Una variación del algoritmo LSM: discretizaciónes asociadas a sistemas Haar

Agustín H. Arias

Tesis de Grado, Licenciatura en Economía

Universidad Nacional de Mar del Plata
Facultad de Ciencias Económicas y Sociales
Mar del Plata, febrero 2007

Una variación del algoritmo LSM: discretizaciónes asociadas a sistemas Haar

Agustín H. Arias

Director: Dr. Paulino Eugenio Mallo

Comité Evaluador:
Lic. Ricardo Panza
Cr. José A. Castro

Resumen El algoritmo LSM, introducido por Longstaff y Schwartz, es un simple aunque eficiente método para la aproximación del valor de una opción americana vía simulación. EL algoritmo en cuestión implica dos tipos de aproximaciones, la primera, reemplazar las esperanzas condicionales por proyecciones sobre un conjunto finito de funciones; la segunda, utilizar simulación y mínimos cuadrados para computar el valor de la función en la primera aproximación. En el siguiente trabajo se presenta una variación sobre dicho algoritmo. Se realiza una discretización del espacio de probabilidad subyacente al instrumento financiero, asociado al proceso de precios relevante, para luego generar un sistema H, el cual nos permitirá aproximar las esperanzas condicionales de los flujos futuros de fondos descontados, generados por la opción, respecto de la σ-álgebra inducida por el proceso de precios del subyacente. Es decir, creamos un conjunto particular de funciones donde proyectaremos la esperanza condicional, lo que nos permitirá, entre otras cosas, deshacernos de la condición (necesaria en el LSM) de que el proceso sea markoviano.

Palabras clave: Opciones Americanas-Precio-LSM-Haar systems-Discretización espacio temporal.

Abstract The LSM algorithm, due to Longstaff y Schwartz, is a simple yet effective method for approximating the value of American options by simulation. This algorithm involves two types of approximation, the first one, replace the conditional expectations by projections on a finite set of functions; the second one, use simulation and least squares regression to compute the value function of the first approximation. The following article presents a variation on this algorithm. We propose a discretization of the probability space underlying the financial derivative, representative of the price process, so we can construct an H-system, which will allow us to approximate the conditional expectations of the discounted future cash flows, generated by the option, conditional on the σ-algebra induced by the underlying assets price process. That is, we construct a particular set of functions on which we project the conditional expectation, this will permit us, among other things, to get rid of the markovian condition of the process (necessary in the LSM).

Key words: American Options-Price-LSM-Haar wavelets systems-Space-time discretization.

1. Introducción

Tema general

La valoración de opciones de tipo americano dependientes de varios stocks se yergue como uno de los tópicos fundamentales dentro de la teoría de finanzas. Es, obviamente, esencial para la existencia de un mercado y es lo que permite a los participantes elaborar sus estrategias. Entre esos participantes encontramos a los hedgers, caracterizados por el diseño de estrategias de cobertura, es decir, planes de acción para evitar el riesgo inherente a la fluctuación de los precios de los subyacentes (para una profundización de este tema, así como una introducción a los conceptos fundamentales de estos mercados recomendamos al lector recurrir a [1]). Ahora bien, es la complejidad de las ecuaciones diferenciales asociadas a este problema (precificación de las opciones de tipo americano), así como la impracticabilidad computacional de las técnicas binomiales y de diferencias finitas, las que plantean la necesidad de buscar nuevas sendas para la consecución de una solución. De esta manera, F.A. Longstaff y E.S. Schwartz [6], en el año 2000 propusieron una forma efectiva de aplicar el método de Montecarlo para el cálculo del tiempo de parada óptimo para una opción americana y a partir de ello establecer su precio. La ventaja que ofrece la simulación es que permite a las variables estado seguir procesos estocásticos generales. Como punto de partida de este -y de otras técnicas basadas en simulación- encontramos la discretización del periodo continuo de ejercicio (definición del tipo americano) por un conjunto finito de fechas (i.e. la aproximación se plantea mediante una opción de tipo Bermuda). Existen diversos antecedentes de esta metodología, entre otros encontramos los trabajos de J. Carriere [11] y de J. Tsitsiklis & B. van Roy [12].

Tesis general

El algoritmo LSM, ha sido y sigue siendo estudiado en todos los centros de finanzas, así como también es utilizado por los operadores del mercado. Esto nos da una idea de la importancia y la novedad de los conceptos que introduce este algoritmo. Una simple búsqueda en internet nos arroja cientos de artículos sobre explicaciones, mejorías y aplicaciones del mismo. Una de sus principales limitaciones, es la necesidad de trabajar con stocks que presentan una dinámica markoviana.
En el año 2004, P.J. Catuogno, S.E. Ferrando y A.L. González [5] introdujeron una nueva discretización para el espacio de probabilidad que gobierna la evolución de instrumentos financieros en términos de una expansión martingala construida mediante la utilización de Haar Wavelets systems. Es a partir de una propuesta suya, que se abordó este trabajo: el estudio del algoritmo LSM en el contexto de los sistemas Haar para las discretizaciones introducidas en [5].
En el ámbito descriptivo pretendemos comprender el algoritmo. Esto implica, entender su fundamento e interpretación, su funcionamiento y sus propiedades (e.g. convergencia). En el ámbito prescriptivo esperamos simplificar la matemática del algoritmo y aplicarlo a nuevos contextos (e.g. procesos no markovianos).
Observamos que las bases utilizadas en la literatura son polinomios de Hermite, polinomios de Laguerre, sistemas spline, entre otros, todas clásicas. En consecuencia el objetivo que anhelamos alcanzar es conocer si la utilización de sistemas H (o Haar) asociadas a las discretizaciones representa un avance en esos aspectos.

Marco general de Análisis

El marco general está dado por las teorías de las diciplinas económicas y matemáticas.

Campo de conocimiento

El núcleo central de la investigación se apoyará en las teorías la matemática financiera, de la medida, de wavelets y de estadística.

Desarrollo del discurso

Podemos resumir el problema que encaramos en este trabajo mediante las siguientes preguntas, permite la utilización de los Haar systems ampliar el rango de aplicación del algoritmo LSM? simplifica esto la matemática?
En virtud del problema planteado vamos a formular las siguinetes hipótesis de trabajo:

- El algoritmo LSM utilizado en el contexto de los Haar Systems nos permite aplicarlo a nuevos casos.
- La variación introducida en el algoritmo LSM no mejora la matemática.

Vamos a seguir el siguiente lineamiento para la consecución de los objetivos. Para comenzar incluimos una sección en la que se introduce una idea básica de los mercados de Opciones y Futuros, luego presentamos la formalización matemática de los principales conceptos y propiedades del mercado de derivados financieros (en en el anexo I se da un breve repaso por los conceptos básicos de la teoría de probabilidad, dichos anexos así como la primera sección, están basados directamente en [1], [2] y [3]). En el mismo apartado incluimos además un panorama general de la valoración de opciones de tipo europeo. Pasaremos inmediatamente al caso de opciones de tipo americano en un modelo discreto y abordaremos el análisis del algoritmo LSM, esto último incluye la presentación del modelo, el estudio de su convergencia y un simple ejemplo que nos servirá para ilustrar su aplicación práctica y como punto de referencia. Luego abordaremos la construcción de las discretizaciones del espacio de probabilidad subyacente, para poder trabajar con él. Finalizaremos con los resultados obtenidos y dos ejemplos para ilustrar su funcionamiento.

Justificación del tema elegido

El tema elegido tiene las siguientes justificaciones:

- Justificación teórica: permitirá conocer la utilidad de la discretización mediante el uso de Haar systems para la matemática financiera.
- Justificación metodológica: crear el modelo (la variación sobre el mismo) y reglas de acción necesarias.
- Justificación práctica: mejorar la información a suministrar para la toma de decisiones en el desarrollo de estrategias dentro de los mercados de futuros.

2. Mercados de Futuros y Opciones

Los instrumentos financieros que vamos a analizar se agrupan bajo la denominación de *derivados financieros (financial derivatives)*. Esta denominación responde a que su valor se *deriva* del valor de algún otro activo, al que llamamos *subyacente* (underlying equity), e.g. acciones, bonos, divisas, commodities etc. Por ejemplo, un derivado sobre una acción, es un contrato específico cuyo valor dependerá enteramente del valor futuro de la acción en cuestión. Básicamente existen tres tipos de derivados financieros: *futuros*, *forwards* y *opciones*.

Primero detengámonos en el estudio de una estratégia muy utilizada e importante a la hora de trabajar en los mercados financieros. Esta transacción, normalmente conocida bajo el nombre de 'shorting', implica vender un activo del cual no somos aún propietarios. Esto es posible para algunos, pero no todos, los activos de inversión. Intentaremos, mediante el siguiente ejemplo, ilustrar su funcionamiento. Supongamos que un inversor le pide a su agente que realice una venta 'corta' de 100 acciones de Microsoft. El agente lleva a cabo las instrucciones del inversor tomando las acciones de otro cliente y vendiéndolas de la manera usual en el mercado. El inversor puede mantener su posición corta durante el tiempo que desee, siempre y cuando haya acciones para tomar 'prestadas' por parte del agente. En algún momento, sin embargo el inversor cerrará su posición comprando 100 acciones de Microsoft. Luego, estas son restituidas al cliente de quien fueron, en un principio, tomadas. El inversor, en tanto, habrá realizado una ganancia si el precio de las acciones descendió durante el tiempo en que mantuvo la posición corta (short position), claramente afrontará una pérdida en el caso contrario. Si en algún momento el agente se ve imposibilitado a seguir tomando prestadas acciones de sus otros clientes, el inversor se ve forzado a cerrar su posición inmediatamente (short-squeezed), a pesar de que no lo desee. Por supuesto esta operación tiene un costo para quien mantiene una posición corta, se debe pagar al agente un ingreso, i.e. un interés o dividendos. El agente transferirá luego este a la cuenta del cliente de quien los activos fueron tomados. Así mismo el inversor debe mantener una cuenta de margen con el agente. Esta cuenta trabaja de manera similar a las cuentas de margen que se mantienen en la cámara de compensaciones en una bolsa (ver más adelante). Esta última cuenta no representa un costo para el inversor pues normalmente se paga un interés sobre esos fondos.

2.1. **Futuros.** Los mercados de futuros se originaron con la necesidad de facilitar la comercialización de productos primarios, especialmente cereales y pueden, de hecho, ser rastreados sus orígenes hasta los tiempos de la Edad Media, cuando facilitaban la planificación a los granjeros y mercaderes. De esta manera, en 1848, nace el *Chicago Board of Trade*, donde en la actualidad, se manejan contratos de maíz, soja, trigo y varios subproductos. Su tarea inicial era la de estandarizar las cantidades y calidades de los granos que se negociaban. A penas unos pocos años después aparecieron los primeros contratos futuros, se los conocía como *to-arrive contract*. Rápidamente los especuladores se sintieron atraídos por estos contratos, y los encontraron una alternativa al intercambio de los granos mismos. Luego surgieron otros, como el *Chicago Mercantile Exchange* en 1919 que comenzó con la operatoria de manteca, carnes bovina y porcina para luego incluir a sus productos negociados harina de soja, aceite de soja, títulos públicos, metales e índices búrsatiles (en 1982

se introduce un contrato futuro sobre el Standar & Poor's 500 Stock Index). Fue con la desaparición de los tipos de cambio fijos a nivel internacional que en 1972 surgieron los primeros mercados a futuro de divisas, los que permitieron la cobertura de los operadores. Actualmente se considera que, los tipos de cambio fluctuantes, solo pueden sostenerse gracias a las coberturas que proporcionan dichos mercados.
Un futuro es un acuerdo a plazo, celebrado entre dos partes, mediante el cual se obligan mutuamente a comprar o vender en una determinada fecha futura un determinado activo a un determinado precio. El participante (ya sea una persona o una empresa) que se obliga a comprar, tiene lo que se denomina una *posición larga* (long futures position), mientras que quien se obliga a vender, posee una *posición corta* (short futures position); el precio al que se acordó realizar la transacción se denomina *precio futuro* (futures price). Así, llegada la fecha de expiración del contrato quien posee una posición larga tiene el derecho de recibir el activo subyacente y quien posee una posición corta posee el derecho a recibir el precio futuro a cambio del activo subyacente entregado (análogamente poseen la obligación que satisface el derecho de la contraparte).
Una característica práctica de los futuros, es su tendencia a no finalizar con la entrega del activo subyacente. La entrega del subyacente bajo los términos del contrato futuro, suele resultar inconveniente y en algunas circunstancias muy costosa. Esto es cierto inclusive para un hedger (ver su definición más adelante) que desea comprar o vender el activo subyacente. Por lo general se desea cerrar la posición y luego proceder, en el caso del hedger, a comprar o vender el activo de la forma usual.
Cerrar una posición implica entrar en la posición opuesta a la original. Por ejemplo, un inversor que compra cinco contratos futuros de julio el 6 de mayo, puede cerrar su posición el 20 de junio vendiendo cinco contratos futuros de julio. La ganancia o perdida del inversor es determinada por la diferencia de los precios del contrato entre el 6 de mayo y el 20 de junio.
Algunos futuros financieros, como aquellos sobre índices de acciones, se cierran en dinero puesto que resultaría inconveniente o directamente imposible entregar el subyacente. En el caso de un futuro sobre el S&P 500, por ejemplo, entregar el subyacente implicaría entregar un portafolio de 500 acciones.
Cuando se realiza un nuevo contrato (futuro), la 'exchange' (nombre que recibe el recinto donde se realizan las transacciones) debe especificar como mínimo lo siguiente: cual es el subyacente (si el subyacente es un commodity, se debe especificar la calidad)—, el tamaño del contrato (i.e. cuanto del subyacente será entregado por cada contrato, esto es una importante decisión, ya que si la cantidad es muy grande muchos pequeños inversores que desean proteger relativamente pequeñas cantidades o que desean tomar posiciones especulativas en el mercado no podrán hacerlo. Por otro lado si el contrato es demasiado pequeño su intercambio puede resultar costoso, dado que existe un costo asociado a cada contrato intercambiado), donde se llevará a cabo la entrega (especialmente en el caso en que los costos de entrega sean elevados) y cuando (normalmente se permite la entrega a lo largo de todo un mes, en consecuencia, a los futuros se los identifica por su mes de entrega).
Por lo general, los contratos poseen un *daily price movement limit*, esto es un valor que establece cotas dentro de las cuales la cotización del contrato puede moverse. Si el contrato alcanza alguno de esos limites, su intercambio cesa (e.g. si el precio

de un determinado contrato es de \$10 y el limite es de \$2, el intervalo en el cual la cotización puede moverse es [\$8,\$12]).
Si dos inversores se ponen en contacto y deciden realizar un contrato a futuro (como es el caso de un forward que veremos a continuación) existen obvios riesgos asociados. Uno de los inversores puede arrepentirse del trato e intentar salirse del mismo. En otras circunstancias, uno de los inversores puede simplemente verse imposibilitado a honrar el contrato. Uno de los principales roles de los mercados de valores (exchange) es evitar los defaults de los contratos. Es así que se introducen los llamados *márgenes* (depósitos de grantía). Para ilustrar su funcionamiento, consideremos un inversor que contacta su agente de bolsa el jueves 5 de junio para comprar dos futuros de diciembre cuyo subyacente es oro, en el COMEX (New York Commodity Exchange). Supongamos que el precio futuro ese día es de \$400 por onza. Dado que el tamaño del contrato es 100 onzas, el inversor realiza dos contratos a ese precio. El agente le pedirá al inversor que deposite fondos en una cuenta de margen (*margin account*). La cantidad que se deposita al momento de celebrarse el contrato se denomina margen inicial. Supongamos que es de \$2000 dólares por contrato, o \$4000 en total. Al finalizar cada día de transacciones, la cuenta de margen es ajustada para reflejar las ganancias o perdidas del inversor. Esta práctica es conocida como *marking to market.* Supongamos, por ejemplo, que al finalizar el día 5 de junio el precio futuro cayó de \$400 a \$397. El inversor perdió \$600. En consecuencia el balance de la cuenta de margen se reducirá a \$3400. Similarmente si el precio hubiese subido a \$403 al finalizar el día se depositarían \$600 en la cuenta. (el inversor tiene derecho a retirar cualquier balance en exceso de su margen inicial). Para asegurar que el margen nunca se vuelva negativo existe lo que se conoce como margen de mantenimiento, menor que el *margen inicial.* Si el balance de la cuenta cae por debajo del *margen de mantenimiento*, el inversor recibe una *llamada de margen* y se espera que deposite la diferencia hasta el margen inicial al siguiente día.
La *cámara de compensaciones* (*Clearinghouse*) es un adjunto de la bolsa (exchange) que actúa como un intermediario en las transacciones de futuros. Garantiza el cumplimiento de todas las partes. Esta posee un número de miembros. Los agentes que no son miembros deben canalizar sus negocios a través de uno. La tarea principal de la cámara de compensaciones es la de llevar registro de todas las transacciones que tienen lugar durante el día, para poder calcular la posición de cada miembro. Estos deben mantener una cuenta de margen con la cámara de compensaciones. Esta se conoce como *clearing margin.* Al finalizar cada día transable, dichas cuentas son ajustadas para reflejar las perdidas o ganancias del miembro, del mismo modo que se ajustan la de los inversores. Sin embargo, en el caso de los miembros, existe un margen inicial pero no uno de mantenimiento. Al finalizar el día, en consecuencia y dependiendo de las transacciones y los movimientos de precios ocurridos a lo largo del día, deben realizar un deposito o retirar (si lo desean) el excedente. El propósito de las cuentas de margen, no es otro que el de reducir la probabilidad de que los participantes del mercado afronten pérdidas como consecuencia de defaults. Este sistema ha resultado ser muy exitoso.

