Universidad Nacional Autónoma de México

Facultad de Ciencias

UN MODELO DE OPTIMIZACIÓN DE INVERSIÓN EN UN MERCADO DE VALORES

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

Matemático

PRESENTA:

Orlando Alan Betancourt Ramírez

TUTOR

Dr. Luis Antonio Rincón Solís

Ciudad Universitaria, 2018

Índice general

Capítulo 1

Introducción

Uno de los problemas más estudiados dentro de las finanzas y la teoría económica es encontrar la manera en que después de cierto proceso, una persona finalice con más dinero que con el que inició. El presente trabajo pretende dar una manera de abordar un caso particular del problema de maximización: la manera de maximizar el rendimiento del capital que un inversionista ingresa en un mercado de valores. Por supuesto, y como toda teoría matemática introductoria, hacemos supuestos que, aunque fuertes, permiten un análisis más digerible. No obstante, no perdemos de vista la simplicidad del modelo y por ende, señalamos líneas para profundizar en el estudio del problema abordado.

El presente documento es un trabajo de caracter expositivo. Es una expansión del capítulo titulado «*Information theory and portfolio theory*» del libro [**1**] *Elements of information theory* escrito por Thomas Cover y Joy Thomas, cuyo objetivo es el de presentar el tema en cuestión desde una perspectiva introductoria. Es por esta razón que el presente trabajo tiene por intención el de presentar cómo los conceptos de la teoría de la información tienen una aplicación real en el tema de la teoría de portafolios desde una perspectiva introductoria y básica.

Adicionalmente, el trabajo se ha complementado con ejemplos incluidos a lo largo del desarrollo del documento, con el fin de ilustrar los conceptos que se presentan y también, sin haber tenido la intención al momento de trabajarlos, de indicar posibles variantes, generalizaciones, o cambios de perspectiva para abordar el problema que nos ocupa.

La estructura del trabajo es la siguiente:

Primeramente, se presentan conceptos y definiciones primordiales para el entendimiento del modelo, como lo son las definiciones de esperanza de una variable aleatoria, entropía, entropía relativa, información mutua, conjunto convexo, entre otros. La mayoría de los resultados presentados en esta sección no incluyen demostración, puesto que ese no es el objetivo del trabajo. Sin embargo, se han incluido ejemplos para facilitar la asimilación de los conceptos. Vale la pena indicar que dichos conceptos son en su mayoría pertenecientes a la teoría de la probabilidad. Otros tantos lo son a la teoría de la información y un par de ellos al análisis matemático.

A continuación se presentan las definiciones relativas al problema central del trabajo. Se define lo que es un portafolio log-óptimo, un mercado de valores, el precio relativo de una acción, el capital relativo y la tasa de crecimiento. Se discuten propiedades de algunos de los conceptos antes mencionados, relativos a la linealidad, convexidad y convergencia. Es en este apartado donde se presenta el problema central de una manera formal y donde se indican los supuestos para el análisis. Se presentan además los primeros ejemplos de portafolios óptimos.

Después, se presentan las condiciones de Kuhn-Tucker relativas a un portafolio log-óptimo. Estas condiciones son una caracterización de un portafolio log-óptimo. Además, se muestra cómo estas condiciones permiten obtener conclusiones respecto del capital relativo de un inversionista en relación con el capital que obtendría de usar un portafolio que no sea óptimo. También se incluye un ejemplo que utiliza los conceptos aprendidos en este apartado.

En la siguiente sección, se demuestra que con probabilidad 1, un inversionista que use un portafolio log-óptimo «al menos no le irá peor» que cualquier otro inversionista que use cualquier otro portafolio, o dicho de otra forma, que no hay otro portafolio mejor que el óptimo. Hasta este punto, el análisis se había concentrado en inversiones a un día. Es en este apartado donde se extiende el análisis para el caso en el que una persona invierta más de un día, asumiendo que los mercados en los que se invierte son independientes e idénticamente distribuidos.

Posteriormente, se presenta la situación en que un participante del mercado invierte sin tomar en cuenta información que los demás participantes sí conocen. Concretamente, se dan cotas para la diferencia entre la utilidad que dicho inversionista obtendría contra aquella que obtenga el inversionista bien informado. Matemáticamente, esto se alcanza a través de analizar la relación entre la información mutua y la tasa de crecimiento.

Adicionalmente, se presenta un primer análisis para el caso de debilitar el supuesto de que la distribución del mercado un día es independiente de la distribución del día anterior. Entran en discusión en esta sección los conceptos de estacionariedad y convergencia de Césaro. Se generalizan algunos conceptos presentados en la sección introductoria.

Como último tema, se demuestra que con las condiciones presentadas hasta el momento, no existe una cota para la probabilidad de que a un inversionista sin el portafolio óptimo le vaya mejor que a un inversionista que sí usa un portafolio óptimo. Por ello, se presentan condiciones adicionales para que se pueda encontrar una cota informativa para la probabilidad mencionada.

Finalmente, se presenta un pequeño compendio de posibles generalizaciones o extensiones a los modelos presentados en el documento. Además, esta sección pretende invitar a la reflexión del lector en cuanto a la complejidad que se puede alcanzar utilizando supuestos básicos y por ende, la complejidad que se puede alcanzar al debilitar hipótesis.

Capítulo 2

Preliminares

Recordemos algunos resultados y definiciones de probabilidad y de teoría de la información.

Comencemos con el concepto de esperanza, coloquialmente conocida como promedio. Efectivamente, la esperanza es un promedio de los posibles valores que puede tomar una variable aleatoria, ponderados por la probabilidad de que dicha variable tome esos valores.

Definición 2.1. La esperanza de una variable aleatoria X con función de distribución $F(x)$ se define como:

$$\mathbb{E}(X) = \int_{\mathbb{R}} x dF(x),$$

suponiendo que dicha integral es absolutamente convergente.

Si la variable es continua y tiene función de densidad $f(x)$, entonces la esperanza se puede escribir como:

$$\mathbb{E}(X) = \int_{\mathbb{R}} x f(x) dx.$$

Si además la variable aleatoria toma valores no negativos, la esperanza se puede expresar como

$$\mathbb{E}(X) = \int_0^\infty (1 - F(x)) dx.$$

Por ejemplo, si X tiene distribución uniforme $U(a, b)$, entonces

$$\begin{aligned}\mathbb{E}(X) &= \int_a^b \frac{x}{b-a} dx \\ &= \frac{b+a}{2}.\end{aligned}$$

Un resultado que hace uso del concepto de esperanza y que nos será útil más adelante es el de la desigualdad de Markov.

Proposición 2.2. *Si tenemos una variable aleatoria no negativa X y un número $t > 0$, entonces:*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

Puede consultar la demostración en [**1**].

La importancia de la desigualdad de Markov es que nos permite encontrar una cota para la probabilidad del evento indicado cuando sólo la esperanza de la distribución de probabilidad es conocida.

La desigualdad de Markov es de ayuda para probar la versión débil del siguiente resultado, que es uno de los pilares de la teoría de la probabilidad: la ley fuerte de los grandes números.

La ley fuerte de los grandes números es probablemente el resultado más conocido en la teoría de la probabilidad. Esta ley establece que el promedio de una secuencia independiente de variables aleatorias con una distribución común convergerá, con probabilidad 1, a la esperanza de la distribución.

TEOREMA 2.3. *Sean $X_1, X_2, ...$ variables aleatorias independientes e idénticamente distribuidas con $E(X_i) = \mu$, entonces*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu,$$

con probabilidad 1.

Puede consultar la demostración de la ley de los grandes números en [**2**].

Ahora veamos la desigualdad de Jensen, que relaciona la concavidad o convexidad de una función con el concepto de esperanza de una variable aleatoria. Para ello, veamos primero la definición de función convexa y estrictamente convexa.

Gráficamente, una función convexa es aquella en la que el segmento de la curva entre dos puntos de la gráfica queda por debajo de la recta que une dichos puntos. Por el contrario, una función cóncava es aquella en la que la curva entre dos puntos de la gráfica queda por arriba de la recta que une dichos puntos.

DEFINICIÓN 2.4. Se dice que una función $f(x)$ es convexa sobre el intervalo (a, b) si para cualesquiera $x_1, x_2 \in (a, b)$ y para $\lambda \in [0, 1]$ se tiene que:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2).$$

Si la igualdad se da únicamente cuando $\lambda = 0$ o $\lambda = 1$, se dice que la función es estrictamente convexa. Asimismo, una función $f(x)$ es cóncava si $-f(x)$ es convexa.

Ahora podemos ver la desigualdad de Jensen.

TEOREMA 2.5. *Si $g(x)$ es una función convexa en (a, b) y X es una variable aleatoria con valores en (a, b), entonces:*

$$\mathbb{E}(g(X)) \geq g(\mathbb{E}(X)).$$

Más aún, si $g(x)$ es estrictamente convexa la igualdad en la desigualdad anterior implica que $X = \mathbb{E}(X)$ con probabilidad 1 (es decir, X es constante).

La demostración de este hecho está en [**1**].

Un ejemplo sencillo de la utilidad de la desigualdad de Jensen en el contexto de las inversiones es el siguiente, que puede consultarse en [**2**]. Un inversionista se enfrenta con las siguientes alternativas: invertir todo su dinero en un activo riesgoso que le ofrece un rendimiento X con media m, o bien, invertir su dinero en un activo libre de riesgo que le dará un rendimiento m con probabilidad 1. Supongamos que su decisión estará basada en maximizar $u(R)$, donde R es el rendimiento y u es su función de utilidad. Por la desigualdad de Jensen, y dado que las funciones de utilidad son cóncavas, tenemos que $\mathbb{E}(u(X)) \leq u(m)$, por lo que la alternativa libre de riesgo es preferible.

Otro teorema que se usará en el desarrollo de este trabajo es el famoso lema de Borel-Cantelli.

TEOREMA 2.6. *Sea $E_1, E_2, ...$ una sucesión de eventos en un espacio de probabilidad.*
a) Si la sucesión cumple que

$$\sum_{i=1}^{\infty} \mathbb{P}(E_i) < \infty,$$

entonces:

$$\mathbb{P}\left(\limsup_{n\to\infty} E_n\right) = 0,$$

o equivalentemente:

$$\mathbb{P}\left(\omega \in E_n, \textit{para una infinidad de valores de } n\right) = 0.$$

b) Si la sucesión está compuesta por eventos independientes y cumple que

$$\sum_{i=1}^{\infty} \mathbb{P}(E_i) = \infty,$$

entonces:

$$\mathbb{P}\left(\limsup_{n\to\infty} E_n\right) = 1.$$

Un ejemplo bastante ilustrativo del lema de Borel-Cantelli es el siguiente:

Sea E el evento «Ganar la lotería». Supongamos que $\mathbb{P}(E) > 0$.

Definamos A_k como el evento «Ocurre E en el ensayo k», con $k = 1, 2, 3, ...$

Los eventos A_k tienen las siguientes propiedades:

- $\mathbb{P}(A_k) = \mathbb{P}(E) > 0$.
- $\displaystyle\sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty$.
- $A_1, A_2, A_3...$ son independientes.

Por el lema, tenemos que $\mathbb{P}\left(\limsup_{k\to\infty} A_k\right) = 1$.

Pero recordemos que $\limsup_{k\to\infty} A_k$ es el evento «Ocurren una infinidad de eventos $A_1, A_2, ...$». Esto significa que si nos mantenemos jugando a la lotería de forma permanente, con probabilidad 1, eventualmente ganaremos el primer premio. Y si somos un poco más pacientes, con probabilidad 1, ganaremos el primer premio tantas veces como deseemos.

Ahora introduciremos los conceptos fundamentales de la teoría de la información.

Definición 2.7. Sea X una variable aleatoria discreta con función de probabilidad $p(x)$. La entropía de la variable aleatoria se define como:

$$H(X) = -\sum_x p(x)\log(p(x)) = -\mathbb{E}\left(\log(p(X))\right).$$

Donde log simboliza el logaritmo en base 2 de un número real y x es tal que $p(x) > 0$.

Por ejemplo, la entropía de una variable aleatoria que se distribuye Bernoulli $(\frac{1}{2})$ es:

$$\begin{aligned} H(X) &= -\frac{1}{2}\log(\frac{1}{2}) - \frac{1}{2}\log(\frac{1}{2}) \\ &= 1. \end{aligned}$$

La entropía se suele interpretar como el promedio de la incertidumbre que tiene una variable aleatoria, en el entendido que la incertidumbre representa la dispersión de los valores que la variable aleatoria podría tener.

Mencionemos algunas características de la entropía:

- La entropía es siempre positiva e incluso puede tomar el valor infinito.
- Se puede definir la entropía usando alguna base b en el logaritmo. En este caso, si queremos cambiar la base a una base a, debemos usar la siguiente igualdad:

$$H_b(X) = (\log_b(a))H_a(X).$$

Extendemos ahora la definición de entropía para un par de variables aleatorias.

Definición 2.8. La entropía conjunta $H(X, Y)$ de un par de variables aleatorias discretas (X, Y) con función de probabilidad $p(x, y)$ se define como

$$H(X,Y) = -\sum_x \sum_y p(x,y)\log(p(x,y)),$$

donde x, y son tales que $p(x, y) > 0$. La entropía conjunta también puede expresarse en términos de la esperanza, de la siguiente manera:

$$H(X,Y) = -\mathbb{E}(\log(p(X,Y))).$$

Como en el caso de una varibale aleatoria, esta entropía puede interpretarse como el promedio de la incertidumbre que tiene el vector aleatorio.

También definimos la entropía condicional de una variable aleatoria Y dada la variable aleatoria X. Se define como la esperanza de las densidades condicionales, ponderadas por la densidad de la variable que condiciona.

Definición 2.9. Si (X, Y) tiene función de probabilidad $p(x, y)$, la entropía condicional se define como:

$$H(Y|X) = \sum_x p(x)H(Y|X = x),$$

donde $H(Y|X = x)$ es la entropía de la variable aleatoria $Y|X = x$.

De la definición de $H(Y|X)$, podemos deducir que:

$$\begin{aligned} H(Y|X) &= \sum_x p(x)H(Y|X=x) \\ &= -\sum_x p(x)\sum_y p(y|x)\log(p(y|x)) \\ &= -\sum_x\sum_y p(x,y)\log(p(y|x)) \\ &= -\mathbb{E}(\log(p(Y|X))). \end{aligned}$$

La motivación para las definiciones de entropía condicional y conjunta surge por el hecho de que la entropía de un par de variables aleatorias es la entropía de una de ellas más la entropía condicional de la otra, como se muestra en el siguiente teorema, el cual se conoce como regla de la cadena para la entropía. La demostración de este teorema puede encontrarse en [**1**].

TEOREMA 2.10.

$$H(X,Y) = H(X) + H(Y|X).$$

Note que las definiciones antes mencionadas pueden extenderse sin mayor dificultad para el caso en el que trabajamos con vectores aleatorios de cualquier dimensión finita. No obstante, la regla de la cadena obtiene entonces la siguiente forma:

TEOREMA 2.11. *Suponga que $X_1, X_2, ..., X_n$ son cualesquiera variables aleatorias discretas y que cada una tiene función de probabilidad $p(x)$. Entonces*

$$H(X_1, X_2, ..., X_n) = \sum_{i=1}^{n} H(X_i|X_{i-1}, ..., X_1).$$

Cabe destacar que más adelante se encontrará otra «regla de la cadena», que tendrá similitudes y simetrías con la que acabamos de definir.

Ahora veamos la definición de entropía para variables aleatorias continuas, conocida como entropía diferencial.

DEFINICIÓN 2.12. La entropía diferencial $h(X)$ de una variable aleatoria continua con función de densidad $f(x)$ se define como:

$$h(X) = -\int_S f(x)\log f(x)dx,$$

si la integral existe y en donde S es el soporte de la variable aleatoria, es decir, el conjunto donde $f(x) > 0$.

