversos experimentos computacionales, cada uno de ellos respondiendo diferentes preguntas e hipótesis. En particular se evaluó la convergencia y tiempos de resolución de los diferentes modelos, conceptos como la correlación, distribuciones con colas gordas simétricas y asimétricas, y la comparativa y pertinencia de métodos más sofisticados frente al clásico enfoque de [58]. Adicionalmente, con datos diarios históricos del fondo A, desde el año 2003 al 2013, se resuelve a modo de referencia, el problema de optimización.

Se concluyó que ante instancias simétricas, las soluciones generadas por las diferentes medidas coinciden para algún nivel de aversión. En particular, la noción clásica de dispersión es suficiente. No obstante, ante una instancia asimétrica las diferentes nociones de riesgo se diferencian. Por otro lado, el efecto de la correlación en la instancia normal multivariada evaluada, sólo generó un desplazamiento de los parámetros de aversión de las medidas, efecto semejante a lo ocurrido en distribuciones simétricas con colas gordas¹.

La metodología SAA mostró ser un marco de desarrollo y resolución, con garantías estadísticas de convergencia, muy sólida y general en su aplicación, con la única desventaja directa de una

¹Es importante hacer notar, que la definición de *cola gorda* empleada hace referencia al concepto de fat-tail, donde se hace la diferencia con una distribución normal. Y por el contrario, las *heavy-tail* usa de referencia un decaimiento exponencial de las colas.

necesidad de computo mayor. Por otro lado, el enfoque adoptado de aproximar las restricciones exponenciales como una secuencia de cortes lineales generados y ajustados en la misma ejecución, dió como resultado una solución efectiva y extensible a otro tipo de problemas no lineales, convexos, y de fácil separación, como fue la suma de exponenciales.

Finalmente, a la conocida interpretación de la medida de riesgo Entrópica, como una distorsión de la medida de probabilidad para los casos más desfavorables, se le agrega la evidente interpretación, como observación de una expansión de Taylor, y de su comportamiento en los diferentes experimentos, como una ponderación de los diferentes momentos de la distribución por las potencias del término de aversión. Ello resulta coherente con la conclusión antes expuesta de distribuciones simétricas normales, donde los primeros dos momentos de la distribución son suficientes, para la descripción total de la variable.

6.2. Monotonía de la medida de riesgo entrópica a la adversión al riesgo

Teorema 2. *La medida de riesgo entrópica $\rho^{ent}(X, \theta)$ es monotonamente creciente a la adversión al riesgo $\theta \in [0, \infty)$*

Demostración. Dado $\theta < \theta_1$ y $X \in \mathcal{L}^2$, tomamos $\theta_1 = k\theta$ con $k > 1$, con esto es directo observar que las siguientes inecuaciones son equivalentes:

$$\begin{aligned}\rho^{ent}(X, \theta) &\leq \rho^{ent}(X, \theta_1) \\ \frac{1}{\theta} \log(\mathbb{E}(e^{-\theta X})) &\leq \frac{1}{k\theta} \log(\mathbb{E}(e^{-k\theta X})) \\ \left[\mathbb{E}(e^{-\theta X})\right]^k &\leq \mathbb{E}((e^{-\theta X})^k) \\ \left[\mathbb{E}(Z)\right]^k &\leq \mathbb{E}(Z^k)\end{aligned}$$

Lo que resulta en una desigualdad de Jensen $g(\mathbb{E}(Z)) \leq \mathbb{E}(g(Z))$, con $g(Z) = Z^k$ que es convexa dado que $Z \in \mathbb{R}^+$ como resultado del cambio de variable $Z = e^{-\theta X}$ y $k > 1$ por construcción. \square

6.3. Consistencia intertemporal de las medidas de riesgo

Un aspecto relacionado a las medidas de riesgo, necesario de introducir dadas sus implicancias en el actual trabajo, hace relacion a la consistencia intertemporal.

Antes de definir la consistencia, es necesario ampliar el marco teorico, agregando el concepto de medidas de riesgo dinamicas, para permitir introducir la noción de consistencia intertemporal, propiedad fundamental cuando se pretende entregar una medida de riesgo que sea consistente en problemas multiperiodos, exigiendo un orden en la evolución de la medida de riesgo, a medida que evoluciona la sigma algebra de la filtración asociada hasta el instante t . Esta propiedad es satisfecha por la medida de riesgo entrópica, una de las medidas de riesgo considerada en este trabajo, donde se incorpora y resuelve para un problema de portafolio con restricciones.