2.2. **Forwards.** Un contrato *forward* (contrato a plazo) es similar a un futuro, en el sentido que es un acuerdo para comprar o vender un determinado activo en un determinado momento en el futuro y a un determinado precio. Pero, mientras que

los *futures* son intercambiados en *exchanges*, los *forwards* son intercambiados en mercados del tipo *over-the-counter*. Estos últimos son una importante alternativa a los exchanges y se caracterizan porque las negociaciones son llevadas a cabo por medio de comunicaciones telefónicas y por lo general entre dos instituciones financieras o entre una institución financiera y uno de sus clientes corporativos (normalmente estas negociaciones por teléfono son grabadas, de modo tal que si, eventualmente, surgiese disputa alguna, esta pueda ser resuelta). Las instituciones financieras usualmente actúan como creadores de mercados para los instrumentos más usuales, eso significa que están siempre preparadas para ofrecer tanto un precio de venta como uno de compra (i.e. a los cuales ellos están dispuestos a realizar los intercambios). Una importante ventaja de este tipo de intercambio es que los contratos no deben adecuarse a un modelo previamente estipulado y rígido, dando la posibilidad a los interesados de negociar un contrato con las carcterísticas que a ambos le resulten atractivas y convenientes. Así mismo una importantísima desventaja, es la existencia de la posibilidad (aunque suele ser pequeña) de que el contrato no sea honrado. Veamos las principales diferencias entre un contrato futuro y uno forward. En primer lugar, como ya se mencionó, un forward es un contrato que se realiza entre dos agentes privados, mientras que un futuro se intercambia en un mercado de valores (exchange). Los contratos forward no están estandarizados, lo que reviste su principal ventaja; en cambio, en el caso de los futuros, existe una serie limitada de contratos, obligando a los interesados a adecuarse a los mismos. En el caso de los forwards usualmente se explicita una fecha única de entrega. Por otro lado en los futuros existe un rango de fechas de entrega, normalmente suele ser de un mes. Otra importante diferencia radica en que los forwards son ajustados al finalizar el contrato y los futuros son ajustados diariamente. Recordemos también, la existencia del riesgo asociado a los forwards debido a la posibilidad de que el contrato no se honre. En el caso de lo futuros esto no sucede, debido a la existencia de márgenes y a la regulación de las cámaras de compensación.
A continuación mostraremos algunos procedimientos simples para la valución de forwards. Para ello, consideremos un forward sobre un activo con precio S_0 que no produce ingreso alguno (e.g. acciones que no pagan dividendos). Notemos T al tiempo restante para su madurez (con respecto al significado de T, cabe decir que abusaremos de él, pudiendo utilizarlo indistintamente como la fecha de madurez), r a la tasa de interés libre de riesgo y F_0 al precio del forward. La relación entre S_0 y F_0 esta dada por:

$$F_0 = S_0 e^{rT}$$

Observemos por qué. Si $F_0 > S_0 e^{rT}$, un arbitrageurs podría tomar prestado S_0 a la tasa libre de riesgo, comprar el activo y tomar una posición corta en un forward sobre ese activo. Luego en t=T, pagar el préstamo, el cual habrá crecido para entonces a $S_0 e^{rT}$, entregar el subyacente y recibir a cambio F_0. Obteniendo así una ganancia libre de riesgo de:

$$F_0 - S_0 e^{rT} > 0$$

Procediendo de manera análoga si $F_0 < S_0 e^{rT}$, un arbitrageurs podría vender corto el activo, colocar lo recibido (S_0) a la tasa libre de riesgo y entrar en una posición larga sobre el mismo. En el momento t=T, su colocación habrá crecido a $S_0 e^{rT}$, dipondrá de una suma igual a F_0 para cerrar su posición en el forward y recibirá a cambio el subyacente, que entregará, a su vez, para cerrar la venta corta. De este modo se alzará con una ganancia libre de riesgo igual a

$$S_0 e^{rT} - F_0 > 0$$

Qué sucede en caso de que las ventas cortas no sean posibles? Como vamos a ver, en realidad, no necesariamente se convierte esto en un inconveniente. Solo hace falta que haya un número significante de participantes en el mercado que posean el activo como un bien de inversión. El único caso que debemos analizar es cuando $F_0 < S_0 e^{rT}$. Supongamos que el subyacente es oro (una onza por contrato) y que no existen costos de almacenamiento. En este caso el inversor que posee una onza de oro puede proceder de la siguiente manera:

(1) Vender el oro a S_0.
(2) Invertir lo recaudado a la tasa libre de riesgo r por un plazo igual a T
(3) Tomar una posición larga en un forward sobre una onza de oro.

Luego de transcurrido el plazo T la inversión habrá crecido a $S_0 e^{rT}$. La onza de oro se vuelve a comprar a F_0 y en consecuencia el inversor termina con una ganancia libre de riesgo igual a:

$$S_0 e^{rT} - F_0 > 0$$

Debemos recordar la idea de que son las leyes de la oferta y la demanda las que determinan el precio. Y suponiendo que nos encontramos en un mercado de *competencia perfecta*, la existencia de dos precios distintos para un mismo bien es insostenible en el tiempo.
Analizaremos ahora el caso en el que el subyacente genera un ingreso conocido (e.g. acciones que pagan dividendos conocidos). Llamemos I al valor presente del ingreso provisto por el subyacente durante la vida del forward. Entonces tenemos:

$$F_0 = (S_0 - I) e^{rT}$$

Supongamos que $F_0 > (S_0 - I) e^{rT}$, entonces un arbitrageurs podría conseguir una ganancia al pedir un préstamo (de S_0) a la tasa de interés libre de riesgo r, comprar el subyacente (S_0) y tomar una posición corta en un forward sobre el activo en cuestión. Llegada la fecha de ejercicio (T), pagaría el prestamo, el que habría crecido a $S_0 e^{rT}$, recibiría un total (valuado en T) de Ie^{rT} a cuenta del ingreso generado por el subyacente y recibiría también F_0 en concepto del forward que cerraría entregando el subyacente. Su ganancia resultaría igual a:

$$F_0 - (S_0 - I)e^{rT} > 0$$

En el caso en que $F_0 < (S_0 - I)e^{rT}$ se procede de manera análoga.

2.3. **Opciones.** El irrupción de las opciones se dió en Europa y en los Estados Unidos en el siglo XVIII. En sus primeros años el mercado obtuvo una mala reputación debido a ciertas prácticas corruptas. Entre ellas el entregarle opciones sobre un cierto activo a los agentes de bolsa para que luego recomienden a sus clientes la compra de dicho activo. A comienzos del siglo XIX un grupo de empresas fundo el *Put and Call Brokers and Dealers Association.* El objetivo de esta asociación era el de proveer un mecanismo mediante el cual se pudiera acercar a los compradores y vendedores en un mismo lugar. Aquellos inversores que estaban interesados en comprar una opcón debían contactarse con alguno de los integrantes de la asociación para que luego ellos se encargaran de encontrar un vendedor de la opción entre sus otros clientes. Si se daba el caso en el que no se encontraba una contraparte para la operación, la misma asociación se erguía como vendedora.
El mercado de opciones del *Put and Call Brokers and Dealers Association* sufría de dos deficiencias. Primero, que no existía un mercado secundario, donde el comprador de una opción podía venderla a otro participante antes de la fecha de expiración. Segundo, no existía mecanismo alguno mediante el cual se pudiera garantizar que los participantes honraran los contratos. Si uno de ellos no lo hacía el otro debía embarcarse en una serie de costosos juicios.
Fue entonces, que en abril de 1973 la *Chicago Board of Trade* creó la *Chicago Board Options Exchange*, específicamente para intercambiar opciones. Desde entonces los mercados de opciones se han vuelto cada vez más populares entre los inversores.
Existen dos tipos básicos de *opciones* (*options*): los denominados *calls* y los denominados *puts.* Un call, le brinda al *titular* (*holder*) el derecho de comprar un cierto activo en una determinada fecha futura a un precio específico. Un put, en cambio, le da a su titular, el derecho de vender un cierto activo en una determinada fecha futura a un precio específico. El precio en el contrato se conoce como *precio de ejercicio* (*execise price* o *strike price*), la fecha en el contrato se conoce como *fecha de expiración* (*expiration date*, *exercise date* o *maturity*). Las opciones pueden, principalmente, ser de dos clases, *Europeas* (*European options*), que solo pueden ser ejecutadas en la fecha de expiración; o *Americanas* (*American options*), que son ejecutables a lo largo de toda su vida. Existen cuatro tipos de participantes en los mercados de opciones: compradores de calls, vendedores de calls, compradores de puts y vendedores de puts. Los compradores se dice que poseen una *posición larga* (se los conoce también como *holders* o *titulares*), mientras que los vendedores una *posición corta* (a estos últimos también se los conoce como *writers*). Cabe enfatizar, que aunque el titular de una opción posee el derecho de ejercerla, no tiene la obligación de hacerlo. Esta característica es la que distingue a las opciones de los futuros y forwards. En principio no reviste costo alguno entrar en un mercado de futuros (excepto por los requerimientos de márgenes), mientras que para obtener una opción el inversor debe pagar un precio por el contrato al momento de realizar la operación. Por cada *opción* podemos identificar a dos partes, el que tomó la posición larga (compró la opción) y el que tomó la posición corta (vendió la opción). La ganancia o pérdida del vendedor de la opción es la contraria a la del comprador.

Usualmente resulta útil caracterizar a las distintas posiciónes en una opción en términos de su valor para el inversor al momento de la madurez. El costo inicial de la opción es, por ende, omitida en los cálculos. Si llamamos K al precio de ejercicio y S_T el precio final del subyacente (i.e. al momento t=T), el resultado de una posición larga en un call europeo esta dado por la función $\max(S_T - K, 0)$, esto refleja el hecho de que la opción será ejercida si $S_T > K$ y no será ejercida si $S_T < K$. El resultado para quien posee una posición corta en un call europeo es,

$$-\max(S_T - K, 0) = \min(K - S_T, 0)$$

Por otro lado, y procediendo similarmente obtenemos que el resultado para un inversor que tiene una posición larga en un put europeo es $\max(K - S_T, 0)$, mientras que el resultado para quien posee una posición corta en un put europeo es,

$$-\max(K - S_T, 0) = \min(S_T - K, 0)$$

Los principales subyacentes sobre los que se suscribe una opción, o sea, los que son activamente transados, son acciones, índice de acciones, divisas y futuros. Para el siguiente análisis de los precios de las opciones y con la intención de ilustrar como se determinan utilizaremos acciones como subyacente. Por lo tanto, podemos hablar de los siguientes factores como los que inciden en el precio de una opción sobre una acción:

- El precio actual de la acción, S_0.
- El precio de ejercicio (*strike price*), K.
- El tiempo hasta la expiración, T.
- La volatilidad del precio de la acción, σ.
- La tasa de interés libre de riesgo, r.
- Los dividendos esperados durante la vida de la opción.

A continuación analizaremos en que forma lo hacen. Para eso nos valdremos de un clásico concepto frecuentemente utilizado en economía, la condición *ceteris paribus*. Esta condición consiste en suponer que, excepto las variables que se están estudiando, todas las demás permanecen constantes.

Variable	**call europeo**	**put europeo**	**call americano**	**put americano**
precio actual de la acción	+	-	+	-
precio de ejercicio	-	+	-	+
tiempo hasta la expiración	?	?	+	+
volatilidad	+	+	+	+
tasa de interes libre de riesgo	+	-	+	-
Dividendos	-	+	-	+

Si un call es ejercido en algún momento futuro, el valor de la opción es la cantidad en la que el precio de la acción excede al precio de ejercicio (*strike price*). Por lo tanto el call se revaloriza a medida que el precio de la acción aumenta y se desvaloriza a medida que el precio de ejercicio aumenta. Para un put sucede lo contrario, dado que su valor esta determinado por la cantidad en que el precio de ejercicio excede el de la acción.

Consideremos el efecto de la fecha de ejercicio. En el caso de las opciones americanas, tanto el call como el put aumentan su valor cuando la fecha de ejercicio

se retrasa más en el tiempo. Consideremos dos opciones americanas que solo difieren en la fecha de expiración. El titular de la opción con mayor tiempo hasta la madurez tiene todas las posibilidades de ejercer la opción que las que tiene el titular de la que tiene menor tiempo hasta la madurez, y **más**. Por lo tanto la opción con mayor plazo hasta la fecha de ejercicio debe poseer un valor mayor que aquella con menor tiempo hasta la fecha de ejercicio.
El efecto que el tiempo puede tener sobre el valor de las opciones europeas no es tan claro. Aunque normalmente su valor tiende a aumentar a medida que la fecha de ejercicio se aleja en el tiempo, este no siempre es el caso. Consideremos dos calls europeos sobre una acción, una con una fecha de expiración en un mes y la otra en dos meses. Supongamos que se espera un pago importante en concepto de dividendos en seis semanas. Luego, como el precio de la acción bajará como consecuencia del pago de los dividendos, la opción con madurez en un mes puede terminar valiendo más que la otra.
En términos generales, la volatilidad del precio de una acción es una medida de la incertidumbre que se tiene sobre los futuros movimiento del precio de la acción. Por lo tanto, a medida que la volatilidad aumenta, la probabilidad de que a la acción le vaya muy bien o muy mal aumenta. Para el dueño de la acción, estos resultados tienden a compensarse. Sin embargo, esto no es así para el tenedor del put o del call. El titular de un call se beneficia de los aumentos en el precio, pero posee un riesgo limitado ante las bajas en el precio, porque lo máximo que puede perder es el precio que pagó por el contrato. Similarmente, el titular de un put se beneficia de las bajas en el precio, pero posee un riesgo limitado ante los eventuales aumentos del precio. Por lo que, consecuentemente, tanto el call como el put aumentan su valor a medida que lo hace su volatilidad.
El efecto que la tasa de interés libre de riesgo tiene sobre el precio de las opciones no es tan claro. A medida que las tasas de interés suben en la economía, la tasa de crecimiento esperada para el precio de las acciones tiende a incrementarse. Sin embargo, el valor presente de un flujo de efectivo que el titular de la opción recibiría, en el caso de ejercer la opción, decrece. Ambos efectos tienden a disminuir el valor de un put. Por ende, el precio de un put decrece a medida que la tasa de interés libre de riesgo aumenta. En el caso de un call, el primer efecto tiende a incrementar el precio, mientras que el segundo efecto tiende a hacerlo decrecer. De todas maneras se puede ver que el primer efecto siempre domina al segundo. Entonces se concluye, que el precio de un call siempre aumenta ante el aumento de la tasa de interés libre de riesgo.
Los dividendos tienen el efecto de reducir el precio de la acción una vez que estos son pagados. Por lo que el valor de un put aumenta a medida que lo hace el tamaño de los dividendos y lo contrario sucede con los calls.

A los derivados financieros a los que nos referimos a lo largo de este trabajo se los conoce como *plain vanilla products*. Existen, así mismo, una serie de derivados *exóticos* que son creados por ingenieros financieros y negociados en mercados de tipo *over-the-counter*. Estos son variaciones sobre los estándares *call* y *put* anteriormente tratados.
Se pueden enumerar una cantidad de razones por las cuales aparecen dichos instrumentos financieros. Pueden responder a una necesidad genuina de cobertura (hedge) en el mercado; o a la existencia de regulaciones, impuestos, cuestiones legales o contables que los convierten en herramientas más atractivas.

A modo de mención, pasamos a presentar algunas de los más comunes. En una opción americana estándar la ejecución puede llevarse a cabo en cualquier momento de su vida y el precio de ejercicio es siempre el mismo. En la práctica esto no siempre es así. Por ejemplo, si el ejercicio temprano (anterior a la fecha de expiración) es restringido a ciertas fechas, la opción se conoce como *Bermuda*; otras variaciones son permitir el ejercicio temprano solo durante una parte de la vida de la opción o permitir que el precio de expiración cambie a lo largo del tiempo.
Las opciones *compuestas* (*Compound options*) son opciones sobre opciones. Existen, claramente, cuatro clases de opciones compuestas: call sobre call, put sobre call, call sobre put y put sobre put. Las opciones compuestas poseen dos precios de ejercicio y dos fechas de expiración. Consideremos, por ejemplo, un call sobre un call. En la primera fecha de expiración, T_1, el titular de la opción compuesta tiene derecho a pagar el primer precio de ejercicio, X_1, y recibir un call. Este último call le otorga el derecho a su titular a comprar el subyacente por el segundo precio de ejercicio, X_2, en la segunda fecha de expiración, T_2.
Las opciones *chooser* (conocidas como *as you like it option*) tienen la característica de que luego de un plazo dado, el titular puede elegir (*choose*) si la opción es un call o un put. Supongamos que el momento en que dicha decisión es tomada es T_1. Luego el valor de la opción *chooser* esta dado por

$$\max(c,\ p)$$

donde c es el valor del call y p el del put. Si las opciones subyacentes a la opción *chooser*, son ambas europeas y tienen el mismo precio de ejercicio, se puede utilizar la paridad put-call para determinar una fórmula de valuación.
Otra clase de opciones importantes dentro de las exóticas son las asiáticas. Estas opciones tienen la característica de que su retribución depende del precio medio del subyacente durante, por lo menos, una parte de la vida de la opción. EL resultado final está dado para un *call* por $\max(0, S_{med} - K)$ y para un *put* por $\max(0, K - S_{med})$, donde S_{med} es el precio medio.

Es clara la necesidad de poder determinar los precios de las opciones si pretendemos trabajar con ellas. Para el caso de opciones Europeas el problema queda resuelto por el modelo introducido en 1973 por *Black, Scholes & Merton* y por el introducido por *Cox, Ross & Rubinstein* en 1979 entre otros. No es el caso de las opciones Americanas. Una alternativa, como ya se mencionó, parece ser el modelo LSM (2000). A este modelo es al que pretendemos incorporarle una variación para extender su campo de aplicación. Pasemos a ello.

3. Modelización del mercado de opciones

Es suficiente trabajar con un *espacio de probabilidad* $(\Omega, \mathcal{F}, \mathbb{P})$ finito, con un número $|\Omega|$ de puntos ω, cada uno con probabilidades positivas, i.e. $\mathbb{P}(\{\omega\}) > 0$, para ilustrar estas ideas. Por lo tanto vamos a estudiar mercados *finitos*, i.e. modelos de tiempo-discreto de mercados financieros, en los cuales todas las variables relevantes toman valores finitos. La generalización al caso en que $(\Omega, \mathcal{F}, \mathbb{P})$ no es finito, es inmediata. Es por ello que usamos el parámetro t para denotar el tiempo.
Especificamos un horizonte temporal y lo denotamos con T, la fecha final para todas las actividades económicas consideradas.

Nota: es claro que t (y en particular T) no son las fechas relevantes, sino una indexación de las mismas mediante una suceción, pero nos sentiremos libres de referirnos a ellas como las fechas mismas.

Utilizaremos una *filtración* $\mathcal{F}_t$, consistente de σ-álgebras $\mathcal{F}_0 \subset \mathcal{F}_1 \subset ... \subset \mathcal{F}_t$, donde $\mathcal{F}_0 = \{\emptyset, \Omega\}$ y $\mathcal{F}_T = \mathcal{F} = \mathcal{P}(\Omega)$.
Asumimos también que el mercado consiste de $d+1$ activos financieros. La interpretación usual es asumir que uno es *libre de riesgo* (localmente y lo supraindicamos con 0) y los d activos restantes son riesgosos (e.g. acciones). Los precios de los activos en el instante t son variables aleatorias, $S^0(t,\omega), S^1(t,\omega), ...S^d(t,\omega)$, no negativas y $\mathcal{F}_t$-medibles (i.e. en t conocemos los precios $S^i(t)$). Escribimos $S(t) = (S^0(t), S^1(t), ..., S^d(t))'$ al vector de precios en t (' indica la transpuesta). De aquí en adelante nos vamos a referir al espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, el conjunto de fechas, los procesos de precios S y a la estructura de la información $\mathcal{F}_t$ como el modelo $\mathcal{M}$.