Por ejemplo, si X se distribuye uniforme en el intervalo $(0, a)$ entonces su entropía diferencial es:

$$h(X) = -\int_0^a \frac{1}{a}\log\frac{1}{a}dx = \log a.$$

O si tenemos una variable aleatoria normal con función densidad $\eta(x)$, entonces la entropía diferencial en base e, es:

$$\begin{aligned}
h(\eta) &= -\int_{-\infty}^{\infty} \eta(x)\ln(\eta(x))dx \\
&= -\int_{-\infty}^{\infty} \eta(x)\left[-\frac{x^2}{2\sigma^2} - \ln(\sqrt{2\pi\sigma^2})\right]dx \\
&= \frac{\mathbb{E}(X^2)}{2\sigma^2} + \frac{1}{2}\ln(2\pi\sigma^2) \\
&= \frac{1}{2} + \frac{1}{2}\ln(2\pi\sigma^2) \\
&= \frac{1}{2}\ln(e) + \frac{1}{2}\ln(2\pi\sigma^2) \\
&= \frac{1}{2}\ln(2\pi e\sigma^2).
\end{aligned}$$

A diferencia de la entropía para variables aleatorias discretas, la entropía diferencial puede no ser positiva, como es evidente en el primer ejemplo, cuando $a < 1$. Asimismo, note que la interpretación que se le daba a la entropía para variables aleatorias discretas no tiene sentido en este caso, precisamente gracias a que la entropía diferencial puede ser negativa. Es decir, no puede dársele la interpretación de «promedio de incertidumbre de una variable aleatoria» a un número negativo. No hay un concenso acuerdo sobre qué representaría esta cantidad en el caso continuo

Prosigamos con la definición de la entropía relativa entre dos densidades de probabilidad discretas.

La entropía relativa es una medida de la distancia entre dos distribuciones de probabilidad. Es decir, es una medida de la ineficiencia de asumir que la distribución de una variable aleatoria es q cuando en realidad la distribución es p.

DEFINICIÓN 2.13. La entropía relativa, o «distancia» de Kullback-Leibler, entre dos densidades de probabilidad discretas $p(x)$ y $q(x)$ con soportes F, G respectivamente, se define como:

$$D(p||q) = \sum_{x\in F\cup G} p(x)\log\frac{p(x)}{q(x)} = \mathbb{E}\left(\log\left(\frac{p(X)}{q(X)}\right)\right).$$

Cabe mencionar que en la definición anterior, se usa la convención de que $0 \log \frac{0}{0} = 0$, $0 \log \frac{0}{q} = 0$ y de que $p \log \frac{p}{0} = \infty$, cuando $p > 0$.

Una de las propiedades de la entropía relativa es que siempre es no negativa y es cero si y sólo si $p = q$. No obstante, no es cierto que sea una distancia, de acuerdo a la definición formal de distancia, debido a que no es simétrica ni cumple con la desigualdad del triángulo.

Como ejemplo, consideremos una variable aleatoria X con valores en $\{0, 1\}$ y distribución $p(x)$ uniforme. Sea $q(x)$ otra distribución sobre $\{0, 1\}$ dada por $q(0) = 1/3$, $q(1) = 2/3$. Entonces:

$$\begin{aligned} D(p||q) &= \sum_x p(x) \log \frac{p(x)}{q(x)} \\ &= p(0) \log \frac{p(0)}{q(0)} + p(1) \log \frac{p(1)}{q(1)} \\ &= \frac{1}{2} \log \frac{1/2}{1/3} + \frac{1}{2} \log \frac{1/2}{2/3} \\ &= \log 3 - \frac{3}{2}. \end{aligned}$$

Pasemos ahora a la versión continua de la entropía relativa.

Definición 2.14. La entropía relativa, o «distancia» de Kullback-Leibler, $D(f||g)$ entre dos funciones de densidad f y g se define como:

$$D(f||g) = \int_{\mathbb{R}} f(x) \log \frac{f(x)}{g(x)} dx.$$

Es de notarse que la entropía relativa es finita cuando el soporte de f está contenido en el soporte de g. En este caso también se toman las mismas convenciones indicadas en el caso discreto.

A continuación se define el concepto de información mutua para el caso discreto y continuo. La relación entre entropía diferencial e información mutua es estrecha, pues la información mutua es un caso particular de la entropía diferencial.

Definición 2.15. Considérense dos variables aleatorias discretas X, Y con función de probabilidad conjunta $p(x, y)$ y distribuciones marginales $p(x)$ y $p(y)$. La información mutua $I(X; Y)$ es la entropía relativa entre la función de probabilidad conjunta y la distribución producto

$p(x)p(y)$:

$$\begin{aligned} I(X;Y) &= \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \\ &= D(p(x,y)||p(x)p(y)). \end{aligned}$$

Y para el caso continuo:

DEFINICIÓN 2.16. La información mutua $I(X;Y)$ entre dos variables aleatorias continuas con función de densidad conjunta $f(x,y)$ es:

$$\begin{aligned} I(X;Y) &= \int f(x,y) \log \frac{f(x,y)}{f(x)f(y)} dxdy \\ &= D(f(x,y)||f(x)f(y)). \end{aligned}$$

Como se mencionó anteriormente, la información mutua es un caso particular de la entropía relativa. La información mutua es la entropía relativa entre la distribución conjunta de dos variables aleatorias y la distribución que resulta del producto de sus densidades.

A continuación veamos cómo se relaciona la información mutua con la entropía.

PROPOSICIÓN 2.17. *Si X y Y son dos variables aleatorias, ya sea ambas discretas o ambas continuas, se cumple que:*

- $I(X;X) = H(X)$.
- $I(X;Y) = H(X) - H(X|Y)$.
- $I(X;Y) = H(Y) - H(Y|X)$.
- $I(X;Y) = I(Y;X)$.
- $I(X;Y) = H(X) + H(Y) - H(X,Y)$.

Puede consultarse la demostración de estos resultados en [**1**].

Ahora veamos cómo se extiende el concepto de entropía para cuando estamos trabajando con procesos estocásticos.

DEFINICIÓN 2.18. Sea $\mathcal{X} = (X_n)_{n\in\mathbb{N}}$ un proceso estocástico. La entropía de $\mathcal{X}$ se define como:

$$H(\mathcal{X}) = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, ..., X_n),$$

cuando el límite existe.

Un caso particular en el que este límite existe, es cuando el proceso en cuestión es un proceso estacionario, de acuerdo con la siguiente definición.

DEFINICIÓN 2.19. Se dice que un proceso estocástico es estacionario si la distribución conjunta de cualquier subconjunto de variables aleatorias es invariante con respecto a una traslación en el índice del proceso. Es decir, para cualquier $l \geq 1$ entero,

$$\mathbb{P}(X_1 = x_1, ..., X_n = x_n) = \mathbb{P}(X_{1+l} = x_1, ..., X_{n+l} = x_n).$$

Podemos ahora enunciar el siguiente teorema:

Teorema 2.20. *Para un proceso estocástico estacionario, se tiene que $H(\mathcal{X})$ existe y cumple la siguiente igualdad:*

$$H(\mathcal{X}) = \lim_{n\to\infty} H(X_n|X_{n-1}, ..., X_1).$$

Veamos ahora las definiciones de martingala, submartingala y supermartingala.

Definición 2.21. Se dice que un proceso aleatorio $(X_n)_{n\in\mathbb{N}}$ a tiempo discreto es una martingala respecto de una filtración $(\mathcal{F}_n)_{n\geq 1}$ si cumple las siguientes tres condiciones:

- $\mathbb{E}(X_n) < \infty$.
- X_n es $\mathcal{F}_n$-medible.
- Para cualesquiera $n \leq m$, sucede que con probabilidad 1,

$$\mathbb{E}(X_m|\mathcal{F}_n) = X_n.$$

Cuando en lugar de tener la igualdad en la última ecuación, se tiene que $\mathbb{E}(X_m|\mathcal{F}_n) \geq X_n$, entonces el proceso es una submartingala, y si $\mathbb{E}(X_m|\mathcal{F}_n) \leq X_n$, entonces es una supermartingala.

Las martingalas tienen una interpretación un tanto sencilla en términos de juegos justos. Si X_n denota el capital de un jugador al tiempo n, entonces la igualdad de la definición establece que la fortuna promedio al tiempo m, dado que se conoce la historia del juego al tiempo n, es su capital al tiempo n, es decir, el juego es justo pues en promedio el jugador no pierde ni gana. En este sentido, la desigualdad $\mathbb{E}(X_m|\mathcal{F}_n) \geq X_n$, correspondiente a la definición de submartingala equivale a un juego favorable al jugador. La desigualdad contraria, el caso de submartingala, corresponde a un juego desfavorable para el jugador. Un estudio más detallado de las martingalas se puede encontrar en [**4**].

Un teorema importante concerniente a la convergencia casi segura de martingalas es el siguiente:

Teorema 2.22. *Sea $(X_n)_{n\in\mathbb{N}}$ una submartingala que cumple que $\sup_n \mathbb{E}(|X_n|) < \infty$. Entonces existe una variable aleatoria integrable X tal que:*

$$\lim_{n\to\infty} X_n = X.$$

Como toda martingala es una submartingala, y toda supermartingala se convierte en submartingala a través de un cambio de signo, se tiene que el teorema anterior es válido en cualquiera de los tres casos.

Es decir que toda martingala, submartingala o supermartingala acotada en la forma en la que indica el enunciado del teorema es convergente casi seguramente, y su límite es una variable aleatoria integrable.

Repasemos ahora la definición de conjunto convexo.

DEFINICIÓN 2.23. Un conjunto $A \subseteq \mathbb{R}^n$ se llama convexo si el segmento de recta que une cualesquiera dos puntos en A se queda contenido en A.

Capítulo 3

Modelo de optimización

1. Algunas definiciones

Comencemos esta sección definiendo un mercado de valores.

DEFINICIÓN 3.1. Un mercado de valores se representa como un vector X con entradas $(X_1, X_2, ..., X_m)$.

Supondremos que $(X_1, X_2, ..., X_m)$ es un vector aleatorio y denotaremos por $F(x_1, x_2, ..., x_m)$ su función de distribución. Se considerará que $X_i \geq 0$ y que m es el número de valores en el mercado.

La variable aleatoria X_i, con $i = 1, 2, ..., m$ se llama rendimiento del valor o acción. En la práctica, este rendimiento suele modelarse mediante la razón entre el precio de la acción al final del día y el precio al final del día anterior. No obstante, Cover y Thomas [**1**] sugieren que se tome el precio al final del día entre el precio al inicio del día. En cualquier caso, representa el incremento o pérdida porcentual del capital invertido en la acción en un periodo de un día. Generalmente, el rendimiento diario de una acción suele tener un valor cercano a 1. No obstante, en periodos de estrés financieros o de alta volatilidad, dicho rendimiento puede variar considerablemente. En este caso, se decidió utilizar rendimientos porque representan el cambio en el precio de las acciones independientemente de si la acción cuesta diez o un millón de unidades monetarias.

Note además que no estamos haciendo ninguna suposición sobre la dependencia entre las variables X_i, ni sobre su distribución conjunta. Uno de los atributos del modelo es precisamente que no se ajusta a una distribución particular. Sin embargo, el simple hecho de suponer que el mercado sigue una distribución F es un tema que se puede prestar a discusión. Puede argumentarse que un mercado de valores no sigue una distribución fija y estable, puesto que los precios de las acciones están determinadas por una infinidad de factores externos y cambiantes. Estos factores incluyen desde el humor del *broker* al momento de negociar la operación, la influencia de la situación externa sobre los inversionistas inexpertos, hasta una publicación en redes sociales de una declaración u opinión del presidente de los Estados Unidos. Para efectos de la simplicidad, consideraremos que esta distribución F existe, puede ser estable durante un periodo y de hecho puede ser conocida por uno o varios inversionistas.

Ahora definamos lo que es un portafolio.

Definición 3.2. Un portafolio es un vector $b = (b_1, b_2, ..., b_m)$, con $b_i \geq 0$ y con $b_1 + \cdots + b_n = 1$. Cada b_i representa la fracción del capital de un inversionista en el valor i.

El portafolio es la distribución porcentual del capital de un inversionista en el mercado de valores. No necesariamente es todo el capital del inversionista, sino sólo el que destina al mercado de valores.

Finalmente, definamos el capital relativo:

Definición 3.3. El capital relativo se define como la variable aleatoria

$$S = b \cdot X = \sum_{i=1}^{m} b_i X_i.$$

La variable S representa el rendimiento del mercado en general, siguiendo la estrategia de inversión que indica el portafolio b. Si invertimos una cantidad K en el mercado de valores X, entonces después de un día, terminaremos con un capital de KS. Más adelante se generaliza esta idea para cuando invertimos por más de un día.

El problema principal que nos ocupa es el siguiente. Dada una función de distribución F para X dada, se desea encontrar el valor de b que maximiza el capital S. Aunque algunas entradas del valor X pueden tomar valores muy altos, la probabilidad de que eso ocurra no siempre es alta, por lo que maximizar S depende en gran medida del dinero invertido en ese valor y en la distribución de su precio relativo. Una manera de atacar el problema es tratar de maximizar el valor esperado de la variable S, o equivalentemente, la esperanza del logaritmo de S. Esta idea da lugar a la siguiente definición.

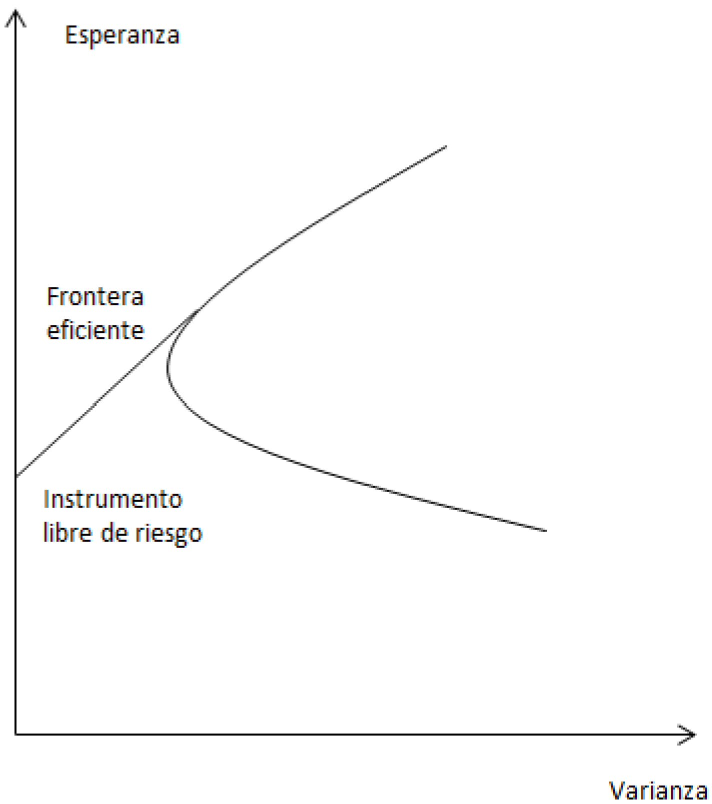
Definición 3.4. La tasa de crecimiento de un portafolio b en un mercado de valores X respecto de una distribución $F(x)$ se define como:

$$W(b, F) = \int_{\mathbb{R}} \log(b \cdot X) dF(x) = E(\log(b \cdot X)).$$

Si el logaritmo es en base 2, la tasa de crecimiento también es llamada la tasa duplicadora.

Cabe recalcar que la teoría estándar de inversión en un mercado de valores se centra en el primer y segundo momento de la variable S. El objetivo se vuelve entonces tratar de maximizar el valor esperado de S con restricciones respecto a la varianza. Dado que generalmente es más fácil calcular dichos momentos, la teoría estándar es más simple que una teoría que se centra en toda la distribución de S.

El enfoque centrado en la media y la varianza es la base de la teoría de inversión en un mercado de valores de Sharpe-Markowitz. Esta teoría se ilustra en la figura siguiente. La figura muestra el conjunto de posibles pares de media y varianza usando distintos portafolios. El conjunto de portafolios en la frontera de la región corresponde al de los portafolios «no dominados»: los portafolios que tienen la esperanza dada cierta varianza. Esta frontera se conoce como la frontera eficiente.



La teoría estándar se simplifica con la introducción de instrumentos libres de riesgo. Estos instrumentos corresponden a un punto en el eje vertical en la figura. Al combinar el instrumento libre de riesgo con otros instrumentos, uno obtiene todos los puntos debajo de la tangente desde el punto del instrumento libre de riesgo hacia la frontera eficiente. La línea entonces se vuelve parte de la frontera eficiente.

Si bien el utilizar la función logaritmo en la definición de tasa de crecimiento permite el uso de ciertas herramientas que facilitan la demostración de teoremas presentados más adelante, es cierto que podría haberse utilizado cualquier otra función creciente de S. Este punto sugiere un posible camino para ampliar el estudio, que es maximizar S. Para una discusión acerca de las condiciones que podrían requerirse a la función de S, vea la sección 7.

Definición 3.5. La tasa de crecimiento óptima se define como

$$W^*(F) = \max_b W(F, b),$$

cuando el máximo existe, con $b_i \geq 0$ y $\displaystyle\sum_{i=1}^{n} b_i = 1$.