Dado el espacio de probabilidad $(\Omega, \mathcal{F}, (\mathcal{F})_{t=0, \dots, T-1}, P)$ con filtraciones sobre $t \in \{0, \dots, T-1\}$, talque $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F} = \mathcal{F}_T$ y $\mathcal{F}_t = \cup_{\{t' < t\}} \mathcal{F}_{t'}$. Se formula la pregunta de como evoluciona la medida de riesgo ρ condicional a la evolución en la filtración, dando origen a las medidas de riesgo dinamicas.

Las medidas de riesgo, las interpretaremos como una sucesión de medidas, indexadas por T , $(\rho_t)_{t \in \{0, \dots, T-1\}}$ talque cada medida $\rho(X)_t$ es la medida de riesgo ρ condicional a la sigma algebra hasta t , ie $\rho(X|\mathcal{F}_t)$.

El concepto de consistencia intertemporal, formaliza la intuición de riesgo en un proceso multiperiodo, de que si un activo Z_1 es más riesgoso en un instante t a un activo Z_2 , entonces en todo momento anterior a t , el activo Z_1 debio ser más riesgoso que Z_2 , y en particular, si no se agrega un evento con probabilidad no nula, donde Z_2 es más riesgoso que Z_1 , entonces en $t + 1$ tambien Z_1 será el más riesgoso. Formalmente, diremos que una medida de riesgo dinamica es consistente intertemporalmente si satisface que:

1. $\forall Z_1, Z_2 \in \mathcal{L}^2(\mathcal{F}_T)$ y $\forall t \in \{0, 1, \dots, T - 2\}$ se tiene que $\rho_{t+1}(Z_1) \geq \rho_{t+1}(Z_2)$ implica que $\rho_t(Z_1) \geq \rho_t(Z_2)$.

6.4. Intento de formulación espectral de la medida entrópica

6.4.1. Formulación LP de las medidas de riesgo entrópicas

$$\begin{aligned} & \min_x \rho(z) \\ \text{s.t.} \quad & z = Ax \\ & x \in X \end{aligned} \tag{6.1}$$

que es equivalente con un N suficientemente grande a:

$$\begin{aligned} & \min_x \rho_g^N(z) \\ \text{s.t.} \quad & z = Ax \\ & x \in X \end{aligned} \tag{6.2}$$

$$\text{con } \rho_g^N(z) = \sum_{i \in [N]} g_i z_{(i)}$$

Lo que no es una formulación lineal, por la operación de orden implícita en los $z_{(i)}$. Buscando replicar lo realizado por Rockafellar y Uryasev en 1999 con el CVAR, debemos buscar $\{t_i\}_{i=1}^N$ talque:

$$t_i + \frac{N}{i} \sum_{j \in [N]} \frac{1}{N} (z_j - t_i)_+ = t_i + \frac{1}{i} \sum_{j \in [N]} (z_j - t_j)_+ = C_i(z, t)$$

Con esto la formulación anterior, es equivalente a:

$$\begin{aligned} & \min_{x,t} \sum_{i \in [N]} C_{N-i+1}(z, t)(g_i - g_{i-1}) \\ \text{s.t.} \quad & z = Ax \\ & x \in X \end{aligned} \tag{6.3}$$

que explícitamente la podemos escribir como el siguiente LP:

$$\min_{x,t} \sum_{i \in [N]} (t_{N-i+1} + \frac{1}{N-i+1} \sum_{j \in [N]} w_j)(g_i - g_{i-1}) \quad (6.4)$$

$$\begin{aligned} s.t. \quad z &= Ax \\ w_j &= z_j - t_j \quad \forall j \in [N] \\ x &\in X \\ \vec{w} &\geq \vec{0} \end{aligned} \quad (6.5)$$

Donde para el caso de las medidas de riesgo entrópicas $g_i = \frac{\beta}{e^\beta - 1} e^{\beta \frac{i}{N}}$