Definición 3.1. *Un* numéraire *(numerario) es un proceso de precios* $(\mathrm{X}(t))_{t=0}^{\mathrm{T}}$ *(una sucesión de* v.a.*), que es estrictamente positivo para todo* $t \in \{0, 1, ..., \mathrm{T}\}$.

Para la aproximación estándar, el proceso determinado por una cuenta en un banco, *libre de riesgo*, es utilizado como *numéraire*. En algunas aplicaciones, sin embargo, resulta más conveniente utilizar una *seguridad* y por lo tanto se utiliza S_0 sin mayores especificaciones.
Vamos a definir $S^0(0) = 1$ y $\beta(t) := 1/S^0(t)$ como factor de descuento. Nos vamos a refierir a los demás activos $1, 2, ..., d$ como los activos riesgosos.
Una *estrategia de inversión* (*trading strategy* o *dynamic portfolio*) φ es un vector estocástico de $\mathbb{R}^{d+1}$, $\varphi = (\varphi(t))_{t=1}^T = ((\varphi^0(t,\omega), \varphi^1(t,\omega), ..., \varphi^d(t,\omega)))_{t=1}^T$, que es predecible, i.e. cada $\varphi^i(t)$ es $\mathcal{F}_{t-1}$-medible para $t \geq 1$. Cada $\varphi^i(t)$ denota la cantidad de unidades del activo i que hay en el portafolio en el momento t, cantidad que es determinada por la información disponible antes del momento t, i.e. el inversor selecciona el portafolio que mantendrá en t luego de observar los precios $S(t-1)$. Sin embargo, el portafolio $\varphi(t)$ debe establecerse antes y mantenido hasta después del anuncio de los precios $S(t)$. Las componentes $\varphi^i(t)$ pueden asumir tanto valores positivos como negativos, reflejando el hecho de que se permite las ventas 'cortas' (*short selling*) y asumiendo que los activos son perfectamente divisibles.

Definición 3.2. *El* valor *de un portafolio en t es el producto escalar*

$$\mathrm{V}_{\varphi}(t)=\varphi(t)\cdot\mathrm{S}(t):=\sum_{i=0}^{\mathrm{d}}\varphi^{i}(t)\mathrm{S}^{i}(t),\ (t=1,2,...,\mathrm{T})\ y\ \mathrm{V}_{\varphi}(0):=\varphi(1)\cdot\mathrm{S}(0)$$

$V_{\varphi}(t,\omega)$ se denomina proceso estocástico del *valor de la estrategia de inversión* φ. El valor inicial del portafolio $V_{\varphi}(0)$ se llama *inversión inicial*.

Ahora, $\varphi(t)\cdot S(t-1)$ refleja el valor de mercado del portafolio, luego de que este fue establecido en $t-1$, mientras que $\varphi(t)\cdot S(t)$ es el valor del portafolio un instante después de que los precios $S(t)$ fueron observados, pero antes de que se realicen cambios en la composición del portafolio. Por lo tanto, tenemos que

$$\varphi(t)\cdot(S(t)-S(t-1))=\varphi(t)\cdot\triangle S(t)$$

es el cambio en el valor de mercado del portafolio, debido a la variación de los precios ocurrida entre los momentos $(t-1)$ y t.

Definición 3.3. *El proceso estocástico,* G_{φ} *que representa las ganancias de una estrategia de inversión,* φ *está dada por*

$$\mathrm{G}_{\varphi}(t):=\sum_{\tau=1}^{t}\varphi(\tau)\cdot(\mathrm{S}(\tau)-\mathrm{S}(\tau-1))=\sum_{\tau=1}^{t}\varphi(\tau)\cdot\triangle\mathrm{S}(\tau),\ (t=1,2,...,\mathrm{T})$$

definimos $\tilde{S}(t):=(1,\beta(t)S^{1}(t),...,\beta(t)S^{d}(t))'$ como el *vector de precios descontados*, y consideramos el proceso de *valor descontado*

$$\tilde{V}_{\varphi}(t)=\beta(t)(\varphi(t)\cdot S(t))=\varphi(t)\cdot\tilde{S}(t),\ (t=1,2,...,T).$$

y el *proceso de ganancias descontadas*

$$\tilde{G}_{\varphi}(t):=\sum_{\tau=1}^{t}\varphi(\tau)\cdot(\tilde{S}(\tau)-\tilde{S}(\tau-1))=\sum_{\tau=1}^{t}\varphi(\tau)\cdot\triangle\tilde{S}(\tau),\ (t=1,2,...,T)$$

Observese que este último proceso solo refleja las ganancias provenientes del intercambio de los d activos riesgosos (no del activo 0).

Definición 3.4. *Una estrategia de inversión* φ *se dice* autofinanciada*,* $\varphi\in\Phi$*, si*

$$\varphi(t)\cdot\mathrm{S}(t)=\varphi(t+1)\cdot\mathrm{S}(t)\ \ (t=1,2,...,\mathrm{T}-1) \tag{1}$$

Esto se interpreta de la siguiente manera, cuando se conocen los precios $S(t)$ en el momento t, el inversor ajusta su portafolio de $\varphi(t)$ a $\varphi(t+1)$, utilizando, ni más ni menos, el valor del portafolio.

Proposición 3.1. *Sea* $\mathrm{X}(t)$ *un* numéraire*. Una estrategia de inversión* φ *es* autofinanciada *con respecto a* $\mathrm{S}(t)$ *sii* φ *es* autofianciada *con respecto a* $\mathrm{X}(t)^{-1}\mathrm{S}(t)$*.*

Dem: Dado que $X(t)$ es estrictamente positivo para todo $t=1,2,...,T$, tenemos la siguiente equivalencia, que demuestra la proposición:

$$\varphi(t)\cdot S(t) = \varphi(t+1)\cdot S(t)\ (t = 1, 2, ..., T-1)$$

$$\Longleftrightarrow$$

$$\varphi(t)\cdot X(t)^{-1}S(t) = \varphi(t+1)\cdot X(t)^{-1}S(t)\ (t = 1, 2, ..., T-1) \quad \square$$

Corolario 3.1. *Una estrategia de inversión φ es* autofinanciada *con respecto a* $\mathrm{S}(t)$ *si y solo si φ es* autofinanciada *con respecto a* $\tilde{\mathrm{S}}(t)$

Proposición 3.2. *Una estrategia de inversión φ pertenece a Φ si y solo si*

$$\tilde{\mathrm{V}}_\varphi(t) = \mathrm{V}_\varphi(0) + \tilde{\mathrm{G}}_\varphi(t),\ (t = 0, 1, ..., \mathrm{T}).$$

Dem: Asumamos que $\varphi \in \Phi$. Luego utilizando la relación dada por (1), la proposición (2.1) y el hecho de que $S^0(0) = 1$

$$V_\varphi(0) + \tilde{G}_\varphi(t) = \varphi(1)\cdot S(0) + \sum_{\tau=1}^{t} \varphi(\tau)\cdot(\tilde{S}\ (\tau) - \tilde{S}\ (\tau-1))$$

$$= \varphi(1)\cdot\tilde{S}\ (0) + \varphi(t)\cdot\tilde{S}\ (t) + \sum_{\tau=1}^{t-1}(\varphi(\tau) - \varphi(\tau+1))\cdot\tilde{S}\ (\tau) - \varphi(1)\cdot\tilde{S}\ (0)$$

$$= \varphi(t)\cdot\tilde{S}\ (t) = \tilde{V}_\varphi(t)$$

Asumamos ahora que $\tilde{V}_\varphi(t) = V_\varphi(0) + \tilde{G}_\varphi(t)$, $(t = 0, 1, ..., T)$, se cumple. Por la proposición (2.1), es suficiente mostrar que se cumple la condición de *autofinanciamiento* de φ con respecto a el proceso descontado de precios $\tilde{S}(t)$ para cada $t = 1, 2, ..., T-1$.

$\tilde{V}_\varphi(2) = \varphi(2)\cdot\tilde{S}\ (2) = \varphi(1)\cdot\tilde{S}\ (0) + \varphi(1)\cdot(\tilde{S}\ (1) - \tilde{S}\ (0)) + \varphi(2)\cdot(\tilde{S}\ (2) - \tilde{S}\ (1))$.
Reordenando, tenemos

$$\varphi(2)\cdot\tilde{S}\ (1) = \varphi(1)\cdot\tilde{S}\ (1)$$

Luego, por inducción se muestra que $\varphi(t)\cdot\tilde{S}\ (t) = \varphi(t+1)\cdot\tilde{S}\ (t)$, para todo $t = 2, ..., T-1$ como queriamos. $\square$

Proposición 3.3. *Si $(\varphi^1(t), ..., \varphi^{\mathrm{d}}(t))$ es predecible y V_0 es $\mathcal{F}_0$-medible, existe un único proceso $(\varphi^0(t))_{t=1}^{\mathrm{T}}$ tal que $\varphi = (\varphi^0, \varphi^1, ..., \varphi^{\mathrm{d}})$ es* autofinanciada *con valor inicial del correspondiente portafolio* $\mathrm{V}_\varphi(0) = \mathrm{V}_0$.

Dem: Si φ es *autofinanciada*, luego por la proposición (2.2),

$$\tilde{V}_\varphi(t) = V_\varphi(0) + \tilde{G}_\varphi(t) = V_0 + \sum_{\tau=1}^{t}(\varphi^1(\tau)\triangle\tilde{S}^{\ 1}(\tau) + ... + \varphi^d(\tau)\triangle\tilde{S}^{\ d}(\tau)).$$

Por otro lado tenemos que

$$\tilde{V}_{\varphi}(t) = \varphi(t) \cdot \tilde{S}(t) = \varphi^0(t) + \varphi^1(t)\tilde{S}^{\,1}(t) + ... + \varphi^d(t)\tilde{S}^{\,d}(t)$$

luego

$$\varphi^0(t) = V_0 + \sum_{\tau=1}^{t}(\varphi^1(\tau)\triangle\tilde{S}^{\,1}(\tau) + ... + \varphi^d(\tau)\triangle\tilde{S}^{\,d}(\tau)) - (\varphi^1(t)\tilde{S}^{\,1}(t) + ... + \varphi^d(t)\tilde{S}^{\,d}(t))$$

lo que define a $\varphi^0(t)$ de forma única. Por otro lado tenemos que

$$\varphi^i(t)\triangle\tilde{S}^i(t) - \varphi^i(t)\tilde{S}^i(t) = -\varphi^i(t)\tilde{S}^i(t-1)$$

que es $\mathcal{F}_{t-1}$-medible. Luego

$$\varphi^0(t) = V_0 + \sum_{\tau=1}^{t-1}(\varphi^1(\tau)\triangle\tilde{S}^{\,1}(\tau) + ... + \varphi^d(\tau)\triangle\tilde{S}^{\,d}(\tau)) -$$
$$-(\varphi^1(t)\tilde{S}^{\,1}(t-1) + ... + \varphi^d(t)\tilde{S}^{\,d}(t-1))$$

dado que $\varphi^1, ..., \varphi^d$ son predecibles, todos los términos en la derecha son $\mathcal{F}_{t-1}$-medibles y por ende φ^0 es predecible. □

Un principio fundamental para cualquier modelo de mercado es el de *ausencia de oportunidades de arbitraje*, i.e. la ausencia de estrategias libres de riesgo que permitan alzarse con ganancias sin invertir.

Definición 3.5. *Sea* $\tilde{\Phi} \subset \Phi$ *un conjunto de estrategias* autofinanciadas. *Una estrategia* $\varphi \in \tilde{\Phi}$ *es llamada una* oportunidad de arbitraje *o una* estrategia de arbitraje *con respecto a* $\tilde{\Phi}$ *si* $\mathbb{P}\{\mathrm{V}_{\varphi}(0) = 0\} = 1$*, y el valor final de* φ *satisface*

$$\mathbb{P}\{\mathrm{V}_{\varphi}(\mathrm{T}) \geq 0\} = 1 \text{ y } \mathbb{P}\{\mathrm{V}_{\varphi}(\mathrm{t}) > 0\} > 0.$$

Por lo tanto una *oportunidad de arbitraje* es una estrategia *autofinanciada* con valor inicial cero, que produce un resultado final ($V_{\varphi}(T)$) no negativo con probabilidad uno y tiene probabilidad positiva de que el valor final sea positivo.

Definición 3.6. *Decimos que un mercado* $\mathcal{M}$ *no tiene* oportunidades de arbitraje*, si no existen* estrategias de arbitraje *en la clase* Φ*.*

Cuando observamos una manifestación $S(t,\omega)$ del proceso de precios $S(t)$, queremos saber que punto muestral $\omega \in \Omega$ tenemos. La información acerca de ω está contenida en la filtración $\mathcal{F}_t$. Como se menciona en el anexo I, siempre podemos encontrar una sucesión cada vez más fina de particiones $\mathcal{P}_t$, que corresponde a la filtración $\mathcal{F}_t$. Por lo tanto en t conocemos el conjunto $A_t \in \mathcal{P}_t$ con $\omega \in A_t$. Recordemos la estructura de las siguientes particiones. Un conjunto $A \in \mathcal{P}_t$ es unión disjunta de conjuntos $A_1, ..., A_K \in \mathcal{P}_{t+1}$. Dado que $S(u)$ es $\mathcal{F}_{(u)}$-medible $S(t)$ es constante en A y $S(t+1)$ es constante en los A_k, $k = 1, ..., K$. Así que, podemos pensar al conjunto A como el 'estado' en el momento $t = 0$ en un modelo de un solo periodo, y a cada A_k como el correspondiente a un 'estado' en el instante $t = 1$. Podemos pensar, en consecuencia, a un modelo de varios periodos como una colección de modelos de un único periodo. Veamos, entonces, cual es el efecto de la condición '*no poseer oportunidades de arbitraje*' en un mercado 'global' (de varios periodos) en los modelos que abarcan un único periodo.

Lema 3.1. *Si $\mathcal{M}$ no contiene* oportunidades de arbitraje*, luego para todo $t \in \{0, 1, ..., \mathrm{T}-1\}$, para toda estrategia autofinanciada $\varphi \in \Phi$ y para cualquier $A \in \mathcal{P}_t$*

(1) $\mathbb{P}(\tilde{\mathrm{V}}_\varphi(t+1) - \tilde{\mathrm{V}}_\varphi(t) \geq 0|A) = 1 \Longrightarrow \mathbb{P}(\tilde{\mathrm{V}}_\varphi(t+1) - \tilde{\mathrm{V}}_\varphi(t) = 0|A) = 1$
(2) $\mathbb{P}(\tilde{\mathrm{V}}_\varphi(t+1) - \tilde{\mathrm{V}}_\varphi(t) \leq 0|A) = 1 \Longrightarrow \mathbb{P}(\tilde{\mathrm{V}}_\varphi(t+1) - \tilde{\mathrm{V}}_\varphi(t) = 0|A) = 1$

La interpretación económica, es que la no existencia de oportunidades de arbitraje 'globalmente' implica la no existencia de oportunidades de arbitraje 'localmente'.

Dem: Desmostraremos únicamente (1) ((2) se demuestra de manera similar). Fijemos $t \in \{0, ..., T-1\}$ y $\varphi \in \Phi$. Supongamos que $\mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) \geq 0|A) = 1$ para algún $A \in \mathcal{P}_t$ y definamos una nueva estrategia de inversión ψ para los momentos $u = 1, ..., T$, de la siguiente manera:

Para $u \leq t : \psi(u) = 0$ (i.e. 'hacer nada antes del instante t').
Para $u = t+1 : \psi(t+1) = 0$ si $\omega \in A^c$, y

$$\psi^k(t+1, \omega) = \begin{cases} \varphi^k(t+1, \omega) & \text{si } \omega \in A \text{ y } k \in \{1, ..., d\}, \\ \varphi^0(t+1, \omega) - \tilde{V}_\varphi(t, \omega) & \text{si } \omega \in A \text{ y } k = 0. \end{cases}$$

Esta estrategia se puede entender de la siguiente forma, 'si ω resulta pertencer a A en t, se sigue la estrategia φ para los activos riesgosos, pero modificamos nuestra posición sobre el *numéraire* apropiadamente para compensar el hecho de no hacer nada cuando $\omega \in A^c$'. Esto último significa que $\varphi^0(t+1, \omega) - \tilde{V}_\varphi(t, \omega)$ no es otra cosa que lo que debo endeudarme para poder financiar la adquisición de los demás activos, recuerdese que hasta el momento t inclusive no poseemos activos.

Para $u > t+1 : \psi^k(u) = 0$ para $k \in \{1, ..., d\}$ y

$$\psi^0(u, \omega) = \begin{cases} \tilde{V}_\psi(t+1, \omega) & \text{si } \omega \in A, \\ 0 & \text{si } \omega \in A^c. \end{cases}$$

Esto es 'invertir la cantidad $\tilde{V}_\psi(t+1)$ en el *numéraire* si ω resulta pertenecer a A, de lo contrario no hacer nada'.