Definición 3.6. A un portafolio b^* que hace que se alcance el máximo se le llama portafolio log-óptimo o portafolio de crecimiento óptimo.

Note que no nos hemos referido a la existencia del portafolio log-óptimo. Una posible línea de investigación sería encontrar condiciones para la existencia de este portafolio óptimo. En este trabajo, no obstante, se indican condiciones necesarias y suficientes para determinar si un portafolio es log-óptimo, las condiciones de Khun-Tucker, que se explican más adelante.

Ahora veamos un ejemplo en el que se resalta la relación de los conceptos definidos con el concepto de entropía.

Ejemplo 3.7. Sea X una variable aleatoria tal que $X = (1, 2)$ con probabilidad p y $X = (1, \frac{1}{2})$ con probabilidad $1-p$. Este variable representa el caso de dos acciones con únicamente dos comportamientos. La primera acción no cambia de precio. La segunda acción puede duplicar su valor o reducirlo a la mitad. Se trata de un ejemplo muy sencillo de mercado de valores. Sea $B = \{(b_1, b_2) : b_1 + b_2 = 1\}$. Este conjunto de posibles portafolios no incluye la condición $b_i \geq 0$.

- Encontraremos el portafolio log-óptimo y relacionaremos la tasa de crecimiento con la entropía de la distribución.

Para encontrar el portafolio log-óptimo, maximicemos la tasa de crecimiento.

$$\begin{aligned}
\mathbb{E}(\log(b_1X_1 + b_2X_2)) &= \mathbb{E}(\log(b_1 + b_2X_2)) \\
&= p\log(b_1 + 2b_2) + (1-p)\log\left(b_1 + \frac{1}{2}b_2\right) \\
&= p\log(1 - b_2 + 2b_2) + (1-p)\log\left(1 - b_2 + \frac{1}{2}b_2\right) \\
&= p\log(1 + b_2) + (1-p)\log\left(1 - \frac{1}{2}b_2\right).
\end{aligned}$$

Denotemos por $f(b_2)$ a la última expresión. Maximicemos respecto a b_2, para ello, derivemos e igualemos a cero:

$$\begin{aligned} f'(b_2) &= \frac{p}{(1+b_2)\ln(2)} + \frac{(1-p)(-\frac{1}{2})}{(1-\frac{1}{2}b_2)\ln(2)} \\ &= \frac{p}{(1+b_2)} + \frac{(1-p)(-\frac{1}{2})}{(1-\frac{1}{2}b_2)} \\ &= \frac{p}{(1+b_2)} + \frac{-\frac{1}{2}+\frac{1}{2}p}{(1-\frac{1}{2}b_2)}. \end{aligned}$$

Esta cantidad es cero si y sólo si,

$$\begin{aligned} p\left(1-\frac{1}{2}b_2\right) + \left(-\frac{1}{2}+\frac{1}{2}p\right)(1+b_2) &= 0 \\ p - \frac{1}{2}pb_2 - \frac{1}{2} - \frac{1}{2}b_2 + \frac{1}{2}p + \frac{1}{2}pb_2 &= 0 \\ -\frac{1}{2} - \frac{1}{2}b_2 + \frac{3}{2}p &= 0 \\ -1 - b_2 + 3p &= 0 \\ 3p - 1 &= b_2. \end{aligned}$$

Este valor produce un máximo, como se puede comprobar al calcular la segunda derivada, pues resulta negativa. Note que este valor no es necesariamente mayor a cero. De esta última ecuación podemos obtener el valor de $b_1 = 1 - b_2 = 1 - (3p-1) = 2 - 3p$. Así, hemos encontrado el portafolio log-óptimo.

Para encontrar la relación entre la tasa de crecimiento y la entropía, calculemos la primera:

$$\begin{aligned}
W^*(p) &= p\log(1+b_2) + (1-p)\log\left(1 - \frac{1}{2}b_2\right) \\
&= p\log(1+3p-1) + (1-p)\log\left(1 - \frac{1}{2}(3p-1)\right) \\
&= p\log(3p) + (1-p)\log\left(\frac{3}{2} - \frac{3}{2}p\right) \\
&= p\log(3) + p\log(p) + (1-p)\log\left(\frac{3}{2}\right)(1-p) + \log(1-p) \\
&= p\log(p) + (1-p)\log(1-p) + p\log(3) + (1-p)\log\left(\frac{3}{2}\right) \\
&= -H(p) + p\log(3) + (1-p)\log\left(\frac{3}{2}\right) \\
&= -H(p) + p\log(3) + (1-p)\log(3) - (1-p)\log(2) \\
&= -H(p) + \log(3) - (1-p).
\end{aligned}$$

Al interpretar el ejemplo anterior, nos damos cuenta de que se permite que el inversionista venda valores, algo que hasta este punto no habíamos contemplado. Al permitir que las entradas del portafolio puedan ser negativas, estamos abriendo la posibilidad de la venta de valores cuando su rendimiento no parece que vaya a ser prometedor. Es de resaltar la riqueza del modelo, que permite incluir estas situaciones. También debemos notar que no es forzosa la posesión de los valores al momento de la venta, es decir, podría darse el caso del llamado *short selling*.

En adición, vale la pena considerar un caso particular: cuando el valor esperado del segundo valor es 1. Encontremos el valor para p en el que esto ocurre.

$$1 = \mathbb{E}(X_2) = 2 * p + \frac{1}{2} * (1-p) = \frac{3}{2}p + \frac{1}{2}$$

El valor de p que satisface esta ecuación es $p = \frac{1}{3}$. En este caso, podríamos pensar que el inversionista es indiferente entre invertir en el segundo valor o el primero. Sin embargo, la inversión en el segundo valor conlleva un riesgo, mientras que la inversión en el primer valor no. Cuando $p = \frac{1}{3}$, el portafolio óptimo es $(1, 0)$, lo cual confirma nuestra afirmación y nos indica que invirtamos todo en el primer valor.

Veamos ahora un nuevo teorema, en el que se muestra que el capital del inversionista «crece» de manera similar a 2^{nW^*}, cuando se utiliza la misma política de inversión durante varios días y suponiendo que $F(x)$ no cambia en ese periodo.

TEOREMA 3.8. *Sean $X_1, X_2, ..., X_n$ vectores aleatorios independientes e idénticamente distribuidos de acuerdo a $F(x)$, es decir, una colección de mercados de valores independientes e idénticamente distribuidos. Sea*

$$S_n^* = \prod_{i=1}^{n} b^* X_i,$$

el capital después de n días usando el portafolio log-óptimo b^*. Entonces, cuando $n \to \infty$,

$$\frac{1}{n}\log(S_n^*) \to W^*,$$

con probabilidad 1.

Note que el capital relativo después de n días es el producto del capital relativo de cada día dentro del periodo n. Esto se justifica porque si se tiene al inicio del periodo un capital K, entonces después del día 1, se tendrá $K_1 = KS_1$. Si se invierte el día 2, entonces el capital al final del día 2 es de $K_2 = K_1S_2 = KS_1S_2$ y así sucesivamente. Por ello es que $S_n = \prod_{i=1}^{n} b_iX_i$, y en el caso particular en el que todos los días se use el portafolio óptimo, se tiene que $S_n^* = \prod_{i=1}^{n} b^*X_i$.

Veamos la prueba del teorema.

DEMOSTRACIÓN. Como $E(\log(b^*X_i)) = W^*$, por la ley fuerte de los grandes números, tenemos que, con probabilidad 1 y cuando $n \to \infty$,

$$\frac{1}{n}\log(S_n^*) = \frac{1}{n}\sum_{i=1}^{n}\log(b^*X_i) \to W^*.$$

$\square$

Haciendo uso del resultado anterior, podemos calcular el siguiente límite, en el sentido de convergencia casi segura:

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\log\frac{S_n^*}{2^{nW^*}} &= \lim_{n\to\infty}\frac{1}{n}\log(S_n^*) - \lim_{n\to\infty}\frac{1}{n}\log(2^{nW^*}) \\
&= \lim_{n\to\infty}\frac{1}{n}\log(S_n^*) - \lim_{n\to\infty}\frac{1}{n}(nW^*) \\
&= \lim_{n\to\infty}\frac{1}{n}\log(S_n^*) - \lim_{n\to\infty}W^* \\
&= \lim_{n\to\infty}\frac{1}{n}\log(S_n^*) - W^* \\
&= 0.
\end{aligned}$$

La interpretación del límite anterior es que justamente S_n^* crece de la misma manera que 2^{nW^*}.

Ahora veamos un ejemplo donde se hace uso de los conceptos que acabamos de exponer.

EJEMPLO 3.9. Sea $\mathbf{X}$ una variable aleatoria tal que $\mathbf{X} = (1, a)$ con probabilidad $\frac{1}{2}$ y $\mathbf{X} = (1, \frac{1}{a})$ con la misma probabilidad, donde $a \geq 1$

- Encontraremos el portafolio log-óptimo, la tasa de crecimiento asociada a dicho portafolio, así como el comportamiento asintótico de $S_n = \prod_{i=1}^{n} \mathbf{bX_i}$ para toda $\mathbf{b}$.

Queremos maximizar $W(\mathbf{b}, F) = \mathbb{E}(\log(\mathbf{bX}))$, para ello, calculemos

$$\begin{aligned}
\mathbb{E}(\log(\mathbf{bX})) &= \mathbb{E}(\log(b_1X_1 + b_2X_2)) \\
&= \mathbb{E}(\log(b_1 + b_2X_2)) \\
&= \frac{1}{2}\log(b_1 + ab_2) + \frac{1}{2}\log\left(b_1 + \frac{1}{a}b_2\right) \\
&= \frac{1}{2}\log(1 - b_2 + ab_2) + \frac{1}{2}\log\left(1 - b_2 + \frac{1}{a}b_2\right).
\end{aligned}$$

Denotemos por $f(b_2)$ a la última expresión. Calculemos la derivada de f e igualemos a cero para encontrar un máximo.

$$f'(b_2) = \frac{1}{2}\frac{a-1}{\ln 2(1 - b_2 + ab_2)} + \frac{1}{2}\frac{\frac{1}{a} - 1}{\ln 2(1 - b_2 + \frac{1}{a}b_2)} = 0.$$

Por tanto, tenemos que

$$\begin{aligned}
\frac{a-1}{1 - b_2 + ab_2} + \frac{\frac{1}{a} - 1}{1 - b_2 + \frac{1}{a}b_2} &= 0 \\
(a-1)\left(1 - b_2 + \frac{1}{a}b_2\right) + \left(\frac{1}{a} - 1\right)(1 - b_2 + ab_2) &= 0 \\
a - ab_2 + b_2 - 1 + b_2 - \frac{b_2}{a} + \frac{1}{a} - \frac{b_2}{a} + b_2 - 1 + b_2 - ab_2 &= 0 \\
\left(-2a + 4 - \frac{2}{a}\right)b_2 + \left(a + \frac{1}{a} - 2\right) &= 0 \\
-2b_2 + 1 &= 0.
\end{aligned}$$

Por lo tanto, $b_2 = \frac{1}{2}$ es un punto crítico de la función. Se puede comprobar que la segunda derivada de la función es negativa, por lo tanto este punto es un máximo. Por tanto, el portafolio log-óptimo es $\mathbf{b} = \left(\frac{1}{2}, \frac{1}{2}\right)$ Calculemos ahora la tasa de crecimiento asociada a dicho portafolio.

$$W^*(\mathbf{b}, F) = \frac{1}{2}\log\left(\frac{1}{2} + \frac{a}{2}\right) + \frac{1}{2}\log\left(\frac{1}{2} + \frac{1}{2a}\right).$$

Finalmente, para descubrir el comportamiento de $S_n = \prod_{i=1}^{n} \mathbf{bX_i}$, notemos que:

$$\frac{1}{n}\log(S_n) = \frac{1}{n}\log(\prod_{i=1}^{n}\mathbf{bX_i}) = \frac{1}{n}\sum_{i=1}^{n}\log(\mathbf{bX_i}) \to \mathbb{E}(\log(\mathbf{bX_i})) = W(\mathbf{b}, F).$$

Por lo tanto, se tiene que, cuando $n \to \infty$, en convergencia casi segura,

$$\begin{aligned}\frac{1}{n}\log(S_n) &\to W\\ \log(S_n) &\to nW\\ S_n &\to 2^{nW}.\end{aligned}$$

Veamos ahora algunas propiedades de la tasa de crecimiento:

Teorema 3.10. *Para un mercado de valores X con función de distribución F:*
1.- $W(b,F)$ es lineal en F, es decir, $W(b,F+\alpha G) = W(b,F) + \alpha W(b,G)$ con F,G cualesquiera dos funciones de distribución de X y α cualquier número real
2.- $W(b,F)$ es cóncava en b.
3.- $W^*(F)$ es convexa en F.

Demostración.

1.- Esto se sigue de que

$$\begin{aligned}W(b,F+\alpha G) &= \int \log(bX)d(F(x)+\alpha G(x))\\ &= \int \log(bX)dF(x) + \alpha\int \log(bX)dG(x)\\ &= W(b,F)+\alpha W(b,G).\end{aligned}$$

Recuerde que en general no se cumple que $F+\alpha G$ sea una función de distribución. Esto puede ser un problema porque $W(b,F) := \mathbb{E}(\log(b\cdot X))$ no tiene sentido cuando en lugar de F se usa $F+\alpha G$. En cambio, si se usa por ejemplo $\alpha F + (1-\alpha)G$, con $0 \leq \alpha \leq 1$, nos mantendremos dentro del espacio de las funciones de probabilidad. No obstante, las propiedades de linealidad de la integral de Riemann-Stieltjes nos permiten llegar a la conclusión.

2.- Sabemos que la función logaritmo es cóncava, por lo tanto, para $0 \leq \lambda \leq 1$,

$$\log((\lambda b_1 + (1-\lambda)b_2)X) \geq \lambda\log(b_1X) + (1-\lambda)\log(b_2X).$$

Si tomamos esperanza de ambos lados, tendremos el resultado, por la definición de $W(b,F)$.

3.- Sean F_1 y F_2 dos distribuciones en el mercado de valores y sean b_1^* y b_2^* sus respectivos portafolios log-óptimos. Sea b_3^* el portafolio log-óptimo correspondiente a la distribución $\lambda F_1 + (1-\lambda)F_2$. Entonces, usando el punto 1, tenemos:

$$\begin{aligned} W^*(\lambda F + (1-\lambda)F_2) &= W(b_3^*, \lambda F + (1-\lambda)F_2) \\ &= \lambda W(b_3^*, F_1) + (1-\lambda)W(b_3^*, F_2) \\ &\leq \lambda W(b_1^*, F_1) + (1-\lambda)W(b_2^*, F_2). \end{aligned}$$

Esto último es gracias a que b_1^* maximiza a $W(b, F_1)$ y b_2^* maximiza a $W(b, F_2)$. □

Ahora veamos un pequeño lema:

Lema 3.11. *El conjunto de los portafolios log-óptimos es convexo.*

Demostración. Supongamos que b_1 y b_2 son portafolios log-óptimos, es decir, que $W(b_1, F) = W(b_2, F) = W^*(F)$, entonces, usando la concavidad de W en b, tenemos:

$$W(\lambda b_1 + (1-\lambda)b_2, F) \geq \lambda W(b_1, F) + (1-\lambda)W(b_2, F) = W^*(F).$$

Esto nos dice que $\lambda b_1 + (1-\lambda)b_2$ es también un portafolio log-óptimo. □

Este último resultado no solo dice que el conjunto de los portafolios log-óptimos es convexo. Implícitamente dice que dicho conjunto o bien es infinito, o bien tiene un solo elemento.

Un ejemplo donde el conjunto tiene un solo elemento es cuando $X = (X_1, X_2)$ y $X_1 = 2, X_2 = \frac{1}{2}$. En este caso, es claro que $b^* = (1, 0)$. Analíticamente esta afirmación puede comprobarse, pues

$$\begin{aligned} \mathbb{E}(b_1X_1 + b_2X_2) &= 2b_1 + \frac{1}{2}b_2 \\ &= 2b_1 + \frac{1}{2}(1 - b_1) \\ &= \frac{3}{2}b_1 + \frac{1}{2}, \end{aligned}$$

que se maximiza si y solo si $b_1 = 1$.

Por otro lado, un ejemplo donde existe una infinidad de portafolios óptimos es cuando $X = (X_1, X_2)$ con $X_1 = X_2 = 1$, pues en este caso,

$$\begin{aligned} \mathbb{E}(b_1X_1 + b_2X_2) &= \mathbb{E}(b_1 + b_2) \\ &= 1. \end{aligned}$$

Por lo que cualquier portafolio es óptimo.

Un problema adicional a trabajar es el de encontrar un ejemplo de mercado de valores donde los portafolios óptimos sean infinitos, pero no sean todo el conjunto de portafolios factibles, como en el ejemplo anterior. La estrategia natural es analizar y tratar de manipular la función de distribución del mercado de valores. No obstante, una conjetura es que los ejemplos factibles no admiten que alguno de los valores sea constante.

2. Caracterización Kuhn-Tucker del portafolio log-óptimo

Denotemos a $\mathcal{B} = \{b \in \mathbb{R}^m : b_i \geq 0, \sum_{i=1}^m b_i = 1\}$ el conjunto de los portafolios permitidos. Hay que resaltar que el problema de encontrar un portafolio que maximice la función $W(b)$ es un problema de optimización de una función cóncava en el conjunto convexo $\mathcal{B}$. Para este tipo de problemas existen las condiciones de Khun-Tucker generales. No obstante, en esta sección derivaremos nuestras propias condiciones de Kuhn-Tucker. Dichas condiciones son una condición necesaria y suficiente para determinar si un portafolio es log-óptimo, dada una función de distribución para el mercado de valores.

TEOREMA 3.1. ***Condiciones de Kuhn-Tucker.***
Sea $b^* = (b_1^*, ..., b_m^*)$ un portafolio para un mercado de valores $X = (X_1, ..., X_m)$ con distribución F. El portafolio b^* es un portafolio log-óptimo si y sólo si satisface las siguientes condiciones para $i = 1, ..., m$:

1.- $E\left(\frac{X_i}{b^*X}\right) = 1$ *si* $b_i^* > 0$.

2.- $E\left(\frac{X_i}{b^*X}\right) \leq 1$ *si* $b_i^* = 0$.

DEMOSTRACIÓN. Como la tasa de crecimiento $W(b)$ es cóncava sobre el conjunto de portafolios, debemos notar que el vector b^* es log-óptimo si y sólo si la derivada direccional de W desde b^* hacia cualquier otro portafolio b es no positiva. Así, haciendo $b_\lambda = (1-\lambda)b^* + \lambda b$, con $0 \leq \lambda \leq 1$, tenemos:

$$\frac{d}{d\lambda} W(b_\lambda)\Big|_{\lambda=0^+} \leq 0.$$

Veamos que esta condición se reduce a las dos condiciones que se plantean en el enunciado del teorema.

$$
\begin{aligned}
\frac{d}{d\lambda}W(b_\lambda)\Big|_{\lambda=0^+} &= \lim_{\lambda\to 0^+} \frac{W(b_\lambda)-W(b_0)}{\lambda}\\
&= \lim_{\lambda\to 0^+} \frac{E(\ln(b_\lambda\cdot X))-E(\ln(b_0 X))}{\lambda}\\
&= \lim_{\lambda\to 0^+} \frac{1}{\lambda}E(\ln(b_\lambda\cdot X)-\ln(b^*\cdot X))\\
&= \lim_{\lambda\to 0^+} \frac{1}{\lambda}E\left(\ln\left(\frac{b_\lambda\cdot X}{b^*\cdot X}\right)\right)\\
&= \lim_{\lambda\to 0^+} \frac{1}{\lambda}E\left(\ln\left(\frac{((1-\lambda)b^*+\lambda b)\cdot X}{b^*\cdot X}\right)\right)\\
&= \lim_{\lambda\to 0^+} \frac{1}{\lambda}E\left(\ln\left((1-\lambda)+\lambda\frac{b\cdot X}{b^*\cdot X}\right)\right)\\
&= \lim_{\lambda\to 0^+} \frac{1}{\lambda}E\left(\ln\left(1+\lambda\left(\frac{b\cdot X}{b^*\cdot X}-1\right)\right)\right)\\
&= E\left(\lim_{\lambda\to 0^+} \frac{1}{\lambda}\ln\left(1+\lambda\left(\frac{b\cdot X}{b^*\cdot X}-1\right)\right)\right)\\
&= E\left(\lim_{\lambda\to 0^+} \frac{\frac{\frac{b\cdot X}{b^*\cdot X}-1}{1+\lambda\left(\frac{b\cdot X}{b^*\cdot X}-1\right)}}{1}\right)\\
&= E\left(\frac{\frac{b\cdot X}{b^*\cdot X}-1}{1}\right)\\
&= E\left(\frac{b\cdot X}{b^*\cdot X}-1\right)\leq 0.
\end{aligned}
$$

Entonces finalmente tenemos:

$$E\left(\frac{b\cdot X}{b^*\cdot X}\right)\leq 1.$$

Hay que aclarar que el intercambio entre el límite y la integral lo podemos hacer gracias al teorema de convergencia dominada de Lebesgue. Además, se usó la regla de L'Hôpital para calcular el límite.

Note que si se toma $\lambda<0$, los vectores que se obtienen no necesariamente estan dentro de $\mathcal{B}$, pero si para cierto b esto se cumple, entonces en la última expresión tendremos una igualdad. Lo último es cierto, porque de no serlo, entonces la derivada existiría en todo λ lo suficientemente cercano a b^* y sería negativa, lo cual implicaría que la función $W(b_l)$ sería decreciente y podríamos encontrar un portafolio elegible b_λ que tendría una tasa de crecimiento mayor a la de b^*, lo cual es una contradicción, por lo que se debe tener una igualdad. Si no podemos tomar $\lambda < 0$ para alguna b, entonces sólo tenemos la expresión a la que llegamos.

Ahora bien, la última condición es cierta para todo vector b si y sólo si es cierta para todo punto extremo de $\mathcal{B}$, esto se debe a la linealidad de $E\left(\frac{b \cdot X}{b^* \cdot X}\right)$, que se aplica como sigue:

$$E\left(\frac{b \cdot X}{b^* \cdot X}\right) = \sum_{i=1}^{m} b_i E\left(\frac{e_i X}{b^* X}\right) \leq \sum_{i=1}^{m} b_i = 1.$$

Donde e_i es el vector de $\mathbb{R}^m$ que tiene 1 en la i-ésima entrada y 0 en las demás. Más aún, el segmento de línea entre b^* y b, (con b tal que $b_j = 1$, $b_i = 0$ para $i \neq j$), puede extenderse con $\lambda < 0$ si y solo si $b_j^* > 0$. Esto porque como $b_\lambda = ((1-\lambda)b_1^*, ..., (1-\lambda)b_j^* + \lambda, ..., (1-\lambda)b_m^*)$, se tiene que $(1-\lambda)b_j^* + \lambda \geq 0$ si y solo si $b_j^* \geq \frac{-\lambda}{1-\lambda}$ y como $\lambda < 0$, lo último ocurre si y solo si $b_j^* > 0$. Por ello, las siguientes condiciones son necesarias y suficientes para caracterizar un portafolio log-óptimo:

1.- $E\left(\frac{X_i}{b^* X}\right) = 1 \ si \ b_i^* > 0.$

2.- $E\left(\frac{X_i}{b^* X}\right) \leq 1 \ si \ b_i^* = 0.$ □

Como hemos mencionado antes, estas condiciones caracterizan un portafolio log-óptimo. No obstante, no se garantiza en ningún momento que el portafolio log-óptimo exista.

Comprobemos el teorema usando uno de los ejemplos anteriores.

Ejemplo 3.2. Sea $\mathbf{X}$ una variable aleatoria tal que $\mathbf{X} = (1, 2)$ con probabilidad p y $\mathbf{X} = (1, \frac{1}{2})$ con probabilidad $1 - p$. En la sección anterior encontramos que el portafolio óptimo era $b = (2 - 3p, 3p - 1)$. Consideremos primero el caso en que $p = \frac{5}{6}$. En este caso, el portafolio óptimo sería $b = (-\frac{1}{2}, \frac{3}{2})$, pero los supuestos que hemos utilizado indican que las entradas del portafolio deben ser positivas, por ello, $b = (0, 1)$.
Ahora bien, calculemos $E\left(\frac{X_i}{b^* \cdot X}\right)$ para $i = 1, 2$.
Si $i = 1$ tenemos que $E\left(\frac{X_i}{b^* \cdot X}\right) = E\left(\frac{X_1}{b^* \cdot X}\right) = E\left(\frac{1}{X_2}\right) = \frac{9}{12} \leq 1$. Este resultado es consistente con el teorema. Ahora, si $i = 2$, tenemos que $E\left(\frac{X_i}{b^* \cdot X}\right) = E\left(\frac{X_2}{b^* \cdot X}\right) = E\left(\frac{X_2}{X_2}\right) = 1$, que también ilustra el teorema.

Ahora analicemos el caso en el que $p = \frac{1}{2}$. En este caso, $b = (\frac{1}{2}, \frac{1}{2})$. Entonces $E\left(\frac{X_1}{b^* \cdot X}\right) = E\left(\frac{2}{1+X_2}\right) = 1$. Similarmente, $E\left(\frac{X_2}{b^* \cdot X}\right) = E\left(\frac{2X_2}{1+X_2}\right) = 1$.

Para continuar con nuestro análisis, demostraremos un pequeño lema.

LEMA 3.3. *Sean $X_1, ..., X_m$ variables aleatorias idénticamente distribuidas y sean $a_1, ..., a_m$ números reales con al menos un valor a_i distinto de cero. Entonces*

$$E(X_i|a_1X_1 + ... + a_mX_m = k) = \frac{k}{a_1 + \cdots + a_m}.$$

DEMOSTRACIÓN. Por la hipótesis de idéntica distribución, tenemos

$$E(X_1|a_1X_1 + ... + a_mX_m = k) = E(X_i|a_1X_1 + ... + a_mX_m = k).$$

Por lo tanto:

$$\left(\sum_{i=1}^{m} a_i\right) E(X_i|a_1X_1 + ... + a_mX_m = k) = E\left(\sum_{i=1}^{m} a_iX_i \middle| a_1X_1 + ... + a_mX_m = k\right) = k.$$

□

Gracias a este lema, podemos probar un resultado que, aunque intuitivo, puede resultar enredoso de demostrar. Cabe mencionar que este resultado es original de este trabajo.

TEOREMA 3.4. *Bajo la hipótesis de idéntica distribución en el mercado de valores, cualquier portafolio es log-óptimo.*

DEMOSTRACIÓN. Para demostrar el resultado, basta demostrar que $W(b, F)$ es constante para cualquier b. Para ello, consideremos las derivadas parciales de $W(b, F)$ con respecto a b_i.

$$
\begin{aligned}
\frac{\partial}{\partial b_i} E\left(\log(\sum_{i=1}^{m} b_i X_i)\right) &= E\left(\frac{X_i}{(\sum_{i=1}^{m} b_i X_i)}\right) \\
&= \int E\left(\frac{X_i}{(\sum_{i=1}^{m} b_i X_i)} | \sum_{i=1}^{m} b_i X_i = k\right) f_S(k) dk \\
&= \int E\left(\frac{X_i}{k} | \sum_{i=1}^{m} b_i X_i = k\right) f_S(k) dk \\
&= \int \frac{1}{k} E\left(X_i | \sum_{i=1}^{m} b_i X_i = k\right) f_S(k) dk \\
&= \int \frac{1}{k} \frac{k}{b_1 + ... + b_m} f_S(k) dk \\
&= \int f_S(k) dk \\
&= 1.
\end{aligned}
$$

Esto implica entonces que $E\left(\log(\sum_{i=1}^{m} b_i X_i)\right) = b_1 + \cdots + b_m + c = 1 + c$ con c constante. □

Veamos ahora otro ejemplo en el que queda en evidencia la utilidad de las condiciones de Khun-Tucker.

Ejemplo 3.5. Considere un mercado de valores de dos acciones $X = (X_1, X_2)$. Supongamos que $X_1 = 2$ con probabilidad 1, es decir, que una inversión en el primer valor con seguridad se duplica al final del día.

- Encontraremos condiciones necesarias y suficientes sobre X_2 para que el portafolio log-óptimo $\mathbf{b}^*$ invierta todo en el segundo valor, es decir $\mathbf{b}^* = (0, 1)$.
- Comprobaremos que para cualquier distribución sobre X_2 se cumple que $W^* \geq 1$.

Este ejemplo da condiciones para que aun cuando la primera acción duplique su valor al final del día con probabilidad 1, no invirtamos en ese valor, sino en la segunda acción. Intuitivamente, se podría pensar que para que esa situación ocurriera, se debería tener que, por ejemplo, la segunda acción triplicará su precio al final del día. De acuerdo con esta línea de pensamiento, tenemos lo siguiente.

Afirmamos que la condición necesaria y suficiente para que el portafolio $b^* = (0, 1)$ sea log-óptimo es que $\mathbb{E}(\frac{2}{X_2}) \leq 1$.

Esto es una consecuencia directa de las condiciones de Khun-Tucker, pues $b^* = (0, 1)$ es óptimo si y solo si:

- $\mathbb{E}\left(\dfrac{X_1}{0 * X_1 + 1 * X_2}\right) \leq 1.$
- $\mathbb{E}\left(\dfrac{X_2}{0 * X_1 + 1 * X_2}\right) = 1.$

Las cuales se traducen en

- $\mathbb{E}\left(\dfrac{2}{X_2}\right) \leq 1.$
- $\mathbb{E}\left(\dfrac{X_2}{X_2}\right) = 1.$

Por ejemplo, si tenemos que $X_2 = c$, con $c \geq 2$, entonces se satisfacen las condiciones de Khun-Tucker y, por ello, para cualquier portafolio (b_1, b_2)se tiene que:

$$\begin{aligned}
2 \leq c &\Rightarrow 2b_1 \leq cb_1 \\
&\Rightarrow 2b_1 \leq c(1 - b_2) \\
&\Rightarrow 2b_1 + cb_2 \leq c \\
&\Rightarrow \log(2b_1 + cb_2) \leq \log(c) \\
&\Rightarrow \mathbb{E}(\log(2b_1 + cb_2)) \leq \mathbb{E}(\log(c)) \\
&\Rightarrow \mathbb{E}(\log(2b_1 + b_2X_2)) \leq \mathbb{E}(\log(X_2)).
\end{aligned}$$

Ahora, para el segundo punto, supongamos que X_2 tiene cualquier distribución. La tasa de crecimiento óptima será mayor que la de cualquier otro portafolio, en particular para la del portafolio $\mathbf{b} = (1, 0)$, cuya tasa de crecimiento es $W = \mathbb{E}(2b_1 + b_2X_2) = \mathbb{E}(\log(2)) = \log(2) = 1$

Veamos ahora un resultado que es consecuencia del teorema de las condiciones de Khun-Tucker. En este resultado, se establece una equivalencia de desigualdades que si bien podría parecer obvia o intuitiva, requiere del teorema principal para su demostración.

Teorema 3.6. *Sea $S^* = b^*X$ el capital aleatorio que resulta de usar el portafolio log-óptimo b^*. Sea $S = b \cdot X$ el capital que resulta de usar cualquier otro portafolio b. Entonces:*

$$E\left(\ln\left(\frac{S}{S^*}\right)\right) \leq 0 \quad \textit{para toda } S \quad \Leftrightarrow \quad E\left(\frac{S}{S^*}\right) \leq 1 \quad \textit{para toda } S.$$

Lo que este teorema indica es que en valor esperado, el cociente entre la ganancia que se alcanza usando un portafolio cualquiera y la ganancia que se alcanza de utilizar un portafolio log-óptimo es menor a 1 si y solo si el logaritmo de dicho cociente es, en valor esperado, menor a cero. Note que si quitásemos los símbolos de esperanza de ambos lado de la desigualdad, el resultado sería trivial, por lo que la importancia del teorema es que asegura la equivalencia de las desigualdades, al menos en valor esperado.