El siguiente paso es mostrar que la estrategia ψ es *autofinanciada* (i.e. pertenece a Φ). Por construcción ψ es predecible, por lo tanto una estrategia de inversión. Para $\omega \in A^c$ $\psi \equiv 0$, por lo que solo debemos considerar los $\omega \in A$. El punto relevante del tiempo es $(t+1)$. Recordemos que $\psi(t) = 0$, por lo que $\psi(t) \cdot \tilde{S}(t) = 0$. Ahora tenemos que,

$$\psi(t+1) \cdot \tilde{S}\,(t) = (\varphi^0(t+1) - \tilde{V}_\varphi(t))\tilde{S}^{\,0}(t) + \sum_{k=1}^{d} \varphi^k(t+1)\tilde{S}^k(t)$$

$$= \sum_{k=0}^{d} \varphi^k(t+1)\tilde{S}^k(t) - \tilde{V}_\varphi(t) = \varphi(t+1)\cdot\tilde{S}\,(t) - \tilde{V}_\varphi(t) = \varphi(t)\cdot\tilde{S}\,(t) - \tilde{V}_\varphi(t) = 0$$

utilizando el hecho de que φ es *autofinanciada.* Dado que $\psi(u)\cdot\tilde{S}\,(u) = 0$ para $u \leq t$ tenemos que $\psi(u+1)\cdot\tilde{S}\,(u) = \psi(u)\cdot\tilde{S}\,(u)$ para todo $u \leq t$ (y para todo $\omega \in \Omega$). Cuando $u > t+1$ y $\omega \in A$ solo mantenemos el *numéraire* (con valor descontado constante igual a uno), entonces

$$\psi(u+1)\cdot\tilde{S}\,(u) = \tilde{V}_\psi(t+1) = \psi(u)\cdot\tilde{S}\,(u).$$

por lo tanto la estrategia ψ es *autofinanciada.*

Analizamos, ahora, el proceso del valor de ψ. Utilizando nuestra hipótesis $\mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) \geq 0|A) = 1$ vemos que para todo $u \geq t+1$ y $\omega \in A$

$$\tilde{V}_\psi(u) = \psi(u)\cdot\tilde{S}\,(u) = \psi(t+1)\cdot\tilde{S}\,(t+1)$$

$$= (\varphi^0(t+1) - \tilde{V}_\varphi(t))\tilde{S}^{\,0}(t) + \sum_{k=1}^{d}\varphi^k(t+1)\tilde{S}^{\,k}(t) = \sum_{k=0}^{d}\varphi^k(t+1)\tilde{S}^{\,k}(t) - \tilde{V}_\varphi(t)$$

$$= \tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) \geq 0$$

Como $\tilde{V}_\psi(T) = 0$ en A^c, ψ define una estrategia de inversión *autofinanciada* con $\tilde{V}_\psi(0) = 0$ y $\tilde{V}_\psi(T) \geq 0$. La hipótesis de un mercado 'libre de oportunidades de arbitraje', implica que $\tilde{V}_\psi(T) = 0$ o

$$\begin{aligned} 0 &= \mathbb{P}(\tilde{V}_\psi(T) > 0) = \mathbb{P}(\{\tilde{V}_\psi(T) > 0\} \cap A) \\ &= \mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) > 0|A)\mathbb{P}(A). \end{aligned}$$

Por lo tanto obtenemos lo que buscabamos, $\mathbb{P}(\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) = 0|A) = 1$

Definición 3.7. *Una medida de probabilidad* $\mathbb{P}^*$ *sobre* $(\Omega, \mathcal{F}_{\mathrm{T}})$ *equivalente a* $\mathbb{P}$ *es llamada una* medida martingala *para* $\tilde{\mathrm{S}}$ *si el proceso* $\tilde{\mathrm{S}}$ *es una* $\mathbb{P}^*$*-martingala con respecto a la filtración* $\mathcal{F}_t$*. Denotamos mediante* $\mathcal{P}(\tilde{\mathrm{S}})$ *a la clase de* medidas martingalas equivalentes.

Proposición 3.4. *Sea* $\mathbb{P}^*$ *una* medida martingala equivalente *(i.e.* $\mathbb{P}^* \in \mathcal{P}(\tilde{\mathrm{S}})$*) y* $\varphi \in \Phi$ *una estrategia autofinanciada cualquiera. Luego el proceso del* valor *de* φ*,* $\tilde{\mathrm{V}}_\varphi(t)$ *es una* $\mathbb{P}^*$*-martingala con respecto a la filtración* $\mathcal{F}_t$*.*

Dem: Por la propiedad de autofinanciamiento de φ (proposición 2.2), tenemos que

$$\tilde{V}_\varphi(t) = V_\varphi(0) + \tilde{G}_\varphi(t),\ (t = 0, 1, ..., T\ ).$$

entonces

$$\tilde{V}_\varphi(t+1) - \tilde{V}_\varphi(t) = \tilde{G}_\varphi(t+1) - \tilde{G}_\varphi(t) = \varphi(t+1)\cdot(\tilde{S}\,(t+1) - \tilde{S}\,(t)))$$

Entonces para $\varphi \in \Phi$, $\tilde{V}_\varphi(t)$ es la *transformada martingala* de la $\mathbb{P}^*$-martingala $\tilde{S}$ por φ y por lo tanto una $\mathbb{P}^*$-martingala. $\square$

Proposición 3.5. *Si existe una* medida martingala equivalente*, i.e.* $\mathcal{P}(\tilde{\mathrm{S}}) \neq \emptyset$*, entonces el mercado* $\mathcal{M}$ *no posee oportunidades de arbitraje.*

Dem: Asumamaos que existe una $\mathbb{P}^* \in \mathcal{P}(\tilde{S})$. Para cualquier estrategia autofinanciada φ, tenemos

$$\tilde{V}_\varphi(t) = V_\varphi(0) + \sum_{\tau=1}^{t} \varphi(\tau) \cdot \triangle \tilde{S}\,(\tau)$$

por la proposición (2.4), al ser $\tilde{S}(t)$ una $\mathbb{P}^*$-martingala, $\tilde{V}_\varphi(t)$ es una $\mathbb{P}^*$-martingala. Por lo tanto las *esperanzas* final e inicial (por $\mathbb{P}^*$) coinciden,

$$\mathbb{E}^*(\tilde{V}_\varphi(T)) = \mathbb{E}^*(\tilde{V}_\varphi(0))$$

Si la estrategia es una *oportunidad de arbitraje*, su valor inicial es cero. Por lo tanto, de acuerdo a la igualdad anterior, la esperanza, por $\mathbb{P}^*$, de $\tilde{V}_\varphi(T)$ es cero, pero $\tilde{V}_\varphi(T) \geq 0$ por definición. Además cada $\mathbb{P}^*(\{\omega\}) > 0$ (asumimos que cada $\mathbb{P}(\{\omega\}) > 0$, entonces por equivalencia cada $\mathbb{P}^*(\{\omega\}) > 0$). Esto y $\tilde{V}_\varphi(T) \geq 0$ hacen que $\tilde{V}_\varphi(T) = 0$. Por ende no es posible la presencia de *oportunidades de arbitraje.* □

Proposición 3.6. *Si el mercado* $\mathcal{M}$ *no posee* oportunidades de arbitraje*, luego la clase* $\mathcal{P}(\tilde{\mathrm{S}})$ *de* medidas martingalas equivalentes *es no vacía.*

Teorema 3.1. *El mercado* $\mathcal{M}$ *no posee* oportunidades de arbitraje *si y solo si existe una medida de probabilidad* $\mathbb{P}^*$ *equivalente a* $\mathbb{P}$ *bajo la cual el proceso* d*-dimensional de los precios de los activos descontados,* $\tilde{\mathrm{S}}$ *es una* $\mathbb{P}^*$*-martingala.*

Definición 3.8. *Un derivado financiero* X*, con fecha de ejercicio* T *es una variable aleatoria no negativa,* $\mathcal{F}_{\mathrm{T}}$*-medible, arbitraria. Denotamos la clase de todos los derivados financieros con* $\mathcal{X}^+$*.*

Un ejemplo lo podemos encontrar en el caso de una opción sobre algún activo subyacente S, tomemos un *call Europeo* con fecha de ejercicio T y precio de ejercicio K. Luego, podemos encontrar una relación funcional $X = f\,(S\,)$ con alguna función f , que en nuestro caso esta dada por, $X = \max\{S(T) - K, 0\}$.

Decimos que un derivado financiero es *alcanzable* (*attainable*) si existe una estrategia $\varphi \in \Phi$, que lo *reproduce*, esto es

$$V_\varphi(T) = X$$

Luego la estrategia φ que *reproduce* al derivado financiero, genera el mismo flujo de efectivo en T que X. La condición de 'no existencia de oportunidades de arbitraje' implica que para cualquier derivado financiero *alcanzable* su precio en t debe estar dado por el valor de cualquier estrategia que lo *reproduzca.* Esta es la idea básica detrás de la *precificación mediante arbitraje.*
Analicemos esto un poco más. La idea es *reproducir* un flujo de efectivo dado en un momento cierto. Utilizando una estrategia autofinanciada, la *riqueza* del inversor

puede volverse negativa en algún instante $t < T$, pero debe coincidir en la fecha final. Para evitar que una estrategia se vuelva negativa, introducimos el concepto de *admisible* (*admissble*). Una estrategia de inversión autofinanciada, $\varphi \in \Phi$, se dice *admisible* si $V_\varphi(t) \geq 0$ para cada $t = 0, 1, ..., T$. Notamos con Φ_a a la clase de estrategias *admisibles.* El supuesto del modelo, de estrategias *admisibles* refleja el hecho económico de que el agente debe estar protegido frente a ilimitadas ventas 'cortas' (*short sales*). Desde el punto de vista matemático no es realmete necesario y utilizamos estrategias autofinanciadas cuando abordamos los aspectos matemáticos de la teoría.
Entonces, dado un derivado financiero X, como podemos determinar su valor en los momentos $t < T$?. Para un derivado financiero alcanzable, este valor debe ser dado por el valor de cualquier estrategia que lo reproduzca, en el momento $t < T$. Debería existir un único proceso de *valor*, digamos $V_X(t)$ representando el valor en t de X. La siguiente proposición nos asegura que los procesos del *valor* de las distintas estrategias que reproducen el derivado financiero coinciden, probando así su unicidad.

Proposición 3.7. *Supongamos que el mercado $\mathcal{M}$ no posee* oportunidades de arbitraje. *Luego cualquier derivado financiero* alcanzable *es* reproducido *de manera única en $\mathcal{M}$.*

Dem: Supongamos que existe un derivado financiero X alcanzable y dos estrategias φ y ψ tales que,

$$V_\varphi(T) = V_\psi(T) = X,$$

pero existe un $\tau < T$ tal que

$$V_\varphi(u) = V_\psi(u) \text{ para cada } u < \tau \text{ y } V_\varphi(\tau) \neq V_\psi(\tau).$$

Definimos $A := \{\omega \in \Omega : V_\varphi(\tau, \omega) > V_\psi(\tau, \omega)\}$, luego $A \in \mathcal{F}_\tau$ y $\mathbb{P}(A) > 0$ (de lo contrario solo renombramos las estrategias). Definimos la variable aleatoria, $\mathcal{F}_\tau$-medible, $Y := V_\varphi(\tau) - V_\psi(\tau)$ y consideramos la estrategia ξ definida por

$$\xi(u) := \begin{cases} \varphi(u) - \psi(u) & u \leq \tau, \\ 1_{A^c}(\varphi(u) - \psi(u)) + 1_A(Y\beta(\tau), 0, ..., 0) & \tau < u \leq T. \end{cases}$$

Luego ξ es predecible y autofinanciada para $t \neq \tau$, y para $t = \tau$ tenemos, utilizando el hecho de que $\varphi, \psi \in \Phi$

$$\begin{aligned} \xi(\tau) \cdot S(\tau) &= (\varphi(\tau) - \psi(\tau)) \cdot S(\tau) = V_\varphi(\tau) - V_\psi(\tau) \\ \xi(\tau+1) \cdot S(\tau) &= 1_{A^c}(\varphi(\tau+1) - \psi(\tau+1)) \cdot S(\tau) + 1_A Y\beta(\tau) S_0(\tau) \\ &= 1_{A^c}(\varphi(\tau) - \psi(\tau)) \cdot S(\tau) + 1_A(V_\varphi(\tau) - V_\psi(\tau))\beta(\tau)\beta^{-1}(\tau) \\ &= V_\varphi(\tau) - V_\psi(\tau) \end{aligned}$$

Por lo tanto ξ es una estrategia autofinanciada con valor inicial cero. Más aún

$$\begin{aligned} V_\xi(T) &= 1_{A^c}(\varphi(\tau+1) - \psi(\tau+1)) \cdot S(\tau) + 1_A(Y\beta(\tau), 0, ..., 0) \cdot S(T) \\ &= 1_A Y\beta(\tau) S^0(\tau) \geq 0 \end{aligned}$$

y

$$\mathbb{P}\{V_\xi(T)>0\}=\mathbb{P}\{A\}>0$$

Entonces el mercado $\mathcal{M}$ contiene una *oportunidad de arbitraje* con respecto a la clase Φ de estrategias autofinanciadas. Pero esto contradice el supuesto de que el mercado $\mathcal{M}$ no posee *oportunidades de arbitraje.* $\square$

Definición 3.9. *Supongamos que el mercado $\mathcal{M}$ no posee* oportunidades de arbitraje*. Sea* X *un derivado financiero* alcanzable *con fecha de ejercicio* T*. Luego el proceso del precio de* arbitraje $\pi_{\mathrm{X}}(t)$*,* $0\le t\le \mathrm{T}$ *o simplemente el precio de* arbitraje *de* X *en* $\mathcal{M}$ *esta dado por el proceso del valor de cualquier estrategia,* φ *que reproduzca a* X*.*

Proposición 3.8. *El proceso del precio de* arbitraje *de cualquier derivado financiero* alcanzable*,* X*, esta dado por la* fórmula de valuación neutral ante el riesgo *(*risk-neutral valuation formula*),*

$$\pi_{\mathrm{X}}(t)=\beta(t)^{-1}\mathbb{E}^*(\mathrm{X}\beta(\mathrm{T})|\mathcal{F}_t),\ \forall t=0,1,...,\mathrm{T},$$

donde $\mathbb{E}^*$ *es el operador esperanza con respecto a una* medida martingala equivalente $\mathbb{P}^*$*.*

Dado que asumimos que el mercado $\mathcal{M}$ no posee oportunidades de arbitraje, existe una *medida martingala equivalente* $\mathbb{P}^*$. Luego por la proposición (2.4) el *proceso del valor descontado*, $\tilde{V}_\varphi$ respecto de cualquier estrategia autofinanciada, φ es una $\mathbb{P}^*$-martingala. Entonces, para cualquier derivado financiero X con fecha de ejercicio T y cualquier estrategia, $\varphi\in\Phi$ que lo *reproduzca*, tenemos para cada $t=0,1,...,T$,

$$\begin{aligned}\pi_X(t)=&\ V_\varphi(t)=\beta(t)^{-1}\tilde{V}_\varphi(t)\\ =&\ \beta(t)^{-1}\mathbb{E}^*(\tilde{V}_\varphi(T)|\mathcal{F}_t),\ (\text{pues } \tilde{V}_\varphi(t) \text{ es una } \mathbb{P}^*\text{-martingala})\\ =&\ \beta(t)^{-1}\mathbb{E}^*(\beta(T)V_\varphi(T)|\mathcal{F}_t),\ (\text{volviendo atrás el 'descuento'})\\ =&\ \beta(t)^{-1}\mathbb{E}^*(\beta(T)X|\mathcal{F}_t),\ (\text{pues } \varphi \textit{ reproduce} \text{ a } X)\end{aligned}$$

Una propiedad muy deseable del mercado $\mathcal{M}$, es el hecho de que todos los derivados financieros sean *alcanzables*, puesto que con esta condición se habrá resuelto el problema de precificación (de derivados financieros) de manera completa.

Definición 3.10. *Un mercado $\mathcal{M}$ se dice* completo *si cada derivado financiero es* alcanzable*, i.e. para cada variable aleatoria, no negativa, $\mathcal{F}_{\mathrm{T}}$-medible,* $\mathrm{X}\in\mathcal{X}^+$*, existe una estrategia autofinanciada que la reproduce,* $\varphi\in\Phi$*, tal que* $\mathrm{V}_\varphi(\mathrm{T})=\mathrm{X}$*.*

En el caso de un mercado $\mathcal{M}$ que no posee *oportunidades de arbitraje*, se puede insistir en reproducir los derivados financieros mediante una estrategia *admisible*, $\varphi\in\Phi_a$. Si φ es una estrategia autofinanciada y $\mathbb{P}^*$ es una *medida martingala equivalente* bajo la cual el proceso de precios *descontado* $\tilde{S}$ es una $\mathbb{P}^*$-martingala ($\mathbb{P}^*$ existe, por el teorema (2.1)), $\tilde{V}_\varphi(t)$ es también una $\mathbb{P}^*$-martingala (pues es la transformada martingala de $\tilde{S}$ por φ). Entonces

$$\tilde{V}_\varphi(t)=\mathbb{E}^*(\tilde{V}_\varphi(T)|\mathcal{F}_t),\ (t=0,1,...,T).$$

Si φ *reproduce* a X, $V_\varphi(T\ ) = X \geq 0$, luego, descontando, $\tilde{V}_\varphi(T) \geq 0$, entonces la ecuación anterior nos dice que $\tilde{V}_\varphi(t) \geq 0$ para cada t. Por lo tanto, todos los valores, en todos los instantes t, son no negativos, y así φ *admisible.*

Teorema 3.2. *(*Teorema de completitud*). Un mercado que no posee* oportunidades de arbitraje*, $\mathcal{M}$ es* completo *si y solo si existe una única medida de probabilidad $\mathbb{P}^*$ equivalente a $\mathbb{P}$ respecto a la cual los precios de los activos descontados son martingalas.*

Ahora veremos como modelar opciones de tipo americano. Nuevamente estamos ante contratos que otorgan, a su tenedor, el derecho a comprar o vender un determinado activo en el futuro, pero que a diferencia de las europeas pueden ser ejercidas a lo largo de toda su vida, i.e. desde que se celebra el contrato hasta su fecha de expiración . Luego, una opción de tipo *americano* se puede representar mediante un proceso estocástico adaptado no-negativo $Z = (Z_t)_{t=0,...,T}$, el cual se puede pensar como el beneficio obtenido por ejercer la opción en el periodo t. Si consideramos una opción put de tipo Bermuda con fechas de ejercicio $1, ..., T$, donde T es la fecha de expiración y con precio de ejercicio K, podemos escribir la opción como $Z_k = \max(K - S_k, 0)$, donde S_k no es otra cosa que el proceso estocástico adaptado que modela el precio del subyacente en los tiempos $k := 1, ..., T$. El valor de la opción Z en el k-ésimo periodo U_k está dado por la fórmula recursiva:

$$U_T = Z_T$$

$$U_{k-1} = \max\{Z_{k-1}, S^0_{k-1}\mathbb{E}^*(\frac{U_k}{S^0_k}|\mathcal{F}_{k-1})\}$$

donde el máximo se toma entre el resultado que se obtiene de ejecutar la opción y el valor esperado de lo que se obtendría si se la ejecutara en el futuro. (el supra-índice 0 indica que es el precio del activo libre de riesgo o *numéraire*)

Lema 3.2. $\tilde{U} = (\tilde{U}_k)_{k=0,...,\mathrm{T}}$ *,* $(\tilde{U}_k = \frac{1}{\mathrm{S}^0_k}U_k)$*es una* $\mathbb{P}^*$-supermartingala *y es la menor supermartingala que domina a* $\tilde{Z} = (\tilde{Z}_k)_{k=0,...,\mathrm{T}}$.