DEMOSTRACIÓN. Notemos que

$$\begin{aligned} E\left(\ln\left(\frac{S}{S^*}\right)\right) \leq 0 &\Leftrightarrow E(\ln(S)) \leq E(\ln(S^*)) \\ &\Leftrightarrow W(b) \leq W(b^*), \end{aligned}$$

por lo que en realidad lo que esa condición nos dice es que b^* es log óptimo, por tanto, debemos demostrar que:

$$b^* \text{es log-óptimo} \quad \Leftrightarrow \quad E\left(\frac{S}{S^*}\right) \leq 1 \quad \text{para toda S.}$$

$\Rightarrow$]
Del teorema anterior, tenemos que

$$E\left(\frac{X_i}{b^{*^t}X}\right) \leq 1,$$

para toda i. Ahora, si multiplicamos esta ecuación por b_i y sumamos sobre todos los valores del mercado, tendremos que

$$\sum_{i=1}^{m} b_i E\left(\frac{X_i}{b^{*^t}X}\right) \leq \sum_{i=1}^{m} b_i = 1.$$

Y como:

$$E\left(\frac{S}{S^*}\right) = E\left(\frac{b^t X}{b^{*^t}X}\right) = \sum_{i=1}^{m} b_i E\left(\frac{X_i}{b^{*^t}X}\right),$$

tenemos el resultado.
$\Leftarrow$]
Aquí usamos la desigualdad de Jensen con la función cóncava logaritmo natural:

$$E\left(\ln\left(\frac{S}{S^*}\right)\right) \leq \ln\left(E\left(\frac{S}{S^*}\right)\right) \leq \ln(1) = 0.$$

□

Analicemos ahora un ejemplo que hace uso de las condiciones de Khun-Tucker y cuya interpretación es muy interesante.

EJEMPLO 3.7. Considere el conjunto de densidades que tienen el mismo portafolio óptimo. Sea $P_{\mathbf{b}_0}$ el conjunto de todas las densidades de probabilidad en $\mathbb{R}^m$ donde $\mathbf{b}_0$ es óptimo, es decir,

$$P_{\mathbf{b}_0} = \{p(x) : \int \ln(\mathbf{b^t}x)p(x)dx \text{ se maximiza con } \mathbf{b} = \mathbf{b}_0\}.$$

Mostraremos que $P_{\mathbf{b}_0}$ es convexo.

Notemos que el enfoque es ahora diferente al del problema original. En este ejemplo, suponemos que tenemos una estrategia de inversión fija b_0 y nos fijamos en todas las densidades del mercado de valores para las que esa estrategia es óptima. En cierto sentido, vamos en busca de aquellos mercados en los que nuestra estrategia de inversión es óptima. En el problema original, buscamos adaptarnos al mercado, al cambiar nuestra estrategia de inversión.

El problema de buscar el mercado en que nuestra estrategia de inversión es óptima puede ser un camino adicional para alguna futura investigación en este análisis.

Para probar que el conjunto es convexo, debemos tomar $r(x) = \lambda p(x) + (1-\lambda)q(x)$ con p, q elementos en el conjunto y $\lambda \in (0,1)$, y probar que $r(x)$ está en el conjunto. Para ver que esta nueva distribución está en el conjunto, hay que ver lo siguiente:

$$\int \ln(\mathbf{b}x)r(x)dx \leq \int \ln(\mathbf{b_0}x)r(x)dx.$$

Lo cual ocurre si y sólo si:

$$\begin{aligned}
&\int (\ln(\mathbf{b}x) - \ln(\mathbf{b_0}x))r(x)dx \leq 0 \\
\Leftrightarrow \quad & \int \ln \frac{\mathbf{b}x}{\mathbf{b_0}x} r(x)dx \leq 0 \\
\Leftrightarrow \quad & \mathbb{E}_r \left(\ln \frac{\mathbf{bX}}{\mathbf{b_0X}} \right) \leq 0 \\
\Leftrightarrow \quad & \mathbb{E}_r \left(\frac{\mathbf{bX}}{\mathbf{b_0X}} \right) \leq 1 \\
\Leftrightarrow \quad & \int \frac{\mathbf{b}x}{\mathbf{b_0}x} r(x)dx \leq 1 \\
\Leftrightarrow \quad & \int \frac{\mathbf{b}x}{\mathbf{b_0}x} (\lambda p(x) + (1-\lambda)q(x))dx \leq 1 \\
\Leftrightarrow \quad & \lambda \int \frac{\mathbf{b}x}{\mathbf{b_0}x} p(x)dx + (1-\lambda) \int \frac{\mathbf{b}x}{\mathbf{b_0}x} q(x)dx \leq 1 \\
\Leftrightarrow \quad & \lambda \mathbb{E}_p \left(\frac{\mathbf{bX}}{\mathbf{b_0X}} \right) + (1-\lambda)\mathbb{E}_q \left(\frac{\mathbf{bX}}{\mathbf{b_0X}} \right) \leq 1.
\end{aligned}$$

La última desigualdad es cierta debido a que $\mathbb{E}_p \left(\ln \frac{\mathbf{bX}}{\mathbf{b_0X}} \right) \leq 0$ y, similarmente, $\mathbb{E}_q \left(\ln \frac{\mathbf{bX}}{\mathbf{b_0X}} \right) \leq 0$, o equivalentemente, $\mathbb{E}_P \left(\frac{\mathbf{bX}}{\mathbf{b_0X}} \right) \leq 1$ y $\mathbb{E}_q \left(\frac{\mathbf{bX}}{\mathbf{b_0X}} \right) \leq 1$. Estas últimas dos desigualdades concluyen la desigualdad de arriba.

Hemos demostrado que el portafolio log-óptimo, además de maximizar la tasa de crecimiento, también maximiza el capital relativo esperado $\left(E\left(\frac{S}{S^*}\right)\right)$ por un día.

Otra consecuencia de la caracterización de Kuhn-Tucker del portafolio log-óptimo es el hecho de que la proporción de capital esperado en cada valor bajo el uso del portafolio log-óptimo no cambia de un día a otro.

La distribución del capital al inicio del día esta en b^*. La proporción del capital en el valor i al final del día es $\frac{b_i^* X_i}{b^* X}$ cuyo valor esperado es

$$\begin{aligned} E\left(\frac{b_i^* X_i}{b^* X}\right) &= b_i^* E\left(\frac{X_i}{b^* X}\right) \\ &= b_i^*. \end{aligned}$$

Esto es independiente de si $b_i^* > 0$ ó $b_i^* = 0$, puesto que si $b_i^* > 0$ entonces $E\left(\frac{X_i}{b^* X}\right) = 1$ y la igualdad anterior se sigue cumpliendo.

Entonces, la proporción de capital en el valor i esperado al final del día es el mismo que la proporción invertida en el valor i al inicio del día.

3. Optimalidad asintótica del portafolio log-óptimo

En esta sección se demostrará que el inversionista que use el portafolio log-óptimo no le irá peor que cualquier otro inversionista que use cualquier otro portafolio.

Primero consideremos:

$$S_n = \prod_{i=1}^{n} b_i X_i,$$

donde $X_1, X_2, ..., X_n$ son mercados de valores independientes e indénticamente distribuidos de acuerdo a la función de distribución F. La variable S_n es el capital relativo después de n días usando el portafolio b. Solamente consideraremos portafolios que quizá dependan de los pasados valores del mercado de valores, pero independientes de los futuros.

Es de resaltar que el supuesto de independencia e idéntica distribución en una secuencia del mercado de valores es muy fuerte. La idéntica distribución podría justificarse para un periodo «corto» de n días, argumentando que por ejemplo, históricamente en una época de turbulencia financiera, las variaciones de un día son similares con respecto a las del día anterior. No obstante, el supuesto de independencia es difícil de justificar, por la misma razón antes expuesta: durante periodos en los que la bolsa está a la «alza» o a la «baja», los incrementos a decrementos porcentuales suelen ser dependientes de los observados en el pasado. Para una introducción al análisis cuyo supuesto es la estacionariedad del mercado, véase la sección 5.

Veamos ahora una pequeña definición.

Definición 3.1. Una estrategia de portafolios causal (o no anticipante) es una secuencia de mapeos $b_i : \mathbb{R}^{m(i-1)} \to \mathcal{B}$, en donde $\mathcal{B} = \{b \in \mathbb{R}^m : b_i \geq 0, \sum_{i=1}^m b_i = 1\}$.

Esta definición se interpreta en el sentido de que el portafolio b_i es usado el día i ($b_i(\mathbf{x_1}, ..., \mathbf{x_m}) = \mathbf{b_i}$), por ello, b_i depende de los anteriores $i-1$ mercados (de m entradas). Note que aun cuando pareciera que la elección de un portafolio depende de los valores anteriores del mercado, la definición anterior de hecho no limita la secuencia de portafolios que pudiera usarse durante el periodo de n días, permitiendo que la secuencia de portafolios sea independiente de los valores del mercado en días pasados. No obstante, se emplea la definición presentada porque será de utilidad más adelante, cuando hablemos de mercados con la propiedad de estacionariedad.

Veamos ahora un pequeño lema que hace uso de la anterior definición:

Lema 3.2. *Sea S_n^* el capital después de n días usando la estrategia no anticipante log-óptima en una colección de mercados de valores i.i.d. $X_1, X_2, ..., X_n$ y sea S_n el capital usando otra estrategia. Entonces:*

$$E(\ln(S_n^*)) = nW^* \geq E(log(S_n)).$$

La interpretación del lema es bastante sencilla. Nos dice que si nos enfrentamos a un mercado cuyos valores diarios son independientes entre sí e idénticamente distribuidos, entonces la estrategia de inversión óptima en el periodo de n días será la que considere el portafolio log-óptimo para el caso de un día y la repita durante la duración del periodo.

Demostración.

$$\begin{aligned}
\max_{\mathbf{b_1},\ldots,\mathbf{b_n}} E(\ln(S_n)) &= \max_{\mathbf{b_1},\ldots,\mathbf{b_n}} E(\sum_{i=1}^{m} \ln(\mathbf{b_i X_i})) \\
&= \sum_{i=1}^{m} \max_{\mathbf{b_i}(\mathbf{X_1},\ldots,\mathbf{X_{i-1}})} E(\ln(\mathbf{b_i}(\mathbf{X_1},...,\mathbf{X_{i-1}})\mathbf{X_i})) \\
&= \sum_{i=1}^{m} E(\ln(\mathbf{b}^* X_i)) \\
&= nW^*.
\end{aligned}$$

□

Hasta ahora hemos probado que b^* maximiza el capital esperado y que $\lim_{n\to\infty} \frac{1}{n} \log \frac{S_n^*}{2^{nW^*}} = 1$ con probabilidad 1.

Ahora probamos un resultado mucho más fuerte, que intuitivamente nos dice que S_n^* excede el capital de cualquier otro inversionista para casi cualquier secuencia de situaciones del mercado de valores. A este resultado se le conoce como la optimalidad asintótica del portafolio log-óptimo.

Teorema 3.3. *Sea $\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n}$ una colección de mercados de valores i.i.d. que siguen la distribución $F(\mathbf{x})$. Sea $S_n^* = \prod_{i=1}^n \mathbf{b}^* \mathbf{X_i}$ donde b^* es el portafolio log-óptimo, y $S_n = \prod_{i=1}^n \mathbf{b_i^t X_i}$ es el capital resultante de usar cualquier otra estrategia de portafolios causal. Entonces:*

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{S_n}{S_n^*} \leq 0, \quad \textit{con probabilidad 1.}$$

Demostración. De las condiciones de Kuhn-Tucker y la optimalidad de S_n^*, tenemos:

$$E\left(\frac{S_n}{S_n^*} \leq 1\right).$$

Entonces, por la desigualdad de Markov, tenemos

$$\mathbb{P}(S_n > t_n S_n^*) = \mathbb{P}\left(\frac{S_n}{S_n^*} > t_n\right) \leq \frac{E\left(\frac{S_n}{S_n^*}\right)}{t_n} \leq \frac{1}{t_n}.$$

Entonces

$$\mathbb{P}\left(\frac{1}{n} \log \frac{S_n}{S_n^*} > \frac{1}{n} \log t_n\right) \leq \frac{1}{t_n}.$$

Haciendo $t_n = n^2$ y sumando sobre n tenemos:

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\frac{1}{n} \log \frac{S_n}{S_n^*} > \frac{2\log n}{n}\right) \leq \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty.$$

Entonces por el lema de Borel-Cantelli, tenemos que:

$$\mathbb{P}\left(\omega \in \left(\frac{1}{n}\log\frac{S_n}{S_n^*} > \frac{2\log n}{n}\right) \text{ para una infinidad de valores de } n\right) = 0.$$

Es decir, para casi toda $\omega \in \Omega$,

$$\omega \notin \left(\frac{1}{n}\log\frac{S_n}{S_n^*} > \frac{2\log n}{n}\right),$$

i.e para casi toda ω,

$$\omega \in \left(\frac{1}{n}\log\frac{S_n}{S_n^*} > \frac{2\log n}{n}\right)^c,$$

i.e

$$\mathbb{P}\left(\left(\omega \in \left(\frac{1}{n}\log\frac{S_n}{S_n^*} > \frac{2\log n}{n}\right) \text{ para una infinidad de valores de } n\right)^c\right) = 1.$$

Por lo tanto, para casi toda ω existe N tal que para cualquier $n \geq N$, no se cumple que $\frac{1}{n}\log\frac{S_n}{S_n^*} > \frac{2\log n}{n}$. Es decir $\frac{1}{n}\log\frac{S_n}{S_n^*} \leq \frac{2\log n}{n}$. Por lo tanto, para estas ω, se tiene que

$$\begin{aligned}
\limsup_{n\to\infty}\frac{1}{n}\log\frac{S_n}{S_n^*} &= \lim_{n\to\infty}\left(\sup_{k\geq n}(\frac{1}{k}\log\frac{S_k}{S_k^*})\right) \\
&\leq \lim_{n\to\infty}\left(\frac{2\log n}{n}\right) \\
&= 0.
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
1 &= \mathbb{P}\left(\left(\omega \in \left(\frac{1}{n}\log\frac{S_n}{S_n^*} > \frac{2\log n}{n}\right) \text{ para una infinidad de valores de } n\right)^c\right) \\
&\leq \mathbb{P}\left(\omega : \limsup_{n\to\infty}\frac{1}{n}\log\frac{S_n}{S_n^*} \leq 0\right) \\
&\leq 1.
\end{aligned}$$

Por lo que

$$\mathbb{P}\left(\omega : \limsup_{n\to\infty}\frac{1}{n}\log\frac{S_n}{S_n^*} \leq 0\right) = 1.$$

$\square$

Veamos ahora un ejemplo relacionado con la situación planteada al inicio de esta sección.

Ejemplo 3.4. Consideremos un mercado de valores $\mathbf{X}$ con distribución $F(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}_+^m$. Suponga que un «experto» sugiere el portafolio $\mathbf{b}$ que resultaría en un capital $\mathbf{b}\cdot\mathbf{X}$. Añadimos esto al mercado y obtenemos $\mathbf{X}' = (X_1, ..., X_m, \mathbf{b}\cdot\mathbf{X})$. Mostraremos que la nueva tasa de crecimiento óptima es la misma que antes.

Como se menciona, este ejemplo modela la situación en la que un experto sugiere una estrategia de inversión. En términos más coloquiales, se trata de modelar la existencia de fondos de inversión $(\mathbf{b}\cdot\mathbf{X})$. Un fondo de inversión es una institución que vende acciones al público inversionista y usa las ganancias para comprar una selección, o portafolio, de varios activos financieros. El inversionista que posee acciones del fondo de inversión acepta la ganancia y el riesgo asociado con el portafolio. Si el valor del portafolio se incrementa, el inversionista se beneficia. Si el valor del portafolio cae, el accionista sufre una pérdida. Para una introducción a los fondos de inversión y una introducción en general a los mercados financieros, véase [**3**].

En nuestro caso, se acota la situación a que el fondo de inversión invierte en el mercado de valores en el que estamos invirtiendo nosotros, dejando de lado otras posibles inversiones en otros activos que no son necesariamente valores (derivados, bonos, etc.)

Notemos que si b^* es el portafolio log-óptimo para W^*_{m+1}, entonces W^*_{m+1} es de la forma:

$$\begin{aligned} W^*_{m+1} &= \mathbb{E}\ln(b^*_1X_1 + ... + b^*_mX_m + b^*_{m+1}(b_1X_1 + ... + b_mX_m)) \\ &= \mathbb{E}\ln(((b^*_1 + b^*_{m+1}b_1)X_1 + ... + (b^*_m + b^*_{m+1}b_m)X_m)). \end{aligned}$$

Es decir, W^*_{m+1} es el máximo de todas las esperanzas que tienen como portafolio uno cuyas entradas son de la forma $(b^*_i + b^*_{m+1}b_i)$ con $i = 1, ..., m$.