Dem: Es inmediato que $\tilde{U}_{k-1} = \max\{\tilde{Z}_{k-1}, \mathbb{E}^*(\tilde{U}_k|\mathcal{F}_{k-1})\} \geq \mathbb{E}^*(\tilde{U}_k|\mathcal{F}_{k-1})$ de donde $\tilde{U}$ es una $\mathbb{P}^*$-supermartingala. Ahora, sea $\tilde{R} = (\tilde{R}_k)_{k=1,...,T}$ una $\mathbb{P}^*$- supermartingala que domina a $\tilde{Z}$. Tenemos que $\tilde{R}_T \geq \tilde{Z}_T = \tilde{U}_T$. Además, $\tilde{R}_k \geq \tilde{U}_k$ implica que

$$\tilde{R}_{k-1} \geq \mathbb{E}^*(\tilde{R}_k|\mathcal{F}_{k-1}) \geq \mathbb{E}^*(\tilde{U}_k|\mathcal{F}_{k-1})$$

y como por hipótesis se tiene que
$\tilde{R}_{k-1} \geq \tilde{Z}_{k-1}$ resulta que $\tilde{R}_{k-1} \geq \max\{\tilde{Z}_{k-1}, \mathbb{E}^*(\tilde{U}_k|\mathcal{F}_{k-1})\} = \tilde{U}_{k-1}$. Luego por inducción inversa obtenemos lo que queríamos. $\square$

La manera en que definimos el valor de la opción de tipo americano, en su forma descontada, coincide con el concepto de *Cápsula de Snell.* A saber, sea $Z = (Z_k)_{k=1,...,T}$ un proceso estocástico adaptado no negativo. La cápsula de *Snell* de Z es el proceso estocástico $U = (U_k)_{k=1,...,T}$ definido recursivamente por

$$U_T = Z_T$$

$$U_{k-1} = \max\{Z_{k-1}, \mathbb{E}(U_k|\mathcal{F}_{k-1})\}$$

Luego, resulta del lema anterior que U es la menor supermartingala que domina a Z.

Lema 3.3. $\nu_0 = \inf\{k : U_k = Z_k\}$ *es un tiempo de parada y* $U^{\nu_0} = (U_{\nu_0 \wedge k})_{k=1,\ldots,\mathrm{T}}$ *es una martingala.*

Dem: Veamos que ν_0 es un tiempo de parada. ν_0 está bien definido, pues $U_T = Z_T$. Tenemos que

$$\{\nu_0 = 0\} = \{U_0 = Z_0\} \in \mathcal{F}_0$$
$$\{\nu_0 = k\} = \{U_0 > Z_0\} \cap \ldots \cap \{U_{k-1} > Z_{k-1}\} \cap \{U_k = Z_k\} \in \mathcal{F}_k$$

por lo tanto ν_0 es un tiempo de parada. Por otro lado tenemos que la representación para el proceso parado viene dado por:

$$U_k^{\nu_0} = U_{\nu_0 \wedge k} = U_0 + \sum_{j=1}^{k} 1_{\{j \leq \nu_0\}} \triangle U_j$$

Entonces considerando que si $m < \nu_0 \Rightarrow U_m > Z_m \Rightarrow U_m = \mathbb{E}(U_{m+1}|\mathcal{F}_m)$, tenemos

$$U_{k+1}^{\nu_0} - U_k^{\nu_0} = 1_{\{k+1 \leq \nu_0\}}(U_{k+1} - U_k) = 1_{\{k+1 \leq \nu_0\}}(U_{k+1} - \mathbb{E}(U_{k+1}|\mathcal{F}_k))$$

luego tomando esperanza condicional resulta que

$$\mathbb{E}(U_{k+1}^{\nu_0} - U_k^{\nu_0}|\mathcal{F}_k) = 1_{\{k+1 \leq \nu_0\}}\mathbb{E}((U_{k+1} - \mathbb{E}(U_{k+1}|\mathcal{F}_k))|\mathcal{F}_k) = 0$$

es decir, U^{ν_0} es una martingala. $\square$

Notación: $\Upsilon_{k,T} = \{\tau$ tiempo de parada :$Im(\tau) \subset \{k, k+1, \ldots, T\}\}$

Corolario 3.2. *El tiempo de parada* ν_0 *verifica*

$$U_0 = \mathbb{E}(Z_{\nu_0}|\mathcal{F}_0) = \sup_{\tau \in \Upsilon_{0,T}} \mathbb{E}(Z_\tau|\mathcal{F}_0)$$

Dem : Sabemos que U^{ν_0} es una martingala, entonces tenemos que

$$U_0 = U_0^{\nu_0} = \mathbb{E}(U_T^{\nu_0}|\mathcal{F}_0) = \mathbb{E}(U_{\nu_0}|\mathcal{F}_0) = \mathbb{E}(Z_{\nu_0}|\mathcal{F}_0)$$

Por otro lado, si $\tau \in \Upsilon_{0,T}$ como U^τ es una supermartingala tenemos que

$$U_0 \geq \mathbb{E}(U_T^\tau|\mathcal{F}_0) = \mathbb{E}(U_\tau|\mathcal{F}_0) \geq \mathbb{E}(Z_\tau|\mathcal{F}_0)$$

luego $U_0 = \sup_{\tau \in \Upsilon_{0,T}} \mathbb{E}(Z_\tau | \mathcal{F}_0)$. □

Es posible obtener la siguiente generalizacin de los resultados anteriores.

Corolario 3.3. *$\nu_n = \inf\{j \geq k : U_j = Z_j\}$ es un tiempo de parada y $U^{\nu_k} = (U_{\nu_k \wedge m})_{m=k,\ldots,\mathrm{T}}$ es una martingala y*

$$U_k = \mathbb{E}(Z_{\nu_k} | \mathcal{F}_k) = \sup_{\tau \in \Upsilon_{k,T}} \mathbb{E}(Z_\tau | \mathcal{F}_k)$$

Dem: análoga a las del caso $k = 0$.

Definición 3.11. *Un tiempo de parada ν es optimal para un proceso adaptado $Z = (Z_k)_{k=1,\ldots,T}$ si*

$$\mathbb{E}(Z_\nu | \mathcal{F}_0) = \sup_{\tau \in \Upsilon_{0,T}} \mathbb{E}(Z_\tau | \mathcal{F}_k)$$

Claramente ν_0 es optimal para Z.

4. El LS-algoritmo

Suponemos un espacio de probabilidad completo $(\Omega, \mathcal{F}, \mathbb{P})$ y un horizonte temporal finito $[0, T]$, donde el espacio muestral Ω es el conjunto de todos los posibles estados que puede adoptar una economía aleatoria entre los tiempos 0 y T y donde $\omega \in \Omega$ representa un estado determinado. $\mathcal{F}$ es la σ-álgebra de eventos distinguibles al momento T y $\mathbb{P}$ es una medida de probabilidad definida sobre los elementos de $\mathcal{F}$. Sea $(\mathcal{F}_k)_{k=0,\ldots,T}$ una filtración que cumple $\mathcal{F}_T = \mathcal{F}$. Siendo consistentes con la no existencia de oportunidades de arbitraje, asumimos la existencia de una medida mutuamente absolutamente continua $\mathbb{Q}$ en esta economía.
Nos interesa la presificación de derivados financieros de estilo americano con flujos de efectivo aleatorios que puedan ocurrir a lo largo del periodo temporal fijado.
Por lo tanto consideramos un proceso estocástico adaptado $Z = (Z_k)_{k=0,\ldots,T}$, no negativo (este proceso representa el resultado obtenido de ejecutar la opción en el instante n). De aquí en más asumimos que $Z_k \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ para todo $k = 0, ..., T$, i.e. el espacio de funciones de varianza finita. Como ya se vió, la cápsula de Snell -que representa el valor de la opción en su forma descontada- $U = (U_k)_{k=0,\ldots,T}$ para Z está definida recursivamente por

$$U_T = Z_T$$

$$U_{k-1} = \max\{Z_{k-1}, \mathbb{E}(U_k|\mathcal{F}_{k-1})\}$$

y verifica que

$$U_k = \sup_{\tau \in \Upsilon_{k,T}} \mathbb{E}(Z_\tau|\mathcal{F}_k)$$

también tenemos que $U_k = \mathbb{E}(Z_{\nu_k}|\mathcal{F}_k)$ donde

$$\nu_k = \inf\{j \geq k : U_j = Z_j\}$$

Nota: Vamos a focalizarnos en el caso en que la opción americana puede ser solamente ejecutada en los K tiempos discretos $0 < t_1 \leq t_2 \leq t_3 \leq ... \leq t_K = T$, y consideraremos la óptima política de parada en cada fecha de ejercicio t_i. En la práctica, por supuesto, existen opciones americanas continuamente ejecutables, sin embargo el LS-algoritmo puede ser utilizado para aproximar el valor de estas opciones simplemente tomando K lo suficientemente grande.

En la fecha de expiración de la opción T, el inversor ejecuta la opción si esta está *in the money* (i.e. si la opción es ejecutada proporcionara a su tenedor un flujo de fondos positivo, una ganancia), o le permite expirar si se encuentra *out of the money*(i.e. análogamente, si la opción es ejecutada proporcionara a su tenedor un flujo de fondos negativo, una perdida, pero claramente, dado que una opción es un derecho, no una obligación para su tenedor, en este caso no la ejecutará). En el momento $k < T$, sin embargo, el inversor debe elegir entre ejecutar la opción inmediatamente o continuar la vida de la opción y nuevamente en la próxima fecha de ejercicio, rever su situación. El valor de la opción es maximizado en cada camino y el inversor ejecuta la opción tan pronto como el valor de ejecución inmediata de la opción es mayor o igual al de continuar con la vida de la misma.

Al momento k, el flujo de efectivo de una ejecución inmediata es conocido para el inversor. El flujo de efectivo asociado a la continuación de la vida de la opción, obviamente, no es conocido en n. No obstante, la teoría de la valución libre de arbitraje, implica que el valor de continuar o equivalentemente, el valor de la opción sujeta a no ser ejecutada hasta después de n, está dado por la esperanza de los restantes flujos de efectivo descontados, $\mathbb{E}(U_{k+1}|\mathcal{F}_k)$ con respecto a la medida $\mathbb{Q}$.

Podemos escribir, entonces, recursivamente los tiempos de parada optimales $\{\nu_k\}_{k=0,...,T}$ por

- $\nu_T = T$
- $\nu_k = k1_{\{Z_k \geq \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}} + \nu_{k+1}1_{\{Z_k < \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}}$

Nota : en la definición de los tiempos de parada se utilizó la siguiente identidad $\mathbb{E}(U_{k+1}|\mathcal{F}_k) = \mathbb{E}(\mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_{k+1})|\mathcal{F}_k) = \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)$

Aquí asumimos que $S = (S_k)_{k=0,...,\infty}$ es una cadena de Markov homogénea con espacio de estados E y matriz de transición P. Sean $\psi : \{0, 1, ..., T\} \times E \longrightarrow \mathbb{R}_{\geq 0}$ y $Z = (Z_k = \psi(k, S_k))_{k=0,...,T}$ un proceso adaptado no negativo. Entonces la cápsula de Snell $U = (U_k)_{k=0,...,T}$ tiene la forma $U_k = u\ (k, S_k)$ donde la función $u : \{0, 1, ..., T\} \times E \longrightarrow \mathbb{R}_{\geq 0}$ es definida recursivamente por

$$u\ (T, s) = \psi(T, s)$$
$$u\ (k, s) = \max\{\psi(k, s), Pu\ (k + 1, s)\}$$

Como estamos con una cadena de Markov sabemos que vale la propiedad fuerte de Markov, esto es

$$\mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k) = \mathbb{E}(Z_{\nu_{k+1}}|S_k)$$

Nota : donde $S = (S_k)_{k=0,...,\infty}$ es el proceso de precios del subyacente.

El LS-algoritmo introducido por F. Longstaff y E. Schwartz en 2001, utiliza las ecuaciones de los tiempos de parada optimales $\{\nu_k\}_{n=0,...,T}$ para calcular recursivamente ν_1 y así $U_0 = \mathbb{E}(Z_{\nu_0}|S_0) = \max(Z_0, \mathbb{E}(Z_{\nu_1}|S_0))$. El problema es reducido, entonces a calcular las esperanzas condicionales $\mathbb{E}(Z_{\nu_{k+1}}|S_k)$. Lo primero que plantea el algoritmo es la aproximación de $\mathbb{E}(Z_{\nu_{k+1}}|S_k)$ por la proyección ortogonal sobre el espacio generado por un número finito de funciones de S_k. Consideremos entonces una sucesión $\{\phi_m(S_k)\}_{m=0,...}$ de funciones medibles a valores reales, definidas sobre E y que satisfacen las siguientes condiciones

A_1**:** Para $k = 1, ..., T-1$, la sucesión $\{\phi_m(S_k)\}_{m=1,...}$ es total en $L^2(\sigma(S_k))$.
A_2**:** Para $k = 1, ..., T-1$ y $M \geq 1$, si $\sum_{m=0}^{M} a_m^n \phi_m(S_k) = 0$ casi seguramente luego $a_m^n = 0$ para $k = 1, ..., M$.

Denotamos, para $k = 1, ..., T-1$, la proyección ortogonal de $L^2(\Omega)$ sobre el espacio generado por $\{\phi_1(S_k), ..., \phi_M(S_k)\}$ por medio de P_k^M. Luego tenemos los tiempos de parada ν_k^M obtenidos de sustituir U_{k+1} por sus aproximaciones $P_k^M(Z_{\nu_{k+1}^M})$ en el sistema de ecuaciones que definen los tiempos de parada optimales,

- $\nu_T^M = T$

- $\nu_k^M = k1_{\{Z_k \geq P_k^M(Z_{\nu_{k+1}^M})\}} + \nu_{k+1}^M 1_{\{Z_k < P_k^M(Z_{\nu_{k+1}^M})\}}$

Nota: observar que $P_k^M(Z_{\nu_{k+1}^M}) = P_k^M(\mathbb{E}(Z_{\nu_{k+1}^M}|\mathcal{F}_k))$ y claramente $P_k^M(\mathbb{E}(Z_{\nu_{k+1}^M}|\mathcal{F}_k))$ tiende a $\mathbb{E}(Z_{\nu_{k+1}^M}|\mathcal{F}_k)$ cuando M tiende a infinito.

es claro que podemos escribir (más adelante explicitamos la forma de los coeficientes),

$$P_k^M(Z_{\nu_{k+1}^M}) = \sum_{m=1}^{M} a_k^m \phi_m(S_k)$$

De estos tiempos de parada, obtenemos una aproximación del valor de la función U_0, dada por:

$$U_0^M = max(Z_0, \mathbb{E}(Z_{\nu_1^M}))$$

claramente $\mathbb{E}(\mathbb{E}(Z_{\nu_1^M}|\mathcal{F}_0)) = \mathbb{E}(Z_{\nu_1^M})$
Luego resta aproximar $E(Z_{\nu_1^M})$ mediante simulación.
Para ello, asumimos que podemos simular $N > M$ caminos ω_n denotados por

$$S(\omega_n) := \{S_k(\omega_n) : k = 0, ..., T\}$$

Asumimos además que $S_0(\omega_n) = s$ para todo $n = 1, ..., N$. Los coeficientes $\alpha_k^{n,M,N}$ son determinados por algún mecanismo de estimación a partir de los N caminos simulados. Luego con estos coeficientes estimados construimos las aproximaciones,

$$\sum_{n=1}^{M} \alpha_k^{n,M,N} \phi_n(S_k(\omega_n))$$

de

$$P_k^M(Z_{\nu_{n+1}^M})$$

y podemos estimar los tiempos de parada

- $\hat{\nu}_T^{n,M,N} = T$
- $\hat{\nu}_k^{n,M,N} = k1_{\{Z_k \geq \sum_{n=1}^{M} \alpha_k^{n,M,N} \phi_n(S_k(\omega_n))\}} + \hat{\nu}_{k+1}^{n,M,N} 1_{\{Z_k < \sum_{n=1}^{M} \alpha_k^{n,M,N} \phi_n(S_k(\omega_n))\}}$

Finalmente a partir de estas estimaciones podemos derivar la siguiente aproximación de a U_0:

$$\hat{U}_0^{M,N} = \max(Z_0, \frac{1}{N} \sum_{n=1}^{N} Z_{\hat{\nu}_1^{n,M,N}}(\omega_n))$$

El mecanismo de estimación del LS-algoritmo es el de mínimos cuadrados, esto es obtener los coeficientes $\alpha_k^{n,M,N}$ como solución del siguiente problema de minimización:

$$\alpha_k^{M,N} = \arg \min_{a \in \mathbb{R}^M} \sum_{n=1}^{N} (Z_{\hat{\nu}_{k+1}^{n,M,N}}(\omega_n) - a \cdot \phi^M(S_k(\omega_n)))^2$$

Donde para $M \geq 1$ denotamos con $\phi^M(S_k) := (\phi_1(S_k), \phi_2(S_k), ..., \phi_M(S_k))$ y $\alpha_k^{M,N}$ es el vector formado por los coeficientes $\alpha_k^{n,M,N}$.

Nota: Cabe mencionar que el algoritmo aquí presentado no es exactamente el LSM, dado que este último solo considera caminos "in the money" en sus regresiones . Esto parece arribar a resultados más eficientes.

5. Convergencia del algoritmo

Como ya vimos, el algoritmo LSM presenta dos aproximaciones. La primera, reemplazar las esperanzas condicionales por sus proyecciones en un conjunto finito de funciones tomadas de una base apropiada. La segunda, utilizar simulaciones y regresión por mínimos cuadrados para computar el valor de las funciones de la primera aproximación. En consecuencia veremos dos resultados que muestran que para cualquier M fijo, $U_0^{M,N}$ converge casi seguramente a U_0^M cuando N tiende a ∞, y que U_0^M converge a U_0 cuando M tiende a ∞. Para esto seguiremos [7].
Comenzaremos introduciendo la notación que utilizaremos.

Para $M \geq 1$ denotamos con $\phi^M(S_k) := (\phi_0(S_k), \phi_1(S_k), ..., \phi_{M-1}(S_k))$ y para $k = 1, ..., T-1$ definimos $\alpha_k^M \in \mathbb{R}^M$ tal que

$$P_k^M(Z_{\nu_{k+1}^M}) = \alpha_k^M \cdot \phi^M(S_k)$$

Notemos que, bajo la condición A_2, el parámetro M-dimensional α_k^M tiene la siguiente forma explicita,

$$\alpha_k^M = (A_k^M)^{-1}\mathbb{E}(Z_{\nu_{k+1}^M}\phi^M(S_k))$$

para $k = 1, ..., T-1$, donde $A_k^M \in \mathbb{R}^{M\times M}$ está definida por,

$$(A_k^M)_{1\leq j,l\leq M} = \mathbb{E}(\phi_j(S_k)\phi_l(S_k))$$

Similarmente, los estimadores $\alpha_k^{(M,N)}$ están dados por

$$\alpha_k^{(M,N)} = (A_k^{(M,N)})^{-1}\frac{1}{N}\sum_{n=1}^{N} Z_{\hat{\nu}_{k+1}^{n,M,N}}(\omega_n)\phi^M(S_k)(\omega_n)$$

para $k = 1, ..., T-1$, donde $A_k^{(M,N)} \in \mathbb{R}^{M\times M}$ está definida por,

$$(A_k^{(M,N)})_{1\leq j,l\leq M} = \frac{1}{N}\sum_{n=1}^{N}\phi_j(S_k(\omega_n))\phi_l(S_k(\omega_n))$$

Nótese que $\lim_{N\longrightarrow\infty} A_k^{(M,N)} = A_k^M$ casi seguramente. Por lo tanto, bajo A_2, la matriz $A_k^{(M,N)}$ es invertible para N lo suficientemente grande. Definimos además, $\alpha^M := (\alpha_1^M, \alpha_1^M, ..., \alpha_{T-1}^M)$ y $\alpha^{(M,N)} := (\alpha_1^{(M,N)}, \alpha_1^{(M,N)}, ..., \alpha_{T-1}^{(M,N)})$.
Dado un parámetro $a^M := (a_0^M, a_1^M, ..., a_{T-1}^M) \in \mathbb{R}^M \times ... \times \mathbb{R}^M$ y dos vectores determinísticos $z = (z_1, ..., z_T) \in \mathbb{R}^T$ y $x = (x_1, ..., x_T) \in E^T$, definimos el siguiente campo de vectores $F = (F_1, ..., F_T)$ por

$$F_T(a^M, z, x) = z_T$$
$$F_k(a^M, z, x) = z_k 1_{\{z_k \geq a_k^M\phi^M(x_k)\}} + F_{k+1}(a^M, z, x)1_{\{z_k < a_k^M\phi^M(x_k)\}}$$

para $k = 1, ..., T-1$.
Tenemos que

$$F_k(a^M, z, x) = z_k 1_{B_k^c} + \sum_{i=k+1}^{T-1} z_i 1_{B_k \dots B_{i-1}B_i^c} + z_T 1_{B_k \dots B_{T-1}}$$

con

$$B_j = \{z_k < a_k^M \phi^M(x_k)\}.$$

Observemos que tenemos las siguientes igualdades,

$$F_k(\alpha^M, Z, S) = Z_{\nu_k^M}$$
$$F_k(a^{(M,N)}, Z(\omega_n), S(\omega_n)) = Z_{\hat{\nu}_k^{n,M,N}}$$

Para $k = 2, ..., T$, denotamos con G_k la siguiente función

$$G_k(a^M, z, x) = F_k(a^M, z, x)\phi^M(x_{k-1}),$$

y definimos las funciónes ϕ_k y ψ_k por

$$\phi_k(a^M) = \mathbb{E}(F_k(a^M, Z, S))$$
$$\psi_k(a^M) = \mathbb{E}(G_k(a^M, Z, S))$$

Observemos que con esta notación, tenemos

$$\alpha_k^M = (A_k^M)^{-1}\psi_{k+1}(a^M)$$

y similarmente, para $k = 1, ..., T-1$,

$$\alpha_k^{(M,N)} = (A_k^{(M,N)})^{-1}\frac{1}{N}\sum_{n=1}^{N} G_{k+1}(a^{(M,N)}, Z(\omega_n), S(\omega_n)))$$

Finalmente estamos en condiciones de abocarnos al estudio de los teoremas de convergencia.

Teorema 5.1. *Asumamos que se satisface A_1 de la sección 1, luego, para $k = 1$ hasta T, tenemos que*

$$\lim_{M\longrightarrow+\infty} \mathbb{E}(Z_{\nu_k^M}|\mathcal{F}_k) = \mathbb{E}(Z_{\nu_k}|\mathcal{F}_k)$$

en L^2.

Dem: Procedemos por inducción sobre k. El resultado es verdadero para $k = T$. Provemos que si se cumple para $k+1$, entonces vale para k ($k \leq T-1$). Dado que

$$Z_{\nu_k^M} = Z_k 1_{\{Z_k \geq P_k^M(Z_{\nu_{k+1}^M})\}} + Z_{\nu_{k+1}^M} 1_{\{Z_k < P_k^M(Z_{\nu_{k+1}^M})\}}$$

para $k \leq T-1$, tenemos

$$\begin{aligned} &\mathbb{E}(Z_{\nu_k^M} - Z_{\nu_k}|\mathcal{F}_k) = \\ &= \mathbb{E}(Z_k 1_{\{Z_k \geq P_k^M(Z_{\nu_{k+1}^M})\}} + Z_{\nu_{k+1}^M} 1_{\{Z_k < P_k^M(Z_{\nu_{k+1}^M})\}} - Z_k 1_{\{Z_k \geq \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}} - \\ &- Z_{\nu_{k+1}} 1_{\{Z_k < \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}}|\mathcal{F}_k) = \\ &= \mathbb{E}(Z_k 1_{\{Z_k \geq P_k^M(Z_{\nu_{k+1}^M})\}} + Z_{\nu_{k+1}^M} 1_{\{Z_k < P_k^M(Z_{\nu_{k+1}^M})\}} - Z_k 1_{\{Z_k \geq \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}} - \\ &- Z_{\nu_{k+1}} 1_{\{Z_k \geq P_k^M(Z_{\nu_{k+1}^M})\}} - \end{aligned}$$

$$-Z_{\nu_{k+1}}1_{\{Z_k<P_k^M(Z_{\nu_{k+1}^M})\}} - Z_{\nu_{k+1}}1_{\{Z_k\geq \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}}|\mathcal{F}_k) =$$
$$= (Z_k - \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k))(1_{\{Z_k\geq P_k^M(Z_{\nu_{k+1}^M})\}} - 1_{\{Z_k\geq \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}})+$$
$$+\mathbb{E}(Z_{\nu_{k+1}^M} - Z_{\nu_{k+1}}|\mathcal{F}_k)1_{\{Z_k<P_k^M(Z_{\nu_{k+1}^M})\}}$$

de acuerdo a nuestra hipótesis inductiva, el segundo término del lado derecho de la igualdad converge a 0. Entonces solo debemos probar que B_k^M definido por

$$B_k^M = (Z_k - \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k))(1_{\{Z_k\geq P_k^M(Z_{\nu_{k+1}^M})\}} - 1_{\{Z_k\geq \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}})$$

converge a 0 en L^2. Observemos que

$$|B_k^M| = |Z_k - \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)||1_{\{\mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)>Z_k\geq P_k^M(Z_{\nu_{k+1}^M})\}} - 1_{\{P_k^M(Z_{\nu_{k+1}^M})>Z_k\geq \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\}}|$$
$$\leq |Z_k - \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)|1_{\{|Z_k-\mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)|\leq|P_k^M(Z_{\nu_{k+1}^M})-\mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)|\}}$$
$$\leq |P_k^M(Z_{\nu_{k+1}^M}) - \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)|$$
$$\leq |P_k^M(Z_{\nu_{k+1}^M}) - P_k^M(\mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k))| + |P_k^M(\mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)) - \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)|$$

pero tenemos que

$$P_k^M(Z_{\nu_{k+1}^M}) = P_k^M(\mathbb{E}(Z_{\nu_{k+1}^M}|\mathcal{F}_k))$$

pues es la proyección ortogonal en un subespacio del espacio de variables aleatorias $\mathcal{F}_k$-medible y tenemos que $\langle Z_{\nu_{k+1}^M}, \phi_m(s_k)\rangle = \langle \mathbb{E}(Z_{\nu_{k+1}^M}|\mathcal{F}_k), \phi_m(s_k)\rangle$, por la definición de esperanza condicional. Por lo tanto tenemos que

$$\|B_k^M\|_2 \leq \|\mathbb{E}(Z_{\nu_{k+1}^M}|\mathcal{F}_k) - \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\|_2 + \|P_k^M(\mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)) - \mathbb{E}(Z_{\nu_{k+1}}|\mathcal{F}_k)\|_2$$

Donde el primer término de la derecha converge a 0 por la hipótesis inductiva y el segundo término lo hace por A_1. □

En lo que sigue, fijamos M, y vemos el resultado de $U_0^{M,N}$ a medida que N, el número de caminos simulados, va a infinito. Para simplificar la notación eliminamos el indice M por el resto del capítulo.

Teorema 5.2. *Asumamos que para $k = 1, ..., T-1$, $\mathbb{P}(\alpha_k \cdot \phi(S_k) = Z_k) = 0$. Luego $U_0^{M,N} = \max(Z_0, \frac{1}{N}\sum_{n=1}^N Z_{\hat{\nu}_1^{n,M,N}}(\omega_n))$ converge casi seguramente a $U_0^M = \max(Z_0, \mathbb{E}(Z_{\nu_1^M}))$ cuando N tiende a infinito. Además tenemos convergencia casi segura de $\frac{1}{N}\sum_{n=1}^N Z_{\hat{\nu}_k^{n,N,M}}(\omega_n)$ hacia $\mathbb{E}(Z_{\nu_k^M})$ cuando N tiende a infinito, para $k = 1, ..., T$.*

Notemos que bajo la nueva notación introducida previamente, tenemos que probar lo siguiente,

$$\lim_{N\longrightarrow\infty} \frac{1}{N}\sum_{n=1}^N F_k(a^{(N)}, Z(\omega_n), S(\omega_n)) = \phi_k(\alpha)$$

para $1 \leq k \leq T$

La demostración de este teorema se basa en los siguientes lemas que presentamos a continuación.

Lema 5.1. *Para $k = 1, ..., T-1$, tenemos:*

$$|F_k(a, Z, S) - F_k(b, Z, S)| \leq (\sum_{i=k}^{T} |Z_i|)(\sum_{i=k}^{T-1} 1_{\{|Z_i - b_i\phi(S_i)| \leq |a_i - b_i||\phi(S_i)|\}})$$

Dem: Sea $B_j = \{Z_k < a_k\phi(S_k)\}$ y $\tilde{B}_j = \{Z_k < b_k\phi(S_k)\}$. Tenemos,

$$F_k(a, Z, S) - F_k(b, Z, S) = Z_k(1_{B_k^c} - 1_{\tilde{B}_k^c}) + \sum_{i=k+1}^{T-1} Z_i(1_{B_k \ldots B_{i-1}B_i^c} - 1_{\tilde{B}_k \ldots \tilde{B}_{i-1}\tilde{B}_i^c}) +$$

$$+ Z_T(1_{B_k \ldots B_{T-1}} - 1_{\tilde{B}_k \ldots \tilde{B}_{T-1}})$$

Pero tenemos que

$$|1_{B_k^c} - 1_{\tilde{B}_k^c}| = 1_{\{a_k\phi(S_k) \leq Z_k < b_k\phi(S_k)\}} + 1_{\{b_k\phi(S_k) \leq Z_k < a_k\phi(S_k)\}} \leq$$

$$\leq 1_{\{|Z_k - b_k\phi(S_k)| \leq |a_k - b_k||\phi(S_k)|\}}$$

más aún, tenemos que

$$|1_{B_k \ldots B_{i-1}B_i^c} - 1_{\tilde{B}_k \ldots \tilde{B}_{i-1}\tilde{B}_i^c}| \leq \sum_{j=k}^{i-1} |1_{B_j^c} - 1_{\tilde{B}_j^c}| + |1_{B_i^c} - 1_{\tilde{B}_i^c}| =$$

$$= \sum_{j=k}^{i} |1_{B_j^c} - 1_{\tilde{B}_j^c}|$$

Combinando estas desigualdades obtenemos lo que queríamos probar. □

Lema 5.2. *Asumamos que para $k = 1, ..., T-1$, $\mathbb{P}(\alpha_k \cdot \phi(S_k) = Z_k) = 0$, luego α_k^N converge casi seguramente a a_k cuando N tiende a infinito.*

Dem: Procedemos por inducción sobre k. Para $k = T-1$, el resultado es una consecuencia directa de la ley de los grandes números. Ahora asumimos que el resultado es cierto para $i = k, ..., T-1$. Queremos ver que es cierto para $k-1$. Tenemos que

$$\alpha_{k-1}^{(N)} = (A_{k-1}^{(N)})^{-1} \frac{1}{N} \sum_{n=1}^{N} G_k(\alpha^{(N)}, Z(\omega_n), S(\omega_n)))$$

Por la ley de los grandes números, sabemos que $A_{k-1}^{(N)}$ converge casi seguramente a A_{k-1}, por lo que resta probar es que $\frac{1}{N}\sum_{n=1}^{N} G_k(\alpha^{(N)}, Z(\omega_n), S(\omega_n)))$ converge a $\psi_k(\alpha)$. De la ley de los grande números, tenemos la convergencia de $\frac{1}{N}\sum_{n=1}^{N} G_k(\alpha, Z(\omega_n), S(\omega_n)))$ a $\psi_k(\alpha)$, por lo que es suficiente para probar lo que queremos, ver que:

$$\lim_{N \longrightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} G_k(\alpha^{(N)}, Z(\omega_n), S(\omega_n)) - G_k(\alpha, Z(\omega_n), S(\omega_n))) = 0$$

Notemos con $G_N = \frac{1}{N} \sum_{n=1}^{N} G_k(\alpha^{(N)}, Z(\omega_n), S(\omega_n)) - G_k(\alpha, Z(\omega_n), S(\omega_n)))$. Tenemos que

$$|G_N| \leq \frac{1}{N} \sum_{n=1}^{N} |\phi(S_{k-1}(\omega_n))||F_k(\alpha^{(N)}, Z(\omega_n), S(\omega_n)) - F_k(\alpha, Z(\omega_n), S(\omega_n)))| \leq$$

$$\leq \frac{1}{N} \sum_{n=1}^{N} |\phi(S_{k-1}(\omega_n))|((\sum_{i=k}^{T} |Z_i(\omega_n)|)(\sum_{i=k}^{T-1} 1_{\{|Z_i(\omega_n) - \alpha_i \phi(S_i(\omega_n))| \leq |\alpha_i^{(N)} - \alpha_i||\phi(S_i(\omega_n))|\}}))$$

donde la segunda desigualdad vale, claramente, por el lema anterior. Puesto que para $i = k, ..., T-1$, $\alpha_i^{(N)}$ converge casi seguramente a α_i, tenemos que para cada $\varepsilon > 0$:

$$\limsup |G_N| \leq \limsup \frac{1}{N} \sum_{n=1}^{N} |\phi(S_{k-1}(\omega_n))|((\sum_{i=k}^{T} |Z_i(\omega_n)|)$$
$$(\sum_{i=k}^{T-1} 1_{\{|Z_i(\omega_n) - \alpha_i \phi(S_i(\omega_n))| \leq \varepsilon |\phi(S_i(\omega_n))|\}})) =$$
$$= \mathbb{E}(|\phi(S_{k-1})|)((\sum_{i=k}^{T} |Z_i|)(\sum_{i=k}^{T-1} 1_{\{|Z_i - \alpha_i \phi(S_i)| \leq \varepsilon |\phi(S_i)|\}}))$$

donde la última igualdad sigue de la ley de los grandes números. Ahora haciendo tender ε a 0, obtenemos la convergencia a cero, puesto que para $k = 1, ..., T-1$, $\mathbb{P}(\alpha_k \cdot \phi(S_k) = Z_k) = 0$. $\square$

6. Ejemplo LSM

A continuación presentamos un simple ejemplo numérico extraido del artículo de Longstaff y Schwartz, "Valuing American Options by Simulation: A simple Least-Squares Approach", que nos permitirá comprender la escencia del algoritmo. Así mismo, incluimos los resultados obtenidos al utilizar otras bases en la estimación de las funciones de esperanza condicional, para comprobar de alguna manera los que se sugiere en dicho artículo.
En la fecha de expiración (T), la estrategia óptima para una opción de tipo americana es ejercer la opción si se encuentra "in the money". Previo a esta fecha, sin embargo, la estrategia óptima surge de comparar el valor del ejercicio inmediato de la opción, con el flujo de fondos descontado obtenido por continuar (i.e. no ejercerla inmediatamente), para luego, ejercerla si el valor de la primera alternativa es mayor. Por lo tanto, el quid del ejercicio óptimo de una opción es identificar la esperanza condicional del valor de continuar. Este método se vale de la información en los caminos simulados para aproximar dicha esperanza condicional.
Consideremos un put americano escrito sobre una acción que no paga dividendos. El precio de ejercicio es de \$1.10 y las fechas de ejercicio son $t = 1, 2, 3$, donde el instante tres es la fecha de expiración de la opción (i.e. T). La tasa de interés libre de riesgo es de 0.06. En este ejemplo se utilizan solamente ocho caminos del precio del subyacente, simulados con la medida neutral al riesgo, para ilustrar el uso del algoritmo. Así obtenemos la siguiente tabla:

Caminos del precio de la acción

Caminos	$t=0$	$t=1$	$t=2$	$t=3$
1	1.00	1.09	1.08	1.34
2	1.00	1.16	1.26	1.54
3	1.00	1.22	1.07	1.03
4	1.00	0.93	0.97	0.92
5	1.00	1.11	1.56	1.52
6	1.00	0.76	0.77	0.90
7	1.00	0.92	0.84	1.01
8	1.00	0.88	1.22	1.34

Nuestro objetivo es encontrar los tiempos de parada que maximizan el valor de la opción en cada instante a lo largo cada camino. Dado que el algoritmo es recursivo, procedemos de la siguiente manera. Condicional a no ejercer la opción antes de la fecha de expiración (instante 3), el único curso de acción que tenemos es ejecutar la opción si está "in the money", o dejarla expirar si no lo está. Los flujos de fondos realizados por el tenedor de la opción (de acuerdo a la estrategia óptima) se presentan en la siguiente tabla, Flujos de fondos en $t = 3$

Caminos	$t = 1$	$t = 2$	$t = 3$
1	-	-	0.00
2	-	-	0.00
3	-	-	0.07
4	-	-	0.18
5	-	-	0.00
6	-	-	0.20
7	-	-	0.09
8	-	-	0.00

Estos resultados son idénticos a aquellos que se hubiesen percibido de ser la opción de tipo europea en vez de americana.