Para demostrar que $W^*_{m+1} = W^*_m$, basta demostrar que las entradas de cualquier portafolio $\mathbf{c} \in \mathcal{B}_m$ son de la forma $c_i = d_i + d_{m+1}b_i$ con $\mathbf{d} \in \mathcal{B}_{m+1}$. Para ello, sólo hace falta poner $d_i = c_i$ para $i = 1, ..., m$ y $d_{m+1} = 0$. Gracias a esto, podemos afirmar que $W^*_{m+1} \geq W^*_m$, pues W^*_{m+1} es mayor o igual a cada uno de las esperanzas de donde se obtiene el máximo. Para concluir que $W^*_{m+1} = W^*_m$, basta probar que el portafolio relativo a W^*_{m+1} es un realidad uno perteneciente a $\mathcal{B}_m$, para ello, sumemos:

$$\begin{aligned} \sum_{i=1}^{m}(b^*_i + b^*_{m+1}b_i) &= \sum_{i=1}^{m}(b^*_i) + \sum_{i=1}^{m}(b^*_{m+1}b_i) \\ &= 1 - b^*_{m+1} + b^*_{m+1} \\ &= 1. \end{aligned}$$

Además, $(b^*_i + b^*_{m+1}b_i) \geq 0$, por tanto, este portafolio pertenece a $\mathcal{B}_m$ y entonces $W^*_{m+1} = W^*_m$.

4. Información mutua y tasa de crecimiento

En esta sección analizaremos el margen de error que se podría esperar derivado de no estimar de manera adecuada la función de distribución del mercado de valores. Es decir, si no estimamos bien la función de distribución del mercado y asumimos una distribución errónea, obteniendo así un portafolio que no es óptimo, entonces nos alejaremos del óptimo en menos que la entropía relativa entre las dos distribuciones de probabilidad.

TEOREMA 3.1. *Sea* $\mathbf{X}$ *un mercado de valores con densidad* $f(\mathbf{x})$*. Sea* $\mathbf{b}_f$ *un portafolio log-óptimo correspondiente a la densidad* $f(\mathbf{x})$ *y sea* $\mathbf{b}_g$ *un portafolio log-óptimo correspondiente a otra densidad* $g(\mathbf{x})$*. Entonces el incremento en la tasa de crecimiento* ΔW*, usando la densidad incorrecta* g*, está acotado por :*

$$\Delta W = W(\mathbf{b}_f, F) - W(\mathbf{b}_g, F) \leq D(f||g).$$

DEMOSTRACIÓN. Para demostrar la desigualdad, se usarán la desigualdad de Jensen y las condiciones de Kuhn-Tucker.

$$\begin{aligned}
\Delta W &= \int f(\mathbf{x}) \log(\mathbf{b_f x}) d\mathbf{x} - \int f(\mathbf{x}) \log(\mathbf{b_g x}) d\mathbf{x} \\
&= \int f(\mathbf{x}) \log\left(\frac{\mathbf{b_f x}}{\mathbf{b_g x}}\right) d\mathbf{x} \\
&= \int f(\mathbf{x}) \log\left(\frac{\mathbf{b_f x}}{\mathbf{b_g x}} \frac{g(\mathbf{x})}{f(\mathbf{x})} \frac{f(\mathbf{x})}{g(\mathbf{x})}\right) d\mathbf{x} \\
&= \int f(\mathbf{x}) \log\left(\frac{\mathbf{b_f x}}{\mathbf{b_g x}} \frac{g(\mathbf{x})}{f(\mathbf{x})}\right) d\mathbf{x} + \int f(\mathbf{x}) \log\left(\frac{f(\mathbf{x})}{g(\mathbf{x})}\right) d\mathbf{x} \\
&= \int f(\mathbf{x}) \log\left(\frac{\mathbf{b_f x}}{\mathbf{b_g x}} \frac{g(\mathbf{x})}{f(\mathbf{x})}\right) d\mathbf{x} + D(f||g) \\
&\leq \log \int f(\mathbf{x}) \frac{\mathbf{b_f x}}{\mathbf{b_g x}} \frac{g(\mathbf{x})}{f(\mathbf{x})} d\mathbf{x} + D(f||g) \\
&= \log \int g(\mathbf{x}) \frac{\mathbf{b_f x}}{\mathbf{b_g x}} d\mathbf{x} + D(f||g) \\
&= \log\left(\mathbb{E}_g\left(\frac{S}{S^*}\right)\right) + D(f||g) \\
&\leq \log(1) + D(f||g) \\
&= D(f||g).
\end{aligned}$$

□

Ahora veamos un resultado un poco más general. La situación es la siguiente. Supongamos que existe información adicional, representada por la variable aleatoria Y. Esta información puede ser relacionada o no con el mercado de valores. Si no conocemos esa información, pero el resto de los participantes del mercado sí la conoce, entonces estaremos en desventaja. Por lo tanto, el mercado tendrá ahora una distribución que tome en cuenta a la variable Y y nosotros seguiremos creyendo que conserva la distribución original. El resultado nos dice que lo más que nos podemos alejar de la ganancia óptima, sin importar el valor que tome la variable Y, está acotada por la información mutua entre $\mathbf{X}$ y Y. Note que, en el caso en que Y y $\mathbf{X}$ son independientes, la cota se hace cero y no nos alejaremos de la ganancia óptima.

TEOREMA 3.2. *El incremento ΔW en la tasa de crecimiento debido a cualquier otra variable adyacente Y esta acotado por:*

$$\Delta W \leq I(\mathbf{X}, Y).$$

DEMOSTRACIÓN. Consideremos el vector aleatorio $(\mathbf{X}, Y)$ con densidad $f(\mathbf{x}, y)$, donde $\mathbf{X}$ es el mercado de valores y Y es la variable adyacente, que nos da alguna información extra. Dado que $Y = y$, usamos el portafolio log-óptimo para la densidad $f(\mathbf{x}, Y = y)$ y por lo tanto, un inversionista que no esté enterado de la variable aleatoria Y, invertirá con el portafolio óptimo que corresponde a f. Por el teorema anterior, tenemos que:

$$\begin{aligned}
\Delta W_{Y=y} &= W(\mathbf{b}_{f_{Y=y}}, F_{Y=y}) - W(\mathbf{b}_f, F_{Y=y}) \\
&= \int f(\mathbf{x}|Y = y) \log(\mathbf{b}_{\mathbf{f}_{\mathbf{Y}=\mathbf{y}}}\mathbf{x}) d\mathbf{x} - \int f(\mathbf{x}|Y = y) \log(\mathbf{b}_{\mathbf{f}}\mathbf{x}) d\mathbf{x} \\
&\leq D(f(\mathbf{x}|Y = y)||f(\mathbf{x})) \\
&= \int_{\mathbf{x}} f(\mathbf{x}|Y = y) \log\left(\frac{f(\mathbf{x}|Y = y)}{f(\mathbf{x})}\right) d\mathbf{x}.
\end{aligned}$$

Luego, si promediamos sobre todos los valores que puede tomar la variable Y, tenemos que:

$$\begin{aligned}
\Delta W &= \int f(y) \Delta W_{Y=y} dy \\
&\leq \int_y f(y) D(f(\mathbf{x}|Y = y)||f(\mathbf{x})) d\mathbf{x} dy \\
&\leq \int_y f(y) \int_{\mathbf{x}} f(\mathbf{x}|Y = y) \log\left(\frac{f(\mathbf{x}|Y = y)}{f(\mathbf{x})}\right) d\mathbf{x} dy \\
&= \int_y \int_{\mathbf{x}} f(y) f(\mathbf{x}|Y = y) \log\left(\frac{f(\mathbf{x}|Y = y)}{f(\mathbf{x})} \frac{f(y)}{f(y)}\right) d\mathbf{x} dy \\
&= \int_y \int_{\mathbf{x}} f(\mathbf{x}, y) \log\left(\frac{f(\mathbf{x}, y)}{f(\mathbf{x}) f(y)}\right) d\mathbf{x} dy \\
&= I(\mathbf{X}; Y).
\end{aligned}$$

□

Por tanto, el incremento en la tasa de crecimiento, en promedio, sin importar el valor que tome la variable Y, está acotado por la información mutua entre la variable Y y el vector del mercado.

Note que para el caso particular en que $\mathbf{X}$ y Y son independientes, el incremento en la tasa de crecimiento es 0, al igual que la información mutua entre ambas variables, pues $f(\mathbf{x})f(y) = f(\mathbf{x}, y)$.

Veamos que las desigualdades presentadas se cumplen para una distribución presentada previamente.

Ejemplo 3.3. Supongamos que $\mathbf{X} = (1, a)$ con probabilidad $\frac{1}{2}$, y que $\mathbf{X} = \left(1, \frac{1}{a}\right)$ con la misma probabilidad, donde $a > 1$. Supongamos además que se tiene la variable aleatoria $Y = 1$ si $X_1, X_2 \geq 1$ y que $Y = 0$ si $X_1, X_2 \leq 1$. Comprobaremos que

$$\Delta W_{Y=y} \leq D(f(\mathbf{x}|Y = y)||f(\mathbf{x})).$$

Y usando lo anterior, comprobaremos también que:

$$\Delta W \leq I(\mathbf{X} : Y).$$

Recordemos que para esta distribución de $\mathbf{X}$, el portafolio log-óptimo es $b_f = (\frac{1}{2}, \frac{1}{2})$.

Calculemos ahora $f_Y(y)$:

$$\begin{aligned} f_Y(0) &= \mathbb{P}(X_1, X_2 \leq 1) \\ &= \mathbb{P}\left(X_2 = \frac{1}{a}\right) \\ &= \frac{1}{2}. \end{aligned}$$

Similarmente, $f_Y(1) = \frac{1}{2}$.

Ahora encontremos $f_{\mathbf{X}|Y=y}(x_1, x_2)$, para $y = 0, 1$.

$$\begin{aligned} f_{\mathbf{X}|Y=0}(1, a) &= \mathbb{P}((X_1, X_2) = (1, a)|Y = 0) \\ &= \mathbb{P}(X_2 = a|X_2 \leq 1) \\ &= \mathbb{P}\left(X_2 = a|X_2 = \frac{1}{a}\right) \\ &= 0. \end{aligned}$$

Usando el mismo razonamiento y bajo condiciones similares, obtenemos las siguientes igualdades:

$$\begin{aligned} f_{\mathbf{X}|Y=0}\left(1, \frac{1}{a}\right) &= 1 \\ f_{\mathbf{X}|Y=1}(1, a) &= 1 \\ f_{\mathbf{X}|Y=1}\left(1, \frac{1}{a}\right) &= 0. \end{aligned}$$

Usando las expresiones anteriores calculemos la entropía relativa $D(f_{\mathbf{X}|Y=0}||f_{\mathbf{X}})$.

$$\begin{aligned} D(f_{\mathbf{X}|Y=0}||f_{\mathbf{X}}) &= f_{\mathbf{X}|Y=0}(1, a)\log\left(\frac{f_{\mathbf{X}|Y=0}(1, a)}{f_{\mathbf{X}}(1, a)}\right) \\ &+ f_{\mathbf{X}|Y=0}\left(1, \frac{1}{a}\right)\log\left(\frac{f_{\mathbf{X}|Y=0}(1, \frac{1}{a})}{f_{\mathbf{X}}(1, \frac{1}{a})}\right) \\ &= f_{\mathbf{X}|Y=0}\left(1, \frac{1}{a}\right)\log\left(\frac{f_{\mathbf{X}|Y=0}(1, \frac{1}{a})}{f_{\mathbf{X}}(1, \frac{1}{a})}\right) \\ &= \log\left(\frac{1}{\frac{1}{2}}\right) \\ &= 1. \end{aligned}$$

Similarmente, podemos calcular que $D(f_{\mathbf{X}|Y=1}||f_{\mathbf{X}}) = 1$.

Procedamos a calcular $\Delta W_{Y=0}$. Primero recordemos que:

$$\Delta W_{Y=0} = W(b_{f_{\mathbf{X}|Y=0}}, F_{\mathbf{X}|Y=0}) - W(b_{f_{\mathbf{X}}}, F_{\mathbf{X}|Y=0}).$$

Por lo tanto, debemos de encontrar el portafolio log-óptimo para la distribución $b_{f_{\mathbf{X}|Y=0}}$ para la distribución $F_{\mathbf{X}|Y=0}$. Pero recuerde que $f_{\mathbf{X}|Y=0}\left(1, \frac{1}{a}\right) = 1$. Por lo tanto, la distribución condicional a que $Y = 0$ es constante. Por lo tanto, no es difícil ver que el portafolio log-óptimo para esta distribución es $b_{f_{\mathbf{X}|Y=0}} = (1, 0)$, pues si se invirtiera algo en el segundo valor, perderíamos capital con probabilidad 1. Ahora estamos listos para calcular $\Delta W_{Y=0}$. Primero, calculamos:

$$\begin{aligned} W(b_{f_{\mathbf{X}|Y=0}}, F_{\mathbf{X}|Y=0}) &= \mathbb{E}\left[\log\left(b_{f_{\mathbf{X}|Y=0}} \cdot (\mathbf{X}|Y = 0)\right)\right] \\ &= \log\left((1, 0) \cdot \left(1, \frac{1}{a}\right)\right) \\ &= 0. \end{aligned}$$

Y ahora calculamos:

$$\begin{aligned} W(b_{f_{\mathbf{X}}}, F_{\mathbf{X}|Y=0}) &= \mathbb{E}\left[\log\left(b_{f_{\mathbf{X}}} \cdot (\mathbf{X}|Y=0)\right)\right] \\ &= \log\left(\left(\frac{1}{2}, \frac{1}{2}\right) \cdot \left(1, \frac{1}{a}\right)\right) \\ &= \log\left(\frac{1}{2} + \frac{1}{2a}\right) \\ &= \log\left(\frac{a+1}{2a}\right). \end{aligned}$$

Por lo tanto:

$$\begin{aligned} \Delta W_{Y=0} &= W(b_{f_{\mathbf{X}|Y=0}}, F_{\mathbf{X}|Y=0}) - W(b_{f_{\mathbf{X}}}, F_{\mathbf{X}|Y=0}) \\ &= -\log\left(\frac{a+1}{2a}\right) \\ &= \log\left(\frac{2a}{a+1}\right) \\ &< \log\left(\frac{2a+2}{a+1}\right) \\ &= \log(2) \\ &= 1 = D(f_{\mathbf{X}|Y=0}||f_{\mathbf{X}}). \end{aligned}$$

Ahora bien, para el caso en que $Y = 1$, la distribución condicional se vuelve constante e igual a $(1, a)$, por lo que el portafolio log-óptimo es en este caso $(0, 1)$. Por lo tanto:

$$\begin{aligned} W(b_{f_{\mathbf{X}|Y=1}}, F_{\mathbf{X}|Y=1}) &= \mathbb{E}\left[\log\left(b_{f_{\mathbf{X}|Y=1}} \cdot (\mathbf{X}|Y=1)\right)\right] \\ &= \log((0,1) \cdot (1,a)) \\ &= \log(a). \end{aligned}$$

Y similarmente:

$$\begin{aligned} W(b_{f_{\mathbf{X}}}, F_{\mathbf{X}|Y=1}) &= \mathbb{E}\left[\log\left(b_{f_{\mathbf{X}}} \cdot (\mathbf{X}|Y=1)\right)\right] \\ &= \log\left(\left(\frac{1}{2}, \frac{1}{2}\right) \cdot (1,a)\right) \\ &= \log\left(\frac{1}{2} + \frac{a}{2}\right) \\ &= \log\left(\frac{a+1}{2}\right). \end{aligned}$$

Por lo tanto:

$$
\begin{aligned}
\Delta W_{Y=1} &= W(b_{f_{\mathbf{X}|Y=1}}, F_{\mathbf{X}|Y=1}) - W(b_{f_{\mathbf{X}}}, F_{\mathbf{X}|Y=1}) \\
&= \log(a) - \log\left(\frac{a+1}{2}\right) \\
&= \log\left(\frac{a}{\frac{a+1}{2}}\right) \\
&= \log\left(\frac{2a}{a+1}\right) \\
&< \log\left(\frac{2a+2}{a+1}\right) \\
&= \log(2) \\
&= 1 = D(f_{\mathbf{X}|Y=1}||f_{\mathbf{X}}).
\end{aligned}
$$

Calculemos $I(\mathbf{X}, Y)$, antes de alcanzar la conclusión.