Si el put se encuentra "in the money" en el instante 2, el tenedor de la opción debe decidir si la ejerce inmediatamente o continua para ejercerla más adelante -i.e. en $t = 3$- (en el caso en el que el put no se encuentra "in the money" en $t = 2$, el tenedor no tiene otra alternativa más que seguir). De la tabla de los precios de la acción tenemos que solo hay cinco caminos para los cuales el put se encuentra "in the money" en el instante 2. Denotemos con X los precios de la acción en $t = 2$ para estos cinco caminos y con Y los respectivos flujos de fondos descontados obtenidos en $t = 3$ condicional a que la opción no sea ejercida antes de dicha fecha. Solo se utilizan los caminos para los que la opción está "in the money" ya que esto nos permite estimar la esperanza condicional en la región en la que el ejercicio es relevante y mejora significativamente la eficiencia del algoritmo. Los vectores X e Y se presentan en la siguiente tabla.

Regresión en $t = 2$

Caminos	Y	X
1	0.00×0.94176	1.08
2	-	-
3	0.07×0.94176	1.07
4	0.18×0.94176	0.97
5	-	-
6	0.20×0.94176	0.77
7	0.09×0.94176	0.84
8	-	-

donde 0.94176 es el factor de descuento.
Para estimar la esperanza de los flujos de fondos descontados, generados por continuar la vida de la opción, condicionada al precio de la acción en el instante 2, realizamos la regresión de Y respecto a una constante, X y X^2. Esta expanción es una de las más simples posibles , más adelante se especifican los resultados que se obtienen al utilizar otras bases como Hermite y Laguerre . La función de la esperanza condicional queda caracterizada entonces por $\mathbb{E}(Y|X) \simeq -1.070 + 2.983X - 1.813X^2$. Mediante esta función calculamos el valor de continuar, para luego compararlo con el valor de ejercer la opción inmediatamente en $t = 2$. Así construimos la siguiente tabla.

Decisión óptima de continuar en $t = 2$

Caminos	Ejercer	Continuar
1	0.02	0.0369
2	-	-
3	0.03	0.0461
4	0.13	0.1176
5	-	-
6	0.33	0.1520
7	0.26	0.1565
8	-	-

El valor de ejercer inmediatamente es claramente $1.1 - X$, para los caminos en que la opción está "in the money", mientras que el valor de continuar está dado por especializar nuestra función en X. La comparación de ambos valores implica que es óptimo ejercer la opción en $t = 2$ para los caminos cuatro, seis y siete. Esto nos lleva a la siguiente matriz, la cual nos muestra los flujos de fondos percibidos por el tenedor de la opción, condicional a no ejercer la opción antes de $t = 2$.

Flujos de fondos en $t = 2$

Caminos	$t = 1$	$t = 2$	$t = 3$
1	-	0.00	0.00
2	-	0.00	0.00
3	-	0.00	0.07
4	-	0.13	0.00
5	-	0.00	0.00
6	-	0.33	0.00
7	-	0.26	0.00
8	-	0.00	0.00

Notar que cuando la opción es ejercida en el momento 2, el flujo de fondos en la última columna se vuelve cero. Esto responde a que una vez que la opción es ejercida no hay más flujos de fondos puesto que la opción solo puede ser ejercida una vez a lo largo de su vida.

Procediendo recursivamente, ahora analizamos si la opción debe ser ejercida en $t = 1$. De la tabla de los precios de la acción, nuevamente observamos que existen únicamente cinco caminos para los cuales en el momento 1 la opción se encuentra "in the money". Para estos caminos, definimos Y como el valor descontado de los subsecuentes flujos de fondos asociados a la opción. (Notar que en la definición de Y, utilizamos los flujos de fondos realizados a lo largo de cada camino, no el valor de la esperanza condicional de Y estimada en $t = 2$.).

Dado que la opción solo puede ser ejercida una vez, los flujos de fondos futuros, para cada camino, ocurren en el instante 2 o en el 3, pero no en ambos. Claramente, los flujos de fondos recibidos en $t = 2$ se descuentan un solo periodo, mientras que los recibidos en $t = 3$ son descontados dos periodos hasta $t = 1$. Similarmente X denota el precio de la acción en $t = 1$ para los caminos en que la opción está "in the money". Los vectores X e Y se presentan en la siguiente tabla.

Regresión en $t = 1$

Caminos	Y	X
1	0.00×0.94176	1.09
2	-	-
3	-	-
4	0.13×0.94176	0.93
5	-	-
6	0.33×0.94176	0.76
7	0.26×0.94176	0.92
8	0.00×0.94176	0.88

La función de la esperanza condicional de Y dado X en $t = 1$, la estimamos, otra vez, haciendo la regresión de Y en una constante, X y X^2. Lo que nos arroja la siguiente función, $\mathbb{E}(Y|X) \simeq 2.038 - 3.335X + 1.356X^2$. Sustituyendo los valores de X en esta estimación obtenemos la estimación de la esperanza condicional de los flujos de fondos futuros descontados condicional a cada X. A continuación se presentan los valores estimados de continuar y los valores de ejercer inmediatamente. Comparando ambos valores, vemos que es óptimo el ejercicio de la opción en $t = 1$ para los caminos cuatro, seis, siete y ocho.

Decisión óptima de continuar en $t = 1$

Caminos	Ejercer	Continuar
1	0.01	0.0139
2	-	-
3	-	-
4	0.17	0.1092
5	-	-
6	0.34	0.2866
7	0.18	0.1175
8	0.22	0.1533

Habiendo identificado la estrategia de ejecución en los tiempos 1, 2 y 3, la regla de tiempos de parada puede ser ahora representada por la siguiente matriz, donde los unos denotan cuando la opción debe se ejercida.

Tiempos de parada

Caminos	$t = 1$	$t = 2$	$t = 3$
1	0	0	0
2	0	0	0
3	0	0	1
4	1	0	0
5	0	0	0
6	1	0	0
7	1	0	0
8	1	0	0

A partir de la especificación de los tiempos de parada es simple determinar los flujos de fondos realizados al seguir dicha estrategia. Solo se debe ejercer la opción en la fecha de ejercicio en la que aparece un uno en la tabla anterior. Así obtenemos la siguiente tabla de flujos de fondos.

Flujos de fondos

Caminos	$t = 1$	$t = 2$	$t = 3$
1	0.00	0.00	0.00
2	0.00	0.00	0.00
3	0.00	0.00	0.07
4	0.17	0.00	0.00
5	0.00	0.00	0.00
6	0.34	0.00	0.00
7	0.18	0.00	0.00
8	0.22	0.00	0.00

Lo único que resta es valuar la opción, esto se logra descontando cada flujo de fondos de la opción hasta $t = 0$ y calculando el promedio de todos los caminos. Haciendo esto obtenemos un valor de 0.1144 para el put americano. Aunque muy simple este ejemplo ilustra claramente el mecanismo con que trabaja el algoritmo LSM.

Complementando este ejemplo, se presentan los resultados que se obtienen de utilizar otras bases en el cálculo de la esperanza condicional, como son lo polinomios (ponderados) de Laguerre:

- $L_0(X) = \exp(\frac{-X}{2})$
- $L_1(X) = \exp(\frac{-X}{2})(1 - X)$
- $L_2(X) = \exp(\frac{-X}{2})(1 - 2X + \frac{X^2}{2})$
- $L_n(X) = \exp(\frac{-X}{2})\frac{e^X}{n!}\frac{d^n}{dX^n}(X^n e^{-X})$

o las funciones de Hermite:

- $h_0(X) = \frac{e^{\frac{-X^2}{2}}}{\pi^{\frac{1}{4}}}$
- $h_1(X) = \frac{2Xe^{\frac{-X^2}{2}}}{\pi^{\frac{1}{4}}\sqrt{2}}$
- $h_3(X) = \frac{(4X^2-2)e^{\frac{-X^2}{2}}}{\pi^{\frac{1}{4}}2\sqrt{2}}$
- $h_n(X) = \frac{H_n(X)e^{\frac{-X^2}{2}}}{\pi^{\frac{1}{4}}2^{\frac{n}{2}}(n!)^{\frac{1}{2}}}$

donde $H_n(X)$ se define a partir de las siguientes fórmulas:

$$H_0(X) = 1 \text{ , } H_1(X) = 2X$$

y

$$XH_n(X) = \frac{1}{2}H_{n+1}(X) + nH_{n-1}(X), \ n = 1, 2, ...$$

Entonces, partiendo de la tabla "Regresión en $t = 2$", hacemos la regresión de Y respecto a las primeras tres funciones de Laguerre, de donde obtuvimos $\mathbb{E}(Y|X) \simeq 1.62326L_0(X) + 22.1752L_1(X) + 50.2224L_2(X)$

Como hicimos anteriormente, utilizamos esta función para calcular el valor de continuar, para luego compararlo con el valor de ejercer la opción inmediatamente, así construimos la siguiente tabla:

Decisión óptima de continuar en $t = 2(Laguerre)$

Caminos	Ejercer	Continuar
1	0.02	0.03803
2	-	-
3	0.03	0.04645
4	0.13	0.11486
5	-	-
6	0.33	0.15313
7	0.26	0.15606
8	-	-

Como podemos observar, los valores obtenidos no distan de los calculados con la primera base, y las decisiones de continuar o ejecutar inmediatamente son exactamente las mismas para $t = 2$. Veamos que sucede lo mismo cuando utilizamos las tres primeras funciones de Hermite. La función que resulta es la siguiente, $\mathbb{E}(Y|X) \simeq -3.92074H_0(X) + 4.15231H_1(X) + -2.46758H_2(X)$ de donde se arma la siguiente tabla:

Decisión óptima de continuar en $t = 2(Hermite)$

Caminos	Ejercer	Continuar
1	0.02	0.040141
2	-	-
3	0.03	0.047484
4	0.13	0.111085
5	-	-
6	0.33	0.154512
7	0.26	0.15542
8	-	-

Procediendo análogamente al ejemplo (no tiene mucho sentido mostrarlo aquí), logramos ver que se arriba a resultados igualmente precisos utilizando estas tres diferentes bases y la misma regla de decisión.

De acuerdo al estudio de Longstaff y Schwartz, testeos numéricos extensivos, indican que los resultados del algoritmo LSM son robustos frente a la elección de bases.

7. Haar Wavelets Systems

En esta sección nos prestaremos a introducir una discretización del espacio de probabilidad. Esta discretización define σ-álgebras atómicas, construidas vía variables aleatorias que son representativas del proceso de precios $(S_n)_{0\leq n\leq T}$. La técnica principal es el uso de sistemas H (H-systems) (o sistemas Haar) que son una generalización básica de las *Haar wavelets* en el intervalo $[0,1]$, con la medida de Lebesgue, a un espacio de probabilidad general $(\Omega, \mathcal{F}, \mathbb{P})$. La elección de estas *wavelets* en particular, además de dar una base ortonormal en $L^2(\Omega, \mathcal{F}, \mathbb{P})$, las sumas parciales de las expansiones en esta base forman una martingala.
Fijada una variable aleatoria X, un *numéraire* $\{B_n\}$ (e.g. una cuenta bancaria) y un proceso de precios $\{S_n\}$ donde $(0=)T_0 \leq n \leq T$. La construcción que veremos nos da la siguiente identidad:

$$\mathbb{E}(X|\mathcal{B}_n) = \mathbb{E}(X|u_0, u_1, ..., u_n) = \sum_{k=0}^{n} \langle X, u_k \rangle u_k,\ n \geq 0. \tag{2}$$

Las funciones simples u_n, las funciones *Haar*, son un conjunto ortonormal en $L^2(\Omega, \mathcal{F}, \mathbb{P})$, donde $(\Omega, \mathcal{F})$ es la σ-álgebra generada por el proceso de precios y $\mathbb{P}$ es la medida neutral al riesgo. La σ-álgebra $\mathcal{B}_n$ está generada por $u_0, u_1, ..., u_n$ y contiene $n+1$ átomos, estos átomos forman una discretización espacio-temporal del proceso. El lado izquierdo de la anterior identidad es una martingala la cual, bajo las condiciones apropiadas, converge a X en casi todo punto.
Veamos, ahora, esto más formalmente.

7.1. **Sistemas H.** Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad arbitrario. Notamos con $\|\cdot\|^2 = \langle , \rangle$ el producto interno en $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Definición 7.1. *Un sistema ortonormal de funciones $\{u_k\}k \geq 0$ definido sobre Ω se denomina sistema H si y solo si para cualquier $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$*

$$X_{n+1} \equiv \mathbb{E}(X|u_0, u_1, ..., u_n) = \sum_{k=0}^{n} \langle X, u_k \rangle u_k,\ \textit{para todo } n \geq 0.$$

La siguiente proposición, la cual se prueba en [8 (Richard Gundy (1966))], nos brinda una caracterización alternativa de los sistemas H equivalente a la definición 4.1.

Proposición 7.1. *Un sistema ortonormal $\{u_k\}k \geq 0$ definido sobre Ω se denomina sistema H si y solo si se satisfacen las siguientes tres condiciones,*

1: *Cada u_k asume a lo sumo dos valores distintos de cero con probabilidad positiva.*
2: *La σ-álgebra $\mathcal{B}_n = \sigma(u_0, ..., u_n)$, consiste exactamente de $n+1$ átomos.*
3: *$\mathbb{E}(u_{k+1}|u_0, u_1, ..., u_n) = 0$, para todo $k \geq 0$. Por lo tanto las funciones u_k son diferencias martingalas.*

Corolario 7.1. *Asumamos que* $\{u_k\}_{k\geq 0}$ *es un sistema H. Luego, para cada* $n \geq 0$*,* u_{n+1} *toma dos valores distintos de cero (uno positivo y el otro negativo) únicamente sobre un átomo de* $\mathcal{B}_n$ *(este átomo se convierte en su soporte). Consecuentemente,* $\mathcal{B}_{n+1}$ *consiste de n átomos de* $\mathcal{B}_n$ *y dos átomos más obtenidos de partir el restante átomo de* $\mathcal{B}_n$*.*

Definición 7.2. *Decimos que una sucesión de particiones de* Ω*,* $\mathcal{Q} := \{\mathcal{Q}_j\}_{j\geq 0}$*, es una sucesión binaria si* $\mathcal{Q}_0 = \{A_{0,0} = \Omega\}$ *y para* $j \geq 1$*, existe* $A_{k,i} \in \mathcal{Q}_j$ *y* $A_{k+1,2i}, A_{k+1,2i+1} \in \mathcal{Q}_{j+1}$ *tal que* $Q_{j+1} = \{Q_j \setminus \{A_{k,i}\}\} \cup \{A_{k+1,2i}, A_{k+1,2i+1}\}$

El índice j en $A_{j,i}$ se denominará *parámetro de escala* o *nivel*, e indica el número de veces en que fue partido $A_{0,0}$ para obtener $A_{j,i}$.

Definición 7.3. *Decimos que un sistema H* $\{u_k\}_{0\leq k\leq m}$ *es un sistema Haar si* $m = \infty$ *o* $m = 2^J - 1$ *y cada átomo de* $\sigma(u_0, ..., u_{2^j-1})$ *es la unión de dos átomos de* $\sigma(u_0, ..., u_{2^{j+1}-1})$ *para todo j tal que* $j \leq J - 1$

Definición 7.4. *Dado* $A \in \mathcal{B}$*,* $P(A) > 0$*, una función* ψ *se dice una función de Haar (o Haar) sobre A si existen* $A_i \in \mathcal{B}$*,* $A_0 \cap A_1 = \emptyset$*,* $A = A_0 \cup A_1$*,* $\psi = a 1_{A_0} + b 1_{A_1}$ *y*

$$\int_\Omega \psi(\omega) d\mathbb{P}(\omega) = 0, \int_\Omega \psi^2(\omega) d\mathbb{P}(\omega) = 1$$

Introduciremos, ahora, algunas propiedades elementales de los sistemas H y de las particiones.

Definición 7.5. *Sea* $\mathcal{P} := \{\mathcal{P}_j\}_{j\geq 0}$ *una sucesión de particiones de* Ω*, decimos que es una sucesión diádica si*

$$\mathcal{P}_j = \{A_{j,i}\}_{i=0}^{2^j-1}$$

y

$$A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$$

Definición 7.6. *Decimos que una sucesión de particiones de* Ω*,* $\mathcal{R} := \{\mathcal{R}_j\}_{j\geq 0}$ *una sucesión diádica débil si* $\mathcal{R}_0 = \{\Omega\}$ *y para* $j \geq 1$*,* $\mathcal{R}_j$ *satisface*
$A \in \mathcal{R}_j$ *si*

$$\text{existe otro } \acute{A} \in \mathcal{R}_j \text{ tal que } A \cup \acute{A} \in \mathcal{R}_{j-1}$$

o

$$A \in \mathcal{R}_k \text{ , para todo} k \geq j - 1.$$

Para mejorar el uso de las sucesiones diádicas débiles, $\mathcal{R} := \{\mathcal{R}_j\}_{j\geq 0}$, introducimos la siguiente indexación para los elementos de cada $\mathcal{R}_j$. Escribimos $\Omega \equiv A_{0,0}$, lo que significa que $\mathcal{R}_0 = \{A_{0,0}\}$. Si $\mathcal{R}_{j-1}$ ya fue indexado , para indexar los elementos de $\mathcal{R}_j$, tenemos dos casos:

1.Si existe un par $A, \acute{A} \in \mathcal{R}_j$ y $A_{k,i} \in \mathcal{R}_{j-1}$ tal que

$$A_{k,i} = A \cup \acute{A}$$

entonces escribimos $A_{k+1,2i} \equiv A$ y $A_{k+1,2i+1} \equiv \acute{A}$.
2. Si $A \in \mathcal{R}_k$ para todo $k \geq j-1$, preservamos el índice de A en $\mathcal{R}_{j-1}$.
Observemos que $\mathcal{R}_j$ puede tener a lo sumo 2^j elementos, y si $A_{k,i} \in \mathcal{R}_j$ entonces $k \leq j$ y $0 \leq i \leq 2^k - 1$. En el caso 1, se desprende de la definición que $k = j-1$. Insistimos que toda la información sobre la separación de los átomos está almacenada en esta indexación. Los siguientes conjuntos de índices serán útiles más tarde. Consideremos $j \geq 0$ y sean

$$I_j \equiv \{i : A_{j,i} \in \mathcal{R}_j \text{ y } A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}\}$$

y

$$K_j \equiv \{(k,i) : A_{k,i} \in \mathcal{R}_j\}$$

Teorema 7.1. *Todo sistema H induce naturalmente una sucesión diádica débil de particiones y se cumple la reciproca.*

Dem: Sea $\{u_k\}_{k\geq 0}$ un sistema H, y $A_{0,0} = \Omega$. Definimos recursivamente la siguiente sucesión de particiones.

$$\mathcal{R}_0 = \{A_{0,0}\}$$

Una vez que $\mathcal{R}_j$ fue definida, generamos $\mathcal{R}_{j+1}$ de la siguiente manera:

- Para $A_{k,i} \in \mathcal{R}_j$ con $k < j$ agregamos $A_{k,i}$ a $\mathcal{R}_{j+1}$
- Para $A_{j,i} \in \mathcal{R}_j$, existe algún u_k, por el corolario 4.1, tal que $A_{j,i} = suppu_k$ (soporte). Luego agregamos

$$A_{j+1,2i} = u_k^{-1}((-\infty, 0)) \text{ y } A_{j+1,2i+1} = u_k^{-1}((0,\infty))$$

  a $\mathcal{R}_{j+1}$.
- Si $A_{j,i}$ no es el soporte de ningún u_k, agregamos $A_{j,i}$ a $\mathcal{R}_{j+1}$.