$$
\begin{aligned}
I(\mathbf{X}, Y) &= f_{\mathbf{X},Y}(1, a, 1)\log\left(\frac{f_{\mathbf{X},Y}(1, a, 1)}{f_{\mathbf{X}}(1, a)f_Y(1)}\right) \\
&+ f_{\mathbf{X},Y}(1, a, 0)\log\left(\frac{f_{\mathbf{X},Y}(1, a, 0)}{f_{\mathbf{X}}(1, a)f_Y(0)}\right) \\
&+ f_{\mathbf{X},Y}\left(1, \frac{1}{a}, 1\right)\log\left(\frac{f_{\mathbf{X},Y}\left(1, \frac{1}{a}, 1\right)}{f_{\mathbf{X}}\left(1, \frac{1}{a}\right)f_Y(1)}\right) \\
&+ f_{\mathbf{X},Y}\left(1, \frac{1}{a}, 0\right)\log\left(\frac{f_{\mathbf{X},Y}\left(1, \frac{1}{a}, 0\right)}{f_{\mathbf{X}}\left(1, \frac{1}{a}\right)f_Y(0)}\right) \\
&= f_{\mathbf{X},Y}(1, a, 1)\log\left(\frac{f_{\mathbf{X},Y}(1, a, 1)}{f_{\mathbf{X}}(1, a)f_Y(1)}\right) \\
&+ f_{\mathbf{X},Y}\left(1, \frac{1}{a}, 0\right)\log\left(\frac{f_{\mathbf{X},Y}\left(1, \frac{1}{a}, 0\right)}{f_{\mathbf{X}}\left(1, \frac{1}{a}\right)f_Y(0)}\right) \\
&= \frac{1}{2}\log\left(\frac{\frac{1}{2}}{\frac{1}{2}\frac{1}{2}}\right) + \frac{1}{2}\log\left(\frac{\frac{1}{2}}{\frac{1}{2}\frac{1}{2}}\right) \\
&= 1.
\end{aligned}
$$

Finalmente:

$$\begin{aligned}\Delta W &= f_Y(0)\Delta W_{Y=0} + f_Y(1)\Delta W_{Y=1} \\ &= \frac{1}{2}\left(\log\left(\frac{2a}{a+1}\right)\right) + \frac{1}{2}\left(\log\left(\frac{2a}{a+1}\right)\right) \\ &= \log\left(\frac{2a}{a+1}\right) \\ &< 1 = I(\mathbf{X} : Y).\end{aligned}$$

5. Inversión en mercados estacionarios

Ahora extendemos los resultados de las secciones anteriores sobre mercados independientes e idénticamente distribuidos a mercados que dependen del tiempo. Sea $\mathbf{X_1}, ..., \mathbf{X_n}, ...$ un proceso de vectores aleatorios con $\mathbf{X_i} \geq 0$. En esta sección haremos uso de la definición de estrategia de inversión causal. Es decir, consideramos estrategias de inversión que dependen de los valores pasados del mercado de una manera causal, o equivalentemente, que $\mathbf{b_i}$ puede depender de $\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_{i-1}}$. Sea

$$S_n = \prod_{i=1}^{n} \mathbf{b_i}(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_{i-1}})\mathbf{X_i}.$$

Nuestro objetivo es, naturalmente, maximizar $E(\log(S_n))$, haciendo variar las estrategias causales $\{\mathbf{b_i}(\cdot)\}$. Analicemos entonces el problema:

$$\begin{aligned}\max_{\mathbf{b_1},...,\mathbf{b_n}} E(\log(S_n)) &= \sum_{i=1}^{n} \max_{\mathbf{b_i}(\mathbf{X_1},...,\mathbf{X_{i-1}})} E(\log(\mathbf{b_i^t}\mathbf{X_i})) \\ &= \sum_{i=1}^{n} E(\log(\mathbf{b_i^*}^{\mathbf{t}}\mathbf{X_i})).\end{aligned}$$

donde $\mathbf{b_i^*}$ es el portafolio log-óptimo para la distribución de $\mathbf{X_i}$ dadas las variables $\mathbf{X_1}, ..., \mathbf{X_{i-1}}$, esto es, $\mathbf{b_i^*}(\mathbf{x_1}, ..., \mathbf{x_{i-1}})$ es el portafolio que alcanza el máximo condicional, que se denota por:

$$\begin{aligned}&\max_{\mathbf{b}} E(\log(\mathbf{b^t}\mathbf{X_i}|(\mathbf{X_1}, ..., \mathbf{X_{i-1}}) = (\mathbf{x_1}, ..., \mathbf{x_{i-1}}))) \\ &\qquad\qquad = W^*(\mathbf{X_i}|\mathbf{x_1}, ..., \mathbf{x_{i-1}}).\end{aligned}$$

Es de destacar que, intuitivamente, la suma puede separar el máximo pues la manera óptima de invertir durante n días es invertir en cada día usando el portafolio log-óptimo correspondiente.

Si en la ecuación anterior tomamos la esperanza con respecto al vector aleatorio que condiciona, tenemos:

$$W^*(\mathbf{X_i}|\mathbf{X_1}, ..., \mathbf{X_{i-1}}) = E(\max_{\mathbf{b}} E(\log(\mathbf{b^t}\mathbf{X_i}|(\mathbf{X_1}, ..., \mathbf{X_{i-1}}))),$$

que es la tasa de crecimiento condicionada, donde el máximo es sobre todas las funciones valuadas en portafolios definidos sobre $\mathbf{X_1}, ..., \mathbf{X_{i-1}}$.

Por tanto, de acuerdo a la última ecuación, el capital máximo esperado se alcanza usando el portafolio log-óptimo condicionado correspondiente a cada etapa. Sea

$$W^*(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n}) = \max_{\mathbf{b_1}, \mathbf{b_2}, ..., \mathbf{b_n}} \mathbb{E}(\log(S_n)).$$

Donde el máximo es sobre todas las estrategias de portafolios causales. Entonces, como $S_n^* = \sum_{i=1}^n \log(\mathbf{b_i^*}^{\mathbf{t}} \mathbf{X_i})$, tenemos el siguiente resultado, que se conoce como regla de la cadena para W^*:

$$W^*(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n}) = \sum_{i=1}^n W^*(\mathbf{X_i} | \mathbf{X_1}, ..., \mathbf{X_{i-1}}).$$

Definición 3.1. La tasa de crecimiento W_∞^* se define como:

$$W*_\infty = \lim_{n\to\infty} \frac{W^*(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n})}{n},$$

si el límite existe.

Veamos ahora un pequeño lema.

Lema 3.2. $W^*(\mathbf{X_n} | \mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_{n-1}})$ *es no decreciente en* n *cuando el mercado es estacionario.*

La definición de estacionariedad aquí se refiere a la presentada en la página 13.

Intuitivamente, gracias a que el mercado es estacionario, si aumenta n, conocemos más valores anteriores del mercado, y como el valor actual del mercado depende de los anteriores, poseemos más información, y por lo tanto, tenemos más oportunidad de ampliar nuestras ganancias.

Demostración. Por la estacionariedad del proceso, se tiene

$$W^*(\mathbf{X_n} | \mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_{n-1}}) = W^*(\mathbf{X_{n+1}} | \mathbf{X_2}, \mathbf{X_3}, ..., \mathbf{X_n}).$$

Y gracias a que el condicionamiento aumenta W^*, tenemos

$$W^*(\mathbf{X_{n+1}} | \mathbf{X_2}, \mathbf{X_3}, ..., \mathbf{X_n}) \leq W^*(\mathbf{X_{n+1}} | \mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n}).$$

□

Usemos el lema anterior para demostrar el siguiente teorema, que tiene su análogo para cuando hablamos de la entropía de procesos estocásticos estacionarios, como se puede consultar en la página 13 de este trabajo.

TEOREMA 3.3. *Para un mercado estacionario, la tasa de crecimiento $W*_\infty$ existe y es igual a:*

$$W*_\infty = \lim_{n\to\infty} W^*(\mathbf{X_n}|\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_{n-1}}).$$

DEMOSTRACIÓN. Por el lema, $W^*(\mathbf{X_n}|\mathbf{X_2}, \mathbf{X_3}, ..., \mathbf{X_{n-1}})$ es una sucesión no decreciente de números reales, por lo que si está acotada, converge a un número real, y si no está acotada, converge a infinito. En cualquier caso, tiene un límite, que puede ser infinito. Ahora bien, como:

$$\frac{W^*(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n})}{n} = \frac{1}{n}\sum_{i=1}^{n} W^*(\mathbf{X_i}|\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_{i-1}}),$$

podemos usar el teorema de convergencia de Cesáro, que indica que el límite de una sucesión coincide con el límite de las medias de una sucesión, de la siguiente manera:

$$\begin{aligned} W^*_\infty &= \lim_{n\to\infty} \frac{W^*(\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n})}{n} \\ &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} W^*(\mathbf{X_i}|\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_{i-1}}) \\ &= \lim_{n\to\infty} W^*(\mathbf{X_n}|\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_{n-1}}). \end{aligned}$$

$\square$

Ahora extendemos la propiedad de optimalidad asintótica para mercados estacionarios.

TEOREMA 3.4. *Considérese cualquier proceso estocástico $\{X_i\}$, $X_i \in \mathbb{R}^m_+$, portafolios log-óptimos $\mathbf{b}^*_i(X^{i-1})$ y el capital S^*_n. Sea S_n el capital generado por cualquier otra estrategia de portafolios causal $\mathbf{b}_i(X_i)$. Entonces S_n/S^*_n es una supermartingala con respecto a la filtración canónica $(\sigma(X_1, ..., X_n))_{n\in\mathbb{N}}$. Consecuentemente, existe una variable aleatoria V tal que*

- $\dfrac{S_n}{S^*_n} \to V$ *con probabilidad 1,*
- $E(V) \leq 1$.

Note que el hecho de que el proceso aleatorio sea una supermartingala concuerda con la interpretación que dimos en el segundo capítulo, donde indicamos que una supermartingala se interpreta como un juego desfavorable para el jugador. En este caso ocurre igual, porque el numerador es el capital relativo que resulta de usar una estrategia que no es óptima, haciendo que el cociente se reduzca.

DEMOSTRACIÓN. Para simplificar la notación, diremos que $X^n = (X_1, X_2, ..., X_n)$. $\frac{S_n}{S_n^*}$ es una supermartingala porque

$$\begin{aligned}\mathbb{E}\left[\frac{S_{n+1}(X^{n+1})}{S^*_{n+1}(X^{n+1})}\Big| X^n\right] &= \mathbb{E}\left[\frac{(\mathbf{b_{n+1}X_{n+1}})S_n(X^n)}{(\mathbf{b^*_{n+1}X_{n+1}})S^*_n(X^n)}\Big| X^n\right] \\ &= \frac{S_n(X^n)}{S^*_n(X^n)}\mathbb{E}\left[\frac{(\mathbf{b^t_{n+1}X_{n+1}})}{(\mathbf{b^{*t}_{n+1}X_{n+1}})}\Big| X^n\right] \\ &\leq \frac{S_n(X^n)}{S^*_n(X^n)},\end{aligned}$$

por las condiciones de Kuhn-Tucker sobre los portafolios log-óptimos.

Entonces, usando el teorema de convergencia para martingalas, en particular supermartingalas, tenemos que $\frac{S_n(X^n)}{S_n^*(X^n)}$ tiene límite. Llamemos al límite V. También tenemos que $\mathbb{E}(V) \leq \mathbb{E}(\frac{S_1(X^0)}{S_1^*(X^0)}) = 1$. Podemos usar el teorema de convergencia para martingalas porque para cualquier n tenemos que $\mathbb{E}(\frac{S_n(X^n)}{S_n^*(X^n)}) \leq \mathbb{E}(\frac{S_1(X^0)}{S_1^*(X^0)}) = 1 \ < \infty$.

□

6. Optimalidad competitiva del portafolio log-óptimo

En esta sección analizaremos la probabilidad de que el inversionista con cualquier portafolio obtenga mayor ganancia, en términos del capital invertido, que un inversionista con un portafolio log-óptimo. En nuestro análisis previo, habíamos abordado esta pregunta en términos del valor esperado del cociente de los capitales relativos. Ahora nos enfocaremos en encontrar cotas para la probabilidad mencionada arriba.

Por los resultados que se derivan de las condiciones de Kuhn-Tucker tenemos:

$$\mathbb{E}\left(\frac{S_n}{S_n^*}\right) \leq 1,$$

y entonces por la desigualdad de Markov tenemos:

$$\mathbb{P}(S_n > tS_n^*) \leq \frac{1}{t}.$$

Veamos un ejemplo que pondrá de manifiesto el hecho de que no se puede obtener una cota mejor que 1 para la probabilidad de que $S_n > S_n^*$

Consideremos un mercado de valores con dos acciones y dos posibilidades:

$$(X_1, X_2) = \begin{cases} \left(1, \frac{1}{1-\epsilon}\right) & \text{con probabilidad } 1-\epsilon. \\ (1,0) & \text{con probabilidad } \epsilon. \end{cases}$$

En este mercado, el portafolio log-óptimo es $\mathbf{b} = (1, 0)$ pues cumple con las condiciones de Khun-Tucker. Puede comprobarse que:

$$\mathbb{E}\left(\frac{X_1}{\mathbf{b}\cdot X}\right)=\mathbb{E}\left(\frac{X_2}{\mathbf{b}\cdot X}\right)=1.$$

Al analizar el portafolio log-óptimo, nos damos cuenta de que debemos invertir todo el capital en la primera acción. Sin embargo, un inversionista que invierte todo su dinero a la segunda acción ganará más dinero con probabilidad $1-\epsilon$, debido a que:

$$\begin{aligned}\mathbb{P}\left(S>S^*\right)&=\mathbb{P}\left(b_1+b_2X_2>1\right)\\&=\mathbb{P}\left(b_2X_2>b_2\right)\\&=\mathbb{P}\left(X_2>1\right)\\&=\mathbb{P}\left(X_2=\frac{1}{1-\epsilon}\right)\\&=1-\epsilon.\end{aligned}$$

Y como podemos tomar ϵ tan cercana a 0 como queramos, no es posible encontrar una cota para la probabilidad que acabamos de calcular.

Podemos evitar este tipo de situaciones al poner ciertas condiciones adicionales. En este caso, añadiremos un factor adicional que no habíamos considerado previamente: el capital de los inversionistas (distinto al capital relativo). Si permitimos que ambos inversionistas «apuesten» una cantidad de dinero con distribución uniforme en un mismo intervalo, entonces podemos encontrar una cota informativa para la probabilidad de que el inversionista con un portafolio cualquiera gane más que el inversionista que utiliza el portafolio log-óptimo.

TEOREMA 3.1. *Sea S^* el capital al final de un periodo de inversión en un mercado de valores $\mathbf{X}$ con el portafolio log-óptimo, y sea S el capital obtenido por cualquier otro portafolio. Sea U^* una variable aleatoria independiente de $\mathbf{X}$ que se distribuye uniformemente en $[0,2]$ y sea V otra variable aleatoria independiente de $\mathbf{X}$ y de U^* con $V\geq 0$ y $E(V)=1$. Entonces*

$$\mathbb{P}(VS\geq U^*S^*)\leq\frac{1}{2}.$$

Note que el enunciado del teorema incluye las condiciones que comentábamos en el párrafo previo, de hecho las generaliza. El capital en juego del inversionista con un portafolio cualquiera (V) puede seguir cualquier distribución, puede incluso tomar un valor muy grande, únicamente debe tener esperanza 1, como es el caso de una variable aleatoria uniforme en el intervalo $[0,2]$. El capital del inversionista con portafolio log-óptimo (U^*) es más restringido, únicamente puede tomar valores entre $[0,2]$. Note que aún en estas condiciones, es posible encontrar una cota informativa.

DEMOSTRACIÓN. Tenemos:

$$\begin{aligned}\mathbb{P}(VS \geq U^*S^*) &= \mathbb{P}\left(\frac{VS}{S^*} \geq U^*\right) \\ &= \mathbb{P}(W \geq U^*),\end{aligned}$$

donde $W = \frac{VS}{S^*}$ es una variable aleatoria no negativa y con esperanza:

$$E(W) = E(V)E\left(\frac{S_n}{S_n^*}\right) \leq 1,$$

por la independencia de V y $\mathbf{X}$ y por las condiciones de Khun-Tucker. Sea F la distribución de W. Como U^* es uniforme en $[0,2]$,

$$\begin{aligned}\mathbb{P}(W \geq U^*) &= \int_0^2 \mathbb{P}(W \geq w) f_U(w) dw \\ &= \int_0^2 \mathbb{P}(W \geq w)\frac{1}{2} dw \\ &= \int_0^2 \frac{1-F(w)}{2} dw \\ &\leq \int_0^\infty \frac{1-F(w)}{2} \\ &= \frac{1}{2}E(W) \\ &\leq \frac{1}{2}.\end{aligned}$$

Por lo tanto, tenemos que

$$\mathbb{P}(VS \geq U^*S^*) \leq \frac{1}{2}.$$

□

Comprobemos nuestro resultado con el mercado de valores presentado al inicio de esta sección, supongamos para ello que V y U^* tienen distribución uniforme en el intervalo $[0,2]$. Entonces:

$$\begin{aligned}\mathbb{P}\left(VS \geq U^*S^*\right) &= \mathbb{P}\left(V((b_1,b_2)\cdot(X_1,X_2)) \geq U^*((b_1^*,b_2^*)\cdot(X_1,X_2))\right) \\ &= \mathbb{P}\left(b_1V + b_2VX_2 \geq U^*\right) \\ &= \mathbb{P}\left(b_1V + b_2VX_2 \geq U^* | X_2 = 0\right)\mathbb{P}\left(X_2 = 0\right) \\ &+ \mathbb{P}\left(b_1V + b_2VX_2 \geq U^* | X_2 = \frac{1}{1-\epsilon}\right)\mathbb{P}\left(X_2 = \frac{1}{1-\epsilon}\right) \\ &= \epsilon\mathbb{P}\left(b_1V \geq U^*\right) \\ &+ (1-\epsilon)\mathbb{P}\left(b_1V + b_2\left(\frac{1}{1-\epsilon}\right)VX_2 \geq U^*\right).\end{aligned}$$

Para calcular las probabilidades que arroja el cálculo, usaremos la función de densidad de la resta de dos variables aleatorias. Para la primera probabilidad, consideremos $X = b_1 V$ y $Y = U^*$. X es una variable aleatoria con distribución uniforme en $[0, 2b_1]$ y $2b_1 < 2$.

$$\begin{aligned}
\mathbb{P}\left(b_1 V \geq U^*\right) &= \mathbb{P}\left(b_1 V - U^* \geq 0\right) \\
&= \mathbb{P}\left(X - Y \geq 0\right) \\
&= \int_0^\infty f_{X-Y}(u)du \\
&= \int_0^\infty \left(\int_{-\infty}^\infty f_{X,Y}(u+v, v)dv\right) du \\
&= \int_0^\infty \int_{-\infty}^\infty f_X(u+v) f_Y(v) dv du.
\end{aligned}$$

Para calcular los límites de integración en donde las funciones de densidad son ambas positivas, necesitamos que $u + v \in [0, 2b_1]$ y que $v \in [0, 2]$, además de la condición que ya tenemos $u \in [0, \infty)$. Estas tres condiciones nos determinan la siguiente integral.

$$\begin{aligned}
\int_0^\infty \int_{-\infty}^\infty f_X(u+v) f_Y(v) dv du &= \int_0^{2b_1} \int_0^{2b_1 - u} \frac{1}{2b_1}\frac{1}{2} dv du \\
&= \frac{1}{4b_1} \int_0^{2b_1} \int_0^{2b_1 - u} dv du \\
&= \frac{1}{4b_1} \int_0^{2b_1} (2b_1 - u) du \\
&= \frac{1}{4b_1} \left(\frac{(2b_1)^2}{2}\right) \\
&= \frac{b_1}{2}.
\end{aligned}$$

Similarmente, puede calcularse la otra probabilidad, para el caso en que $\epsilon < 2$, y se obtiene que

$$\begin{aligned}
\mathbb{P}\left(b_1 V + b_2 \left(\frac{1}{1-\epsilon}\right) V X_2 \geq U^*\right) &= \frac{1-\epsilon}{4(1-\epsilon b_1)} \left(\frac{\left(2\frac{1-\epsilon}{1-\epsilon b_1}\right)^2}{2}\right) \\
&= \frac{1}{2}\left(\frac{1-\epsilon}{1-\epsilon b_1}\right).
\end{aligned}$$

Por lo que la probabilidad que nos interesa es

$$\begin{aligned}\mathbb{P}\left(VS \geq U^*S^*\right) &= \epsilon\left(\frac{b_1}{2}\right) + (1-\epsilon)\left(\frac{1}{2}\left(\frac{1-\epsilon}{1-\epsilon b_1}\right)\right) \\ &= \frac{\epsilon b_1}{2} + \frac{(1-\epsilon)^2}{2(1-\epsilon b_1)}.\end{aligned}$$

Demostremos que esta última expresión es menor o igual a $\frac{1}{2}$. Para ello, notemos que:

$$\begin{aligned} & \frac{\epsilon b_1}{2} + \frac{(1-\epsilon)^2}{2(1-\epsilon b_1)} \leq \frac{1}{2} \\ \Leftrightarrow \quad & \epsilon b_1 + \frac{(1-\epsilon)^2}{1-\epsilon b_1} \leq 1 \\ \Leftrightarrow \quad & (1-\epsilon)^2 \leq (1-\epsilon b_1)^2 \\ \Leftrightarrow \quad & (1-\epsilon) \leq (1-\epsilon b_1) \\ \Leftrightarrow \quad & \epsilon b_1 \leq \epsilon \\ \Leftrightarrow \quad & b_1 \leq 1. \end{aligned}$$

La última expresión sabemos que es cierta, pues b_1 es una entrada de un portafolio. Por tanto, hemos comprobado la validez del teorema en este ejemplo particular.

Para finalizar este trabajo, analizaremos una manera distinta de definir el portafolio log-óptimo. Veremos que nuestra definición original implica nuestra nueva definición.

Supongamos que se define el portafolio log-óptimo $\mathbf{b}^*$ como el portafolio que maximiza la expresión

$$\int \ln\left(\frac{\mathbf{bx}}{\frac{1}{m}\sum_{i=1}^m x_i}\right) dF(x_1, ..., x_m).$$

- Mostraremos que si $\mathbf{b}$ maximiza $\int \ln(\mathbf{bx})dF(x)$, entonces maximiza $\int \ln\left(\frac{\mathbf{bx}}{u^t x}\right) dF(x)$, donde $u = \left(\frac{1}{m}, \frac{1}{m}, ..., \frac{1}{m}\right)$. Por lo que la definición original es más fuerte que la que estamos proponiendo.
- Encontraremos el portafolio log-óptimo $\mathbf{b}^*$ bajo esta nueva definición para el mercado de valores: $\mathbf{X} = (2^{2^k+1}, 2^{2^k})$ con probabilidad $2^{-(k+1)}$ y $\mathbf{X} = (2^{2^k}, 2^{2^k+1})$ con la misma probabilidad. Con $k = 1, 2, 3, ...$

Supongamos que $\mathbf{b}$ es el portafolio log-óptimo definido de la manera usual y que $\mathbf{c}$ es cualquier otro portafolio, entonces:

$$\int (\ln(\mathbf{bx}))\, dF(x) \geq \int (\ln(\mathbf{cx}))\, dF(x),$$

lo cual implica que:

$$
\begin{aligned}
&\Rightarrow \int (\ln(\mathbf{bx}))\, dF(x) - \int (\ln(\mathbf{cx}))\, dF(x) \geq 0 \\
&\Rightarrow \int (\ln(\mathbf{bx}) - \ln(\mathbf{cx}))\, dF(x) \geq 0 \\
&\Rightarrow \int \ln\left(\frac{\mathbf{bx}}{\mathbf{cx}}\right) dF(x) \geq 0 \\
&\Rightarrow \int \ln\left(\frac{\left(\frac{\mathbf{bx}}{ux}\right)}{\left(\frac{\mathbf{cx}}{ux}\right)}\right) dF(x) \geq 0 \\
&\Rightarrow \int \left(\ln\left(\frac{\mathbf{bx}}{ux}\right) - \ln\left(\frac{\mathbf{cx}}{ux}\right)\right) dF(x) \geq 0 \\
&\Rightarrow \int \ln\left(\frac{\mathbf{bx}}{ux}\right) dF(x) \geq \int \ln\left(\frac{\mathbf{cx}}{ux}\right) dF(x).
\end{aligned}
$$

Por lo tanto, **b** maximiza la otra expresión también.

Para el segundo punto, primero notemos que efectivamente estamos tratando con una medida de probabilidad, pues:

$$\sum_x p(x) = 2\sum_{k=1}^{\infty} 2^{-(k+1)} = \sum_{k=1}^{\infty} 2^{-k} = 1.$$

Ahora calculemos la tasa de crecimiento para un portafolio en general. Utilizando nuestra nueva definición de portafolio óptimo, consideremos la expresión:

$$\mathbb{E}\left(\ln\left(\frac{b_1X_1 + b_2X_2}{\frac{1}{2}(X_1 + X_2)}\right)\right),$$

después de simplificar, la expresión queda como:

$$\frac{1}{2}\ln\left(\frac{4}{3} - \frac{2}{3}b_2\right) + \frac{1}{2}\ln\left(\frac{2}{3} + \frac{2}{3}b_2\right).$$

Llamemos a esta última expresión $f(b_2)$. Como hemos hecho a lo largo de los ejemplos, el problema se reduce a maximizar esta función. Para ello, debemos derivar e igualar a cero la derivada.

$$\begin{aligned}
f'(b_2) &= \frac{1}{2}\frac{-\frac{2}{3}}{(\frac{4}{3}-\frac{2}{3}b_2)\ln(2)} + \frac{1}{2}\frac{\frac{2}{3}}{(\frac{2}{3}+\frac{2}{3}b_2)\ln(2)} \\
&= -\frac{2}{3}(\frac{2}{3}+\frac{2}{3}b_2) + \frac{2}{3}(\frac{4}{3}-\frac{2}{3}b_2) \\
&= -2(2+2b_2) + 2(4-2b_2) \\
&= -(1+b_2) + (2-b_2) \\
&= -1-b_2+2-b_2 \\
&= 1-2b_2.
\end{aligned}$$

igualando a cero se obtiene $b_2 = \frac{1}{2}$. Este punto produce un máximo, de acuerdo a la forma de la función en cuestión, por tanto, el portafolio log-óptimo es $\mathbf{b} = (\frac{1}{2}, \frac{1}{2})$.

7. Preguntas

A raíz del estudio realizado, surgen varias preguntas que vale la pena hacerse y que podrían indicar líneas de trabajo para profundizar en estos temas:

- Si bien el estudio realizado analiza la función logaritmo del capital relativo $\log(S)$, creemos que otros estudios podrían tratar de generalizar los resultados para cualquier función $\phi(S)$ que cumpla, al menos, con las siguientes características: (i) la función es creciente, (ii) la función es continua y (iii) la función posee la propiedad de concavidad o convexidad. Particularmente, el estudio podría ser similar si se usaran funciones como $\log(1+S)$, $\exp(S)$ o incluso S. El problema concreto sería generalizar los resultados de este trabajo para cierto conjunto de funciones, entre las que se encuentre $\log(S)$.

- El estudio se ha centrado en un portafolio de acciones, pero, ¿puede el mismo análisis extenderse a un portafolio con acciones, bonos, contratos derivados y otros productos financieros? ¿De no ser así, cómo se abordaría el problema? Por ejemplo, en el caso de productos derivados, la pérdida de capital no estaría acotada. En el caso de acciones, uno puede perder como máximo el dinero que invierta, pero con los derivados, las pérdidas no están acotadas.

- ¿Cómo se haría el análisis si permitiésemos que el capital que invirtiésemos fuese negativo? es decir, que permitiésemos la venta de acciones. Esto incluye contemplar la hipótesis del llamado *short selling*, en el que un inversionista vende acciones sin tenerlas, entregando las mismas en una fecha posterior, para poder aprovechar el precio disponible el día de la concertación de la operación. En la realidad operativa, la mayoría de las operaciones accionarias concertadas en bolsas de valores tienen un periodo de liquidación de 2 días hábiles, por lo que permitir el *short selling* es una suposición bastante plausible.

- En general, ¿cómo cambiarían los resultados si cambiáramos o debilitáramos algunas hipótesis? Por ejemplo, conocer más información de la distribución del mercado de valores. En gran parte del análisis no se hizo suposición sobre dicha distribución. Otro ejemplo es considerar variar el grado de dependencia entre los resultados del mercado de valores día con día. Hemos supuesto estacionariedad en una ocasión, pero podríamos suponer que hay dependencia solo para cuando se consideran m resultados del mercado seguidos (m-dependencia). También podemos considerar distintos supuestos para calcular una cota para $\mathbb{P}(S_n > tS_n^*) \leq \frac{1}{t}$ que no se concentre en el capital de los inversionistas, sino en restricciones al portafolio de inversión.

- ¿Cómo se puede introducir en el modelo el riesgo de crédito de contraparte? Es decir, cómo introducir en el modelo la probabilidad de que una de las contrapartes de la operación no cumpla con sus obligaciones derivadas de la operación en tiempo y forma. Esta es una de las más grandes deficiencias del modelo. Una sugerencia es utilizar una variable aleatoria adicional, que represente la probabilidad de incumplimiento de la contraparte central que preste servicios al mercado de valores.

- Como elementos adicionales a considerar, se pueden incluir costos transaccionales, y costos de oportunidad en relación con la inversión en otros mercados.

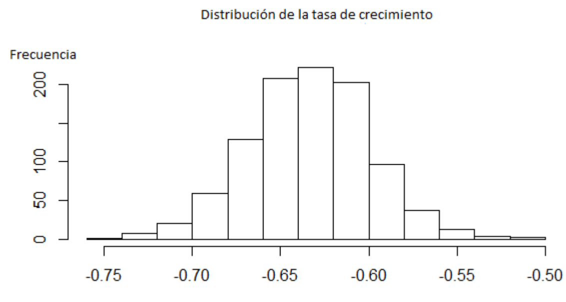
- Finalmente, cabe resaltar que los ejemplos presentados a lo largo de este trabajo corresponden a situaciones muy simplificadas. Por ejemplo, se consideraron acciones con únicamente dos posibles valores. Además, para una $F(x)$ dada, el problema resulta muy difícil de resolver. Creemos que podrían usarse métodos de simulación para tener alguna idea aproximada de la estrategia óptima.

7.1. Problemas de implementación. En este apartado, se muestran los obstáculos a los que nos enfrentríamos en caso de que tuviéramos un mercado con una distribución medianamente conocida y quisiéramos calcular la tasa de crecimiento. De entrada, una pregunta legítima es preguntarse sobre el valor de la tasa de crecimiento cuando el las acciones en el mercado siguen una distribución normal. No obstante, esto no es posible, pues los valores del mercado no pueden ser negativos, ya que representan cocientes de precios. Una manera de darle la vuelta a este inconveniente es suponer que siguen una distribución representada por el valor absoluto de una variable normal estándar. Aun en este caso, para calcular la tasa de cfecimiento, necesitamos saber el grado de dependencia entre las acciones. En el caso má sencillo, cuando todas los valores son independientes entre sí, ya hemos demostrado que cualquier portafolio es log-óptimo. Sin embargo, en ese caso tampoco es trivial calcular el valor de la tasa de crecimiento.

Para facilitar los cálculos, tomemos el portafolio que invierte todo en el primer valor. En este caso, la tasa de crecimiento viene dada por la siguiente expresión:

$$\begin{aligned}\mathbb{E}\left(\log(b \cdot X)\right) &= \int_0^\infty f(x_1, ..., x_n) \log(x_1) \\ &= \int_0^\infty 2\frac{e^{-\frac{x^2}{2}}}{\sqrt{2}} \log(x) dx\end{aligned}$$

Como puede apreciarse, esta no es integral fácil de resolver. El método más apropiado parece ser el de integración por partes, aunque no hemos podido encontrar una solución. No obstante, estimamos esta cantidad mediante un método Monte-Carlo. Obtuvimos un vector de valores estimados para dicha esperanza de la siguiente manera: dada una muestra de valores de una distribución normal, y después de aplicar la correspondiente transformación de valor absoluto y el logaritmo, obtenemos una estimación de la esperanza. Repetimos este proceso y obtenemos distintos valores para la esperanza, cuyo histograma se muestra abajo.



El promedio de estos valores es de $-0{,}63$. Para darnos una idea, esto indica que el capital relativo tendría un valor de aproximadamente $e^{-0{,}63} = 0{,}53$, lo cual indica que nuestro capital disminuiría en un 50 por ciento.

Bibliografía

[1] Thomas M. Cover, Joy A. Thomas. *Elements of information theory.* Wiley, Hoboken, New Jersey, 2006.

[2] Sheldon M. Ross. *A first course in probability.* Pearson, Upper Saddle River, New Jersey, 2010.

[3] N. Gregory Mankiw. *Principles of macroeconomics.* Cengage, Boston, Massachusetts, 2017.

[4] Rincón S. Luis. *Curso intermedio de probabilidad.* Las Prensas de Ciencias, Ciudad de México, México, 2010.