Claramente, esta sucesión de particiones es diádica débil. Recíprocamente, sea $\mathcal{R}$ una sucesión diádica débil de particiones. Vamos a definir una familia de funciones de Haar $\{\psi_{j,i}\}$, asociada con $\mathcal{R}$. Para $A_{j,i} \in \mathcal{R}_j$ tal que $A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$. Si $i \in I_j$, definimos $\psi_{j,i}$ sobre Ω por

$$\psi_{j,i}(\omega) = \begin{cases} a_{j,i} & \text{si } \omega \in A_{j+1,2i}, \\ b_{j,i} & \text{si } \omega \in A_{j+1,2i+1} \ , \\ 0 & \text{si } \omega \in \{A_{j,i}\}^c. \end{cases}$$

donde $a_{j,i}$ y $b_{j,i}$ se eligen satisfaciendo

$$\int_\Omega \psi_{j,i}(\omega) d\mathbb{P}(\omega) = 0 \text{ y } \int_\Omega \psi_{j,i}^2(\omega) d\mathbb{P}(\omega) = 1$$

Las ecuaciones arriba introducidas, definen (salvo el signo) $\psi_{j,i}(\omega)$ para todo $\omega \in \Omega$. Tenemos

$$a_{j,i} = \sqrt{\frac{\mathbb{P}(A_{j+1,2i+1})}{\mathbb{P}(A_{j+1,2i})\mathbb{P}(A_{j,i})}}$$

y

$$b_{j,i} = -\sqrt{\frac{\mathbb{P}(A_{j+1,2i})}{\mathbb{P}(A_{j+1,2i+1})\mathbb{P}(A_{j,i})}}$$

Dando el orden natural al conjunto $N = \{2^j + i : j \geq 0, i \in I_j\}$ y siendo π un isomorfismo, que preserva el orden, de N en el intervalo $[1, N]$ o $\mathbb{N}$. Definimos

$$\upsilon_{\pi(2^j+i)} = u_{2^j+1} \equiv \psi_{j,i}$$

y $\upsilon_0 \equiv 1_\Omega(\omega)$, la sucesión resultante $\{\upsilon_l\}_{l\geq 0}$ es ortonormal y trivialmente satisface (1), (2) y (3) de la proposición 6.1, por lo tanto, es un sistema H.

Nota 1: EL teorema anterior aplicado a una sucesión diádica de particiones $\mathcal{P}$ implica que $\mathcal{P}_j = \{A_{j,i}\}_{i=0}^{2^j-1}$ con $I_j = [0, 2^j - 1]$, el sistema de funciones $\{u_k\}_{k\geq 0}$ definido por

$$u_{2^j+i} \equiv \psi_{j,i}$$

es un sistema Haar. Esto se cumple, porque un entero $k \geq 1$ dado, puede ser escrito como $k = 2^{j_k} + i_k$ donde j_k es el máximo entero que satisface $2^{j_k} \leq k$, resultando en consecuencia $i_k \in I_{j_k} = [0, 2^{j_k} - 1]$. Más aún, $\mathcal{P}_j$ es el conjunto de átomos de $\sigma(u_0, ..., u_{2^j-1})$.

Nota 2: Es también claro que si comenzamos con la sucesión de particiones $\mathcal{R}$ inducida por un sistema H $\{u_k\}_{k\geq 0}$, entonces el sistema H construido por el teorema 6.1 es un reordenamiento $\{\upsilon_{\pi k}\}_{k\geq 0}$ de $\{u_k\}_{k\geq 0}$

Aquí presentamos un teorema de convergencia de martingalas relevante para nuestro análisis. Asumamos una sucesión creciente de σ-álgebras (no necesariamente atómicas), $\mathcal{B}_n$, $n = 0, 1, ...$ relativa al espacio de probabilidad $(\Omega, \mathcal{B}, \mathbb{P})$. Notamos $\mathcal{B}_\infty = \sigma(\bigcup_{n\geq 0} \mathcal{B}_n)$.

Teorema 7.2. *Sea $p \in [0, \infty)$ un número real dado. para todo $X \in L^p$, la sucesión $X_n = \mathbb{E}(X|\mathcal{B}_n)$ es una martingala que converge casi seguramente en L^p a $X_\infty = \mathbb{E}(X|\mathcal{B}_\infty)$*

Cuando uno aplica el resultado anterior a una opción X dada, uno necesita argumentar separadamente que $\sigma(X) \subseteq \mathcal{B}_\infty$, para entonces obtener que la igualdad.

Proposición 7.2. *Sea $X \in L^2$ y $X_n = \mathbb{E}(X|\mathcal{B}_n)$, entonces para cualquier $\lambda > 0$, tenemos*

$$\mathbb{P}(|X - X_n| \geq \lambda) \leq \frac{\sum_{k=n+1}^{\infty} |\langle X, u_k\rangle|^2}{\lambda^2}$$

Claramente la desigualdad anterior nos brinda información sobre el error de una aproximación de un sistema H, o sea de $e_n(X) = X - X_n = \Sigma_{k=n+1}^{\infty}$

A continuación se presentan cuatro ejemplos de sistemas H extraídos del artículo [5]. El primero es un ejemplo clásico de la teoría de *wavelets*, los restantes pertenecen a la esfera de las finanzas. Los ejemplos ilustran construcciones básicas de H-systems.

7.1.1. *El sistema Haar en $L^2([0,1])$.* Sea $\psi_H : \mathbb{R} \longrightarrow \mathbb{R}$ la siguiente función de Haar

$$\psi_H(t) = \begin{cases} 1 & \text{si } 0 \leq t \leq \frac{1}{2}, \\ \text{-1} & \text{si } \frac{1}{2} < t \leq 1. \end{cases}$$

Definimos $u_0 \equiv 1$ y $u_{2^j+i}(t) = 2^{\frac{j}{2}}\psi_H(2^j t - i)$ para $j = 0, 1, ...$ e $i = 0, ..., 2^j - 1$. La familia de funciones $\{u_k\}_{k=0}^{\infty}$ es conocido como "el sistema Haar". Obsérvese que $\{u_k\}_{k=0}^{\infty} \subset L^2([0,1])$. Nótese además que, puesto que $0 \leq 2^j t - i \leq 1$ tenemos que $2^{-j} i \leq t \leq 2^{-j}(i+1)$ y en consecuencia el soporte de u_{2^j+1} es $[2^{-j}i, 2^{-j}(i+1)]$.

De esta manera tenemos que $\{u_k\}_{k=0}^{\infty}$ es un sistema H. Satisface las condiciones de la proposición 6.1, pues (1) es trivial, (2) es una consecuencia de los soportes de los u_k y (3) se cumple por la ortonormalidad de $\{u_k\}_{k=0}^{\infty}$.
Más aún, si elegimos $m = \infty$ o $m = 2^J - 1$, para algún $J \geq 1$, tenemos que $\{u_k\}_{k=0}^{m}$ es también un sistema Haar. Para todo j tenemos que los átomos de $\sigma(u_0, ..., u_{2^j-1})$ están dados por

$$\{[2^{-j}i, 2^{-j}(i+1)]\}_{i=0}^{2^j-1}$$

7.1.2. *Un sistema Haar para el modelo binomial.* Sea S el precio de una acción y $t_0, t_1, ..., t_n$ las fechas de ejercicio. El precio varia de acuerdo a la siguiente regla

$$S_{t_{i+1}} = S_{t_i} H_{i+1},\ i = 0, 1, ..., n-1$$

donde H_{i+1} es una variable aleatoria tal que

$$H_{i+1} = \begin{cases} U & \text{con probabilidad p}, \\ D & \text{con probabilidad q}. \end{cases}$$

donde $0 < D < 1 < U$ y $p + q = 1$. La situación puede ser descrita en términos del espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, donde $\Omega := \{\omega : \{t_0, t_1, ..., t_n\} \longrightarrow \{U, D\}\}$.
$\mathcal{F} := \mathcal{P}(\Omega)$ y $\mathbb{P}$ la correspondiente medida producto de probabilidad. Luego $S : \Omega \times \{t_0, t_1, ..., t_n\} \longrightarrow \mathbb{R}$ y $S_t(\omega) := S(\omega, t) = S_0 \prod_{t_i \leq t} \omega(t_i)$.
Consideremos los conjuntos $A_{j,i}$, $0 \leq j \leq n-1$ y $0 \leq i \leq 2^j - 1$ definido por $A_{0,0} = \Omega$ y

$$A_{j+1,2i+1} = A_{j,i} \cap \{\omega(t_{j+1}) = U\}$$

$$A_{j+1,2i} = A_{j,i} \cap \{\omega(t_{j+1}) = D\}$$

Es claro que $\mathbb{P}(A_{j+1,2i}) = q\mathbb{P}(A_{j,i})$ y $\mathbb{P}(A_{j+1,2i+1}) = p\mathbb{P}(A_{j,i})$, en consecuencia $P(A_{j,i}) = p^{i_0}...p^{i_j} q^{1-i_0}...q^{1-i_j}$ donde $i = \sum_k = 0^j i_k 2^k$ es el desarrollo 2-ádico de i.
Definimos ahora, para $j = 0, ..., n-1$; $i = 0, ..., 2^j - 1$ las funciones normalizadas

$$u_0 \equiv 1$$
$$u_{2^j+i} = \frac{1}{\sqrt{\mathbb{P}(A_{j,i})}}\left(\sqrt{\frac{p}{q}} 1_{A_{j+1,2i}} - \sqrt{\frac{q}{p}} 1_{A_{j+1,2i+1}}\right)$$

Es claro, por la definición de las funciones u_k, que el sistema $\{u_k\}_{0 \leq k \leq 2^n - 1}$ verifica las condiciones de la proposición 6.1 y forma entonces un sistema H para $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Ahora, observando que para cada $j \geq 0$ los átomos de $\sigma(u_0, ..., u_{2^j-1})$

son $\{A_{j+1,i} : i = 0, ..., 2^{j+1} - 1\}$, tenemos que $\{u_k\}_{0 \leq k \leq 2^n - 1}$ es un sistema Haar. En particular tenemos que el subsistema $\{u_0, ..., u_{2^j - 1}\}$ es una base ortonormal de $L^2(\Omega, \mathcal{F}_{t_j}, \mathbb{P})$, donde $F_{t_j} = \sigma(S_{t_0}, ..., S_{t_j})$

7.1.3. *Un sistema que no es **Haar** para el modelo binomial.* A continuación construiremos otro sistema H para el modelo binomial. Esta vez estará asociado a una partición particular de la σ-álgebra final($\sigma(S_{t_n})$). Sea J el menor entero tal que satisfaga $n + 1 \leq 2^J$, luego para $0 \leq j \leq J$ y $0 \leq i \leq 2^j - 1$ definimos los conjuntos $A_{j,i}$ de la siguiente manera. Para $i \neq 0$,

$$A_{j,i} = \{\omega \in \Omega : \frac{i}{2^j} < \frac{1}{n}|\omega|_U \leq \frac{i+1}{2^j}\}$$

siempre y cuando este conjunto sea no vacío, y para $i = 0$

$$A_{j,0} = \{\omega \in \Omega : 0 < \frac{1}{n}|\omega|_U \leq \frac{1}{2^j}\}$$

donde $|\omega|_U$ es el número de t_i tal que $\omega(t_i) = U$. Las probabilidades de los átomos están dadas por

$$\mathbb{P}(A_{j,i}) = \sum_{\frac{i}{2^j} < \frac{s}{n} \frac{i+1}{2^j}} \binom{n}{s} p^s q^{n-s}$$

para $i \neq 0$ y

$$\mathbb{P}(A_{j,0}) = \sum_{0 < \frac{s}{n} \frac{1}{2^j}} \binom{n}{s} p^s q^{n-s}$$

Es importante observar que $A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$ o $A_{j,i} = A_{j+1,i}$. El sistema H correspondiente esta dado por

$$u_0 \equiv 1$$

$$u_{2^j+i} = \frac{1}{\sqrt{\mathbb{P}(A_{j,i})}} \left(\sqrt{\frac{\mathbb{P}(A_{j+1,2i+1})}{\mathbb{P}(A_{j+1,2i})}} 1_{A_{j+1,2i}} - \sqrt{\frac{\mathbb{P}(A_{j+1,2i})}{\mathbb{P}(A_{j+1,2i+1})}} 1_{A_{j+1,2i+1}} \right)$$

si $A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$.

7.1.4. *Un sistema H para el modelo Black-Scholes.* (En el Anexo II se presenta una descripción del mismo) Este ejemplo describe como se construye una clase básica de sistemas Haar asociados al modelo Black-Scholes. Se desprenderá de aquí, que estos sistemas pueden ser usados para aproximar una familia general de opciones europeas y también bermudas. El punto de partida del modelo Black-Scholes es el *movimiento Browniano* B_t, definido sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ provisto de una filtración $(\mathcal{F}_t)_{0 \leq t \leq T}$. La partición de átomos se realizará utilizando los incrementos de dicho *movimiento Browniano.* El proceso de precios bajo la medida neutral al riesgo $\mathbb{P}$ está dada por $S_t : \Omega \longrightarrow \mathbb{R}$, $0 \leq t \leq T$,

$$S_t(\omega) = S_0 \exp(\upsilon t + \sigma\sqrt{t} W_t(\omega))$$

donde $\upsilon = (r - \frac{\sigma^2}{2})$, y $W_t \sim N(0,1)$ definidas sobre $(\Omega, \mathcal{F}_t, \mathbb{P})$.(claramente r, la tasa de interés libre de riesgo, aproxima el retorno esperado de la acción, y σ representa la volatilidad del precio de la misma).
La construcción se basa en dos parámetros, el primero, denotado por n_T, representa el número de fechas durante el periodo $[0,T]$, en el que está permitido el ejercicio de la opción, y el segundo con conjunto de parámetros, $j_1, ..., j_{n_T}$ será la escala o nivel de discretización asociado a cada fecha de ejercicio. Para mayor simplicidad, la partición de los átomos seŕa realizada en pedazos de igual probabilidad, esta restricción puede ser fácilmente removida (de hecho lo haremos más adelante). Primero veremos el caso $n_T = 1$ y luego procederemos a generalizarlo para $n_T \geq 1$. Por lo tanto, primero nos concentramos sobre la σ-álgebra $\sigma(S_T) = S_T^{-1}(\mathcal{B}(0,\infty))$, donde $\mathcal{B}(0,\infty)$ son los borelianos en $(0,\infty)$. la siguiente ecuación especifica $\mathbb{P}$ sobre $\sigma(S_T)$, sea $B = S_T^{-1}((a_1,a_2))$ $(0 < a_1 < a_2 < \infty)$

$$\mathbb{P}(B) = \frac{1}{\sigma\sqrt{2\pi T}} \int_{\ln(a_1)}^{\ln(a_2)} \exp[\frac{-(\ln(\frac{s}{S_0}) - \upsilon T)^2}{2\sigma^2 T}] \frac{ds}{s}$$

Nota: Ver Anexo II
De nuestra notación anterior tenemos que, $W_T : \Omega \longrightarrow \mathbb{R}$

$$\mathbb{P}(W_T^{-1}(A)) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{y^2}{2}} dy$$

para cualquier boreliano $A \subset \mathbb{R}$. Esta ecuación da $\mathbb{P}$ sobre $\sigma(W_T) = W_T^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{F}_T$, claramente, $\sigma(S_T) = \sigma(W_T)$. Denotemos la distribución normal acumulativa estandarizada con

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{y^2}{2}} dy$$

Dado un entero positivo j, definimos los números $-\infty = c_0^j < c_1^j < ... < c_{2^j}^j = \infty$ tal que satisfagan

$$\Phi(c_{i+1}^j) - \Phi(c_i^j) = \frac{1}{2^j}$$

para todo $i = 0, 1, ..., 2^j - 1$
Si en algún momento se encuentra la desigualdad $\leq \infty$, debe ser interpretada por $< \infty$. Definimos la separación binaria de los átomos inductivamente, fijando $A_{0,0} = \Omega$ y para un j dado consideramos $0 \leq i \leq 2^j - 1$

$$A_{j+1,2i} = \{\omega \in A_{j,i} | c_{2i}^{j+1} < W_T(\omega) \leq c_{2i+1}^{j+1}\}$$

,

$$A_{j+1,2i+1} = \{\omega \in A_{j,i} | c_{2i+1}^{j+1} < W_T(\omega) \leq c_{2i+2}^{j+1}\}$$

Nótese que $A_{j,i} = A_{j+1,2i} \cup A_{j+1,2i+1}$, por lo tanto hemos definido una sucesión diádica de particiones $\mathcal{P} = \{\mathcal{P}_j\}_{j\geq 0}$ con $\mathcal{P}_j = \{A_{j,i}\}$, $i = 0, ..., 2^j - 1$, donde los átomos satisfacen

$$\mathbb{P}(A_{j,i}) = \frac{1}{2j}$$

Para tener otra perspectiva de los átomos en cuestión, notemos que se corresponden a la siguiente partición del rango de S_T,

$$A_{j,i} = \{\omega | a_i^j < S_T(\omega) \leq a_{i+1}^j\}$$

donde $i = 0, ..., 2^j - 1$ y los números reales $0 = a_0^j < a_1^j < ... < a_{2^j}^j = \infty$ satisfacen,

$$a_i^j = S_0 \exp(c_i^j \sigma \sqrt{T} + \upsilon T)$$

Escribiendo $m = 2^j$ y $\mathcal{B}_m = \sigma(\{A_{j,i} : i = 0, ..., m - 1\})$ tenemos que $\mathcal{B}_\infty = \sigma(\bigcup_{m \geq 0} \mathcal{B}_m) = \sigma(S_T)$.
Asociada con esta sucesión de particiones $\mathcal{P}$, tenemos las funciones

$$u_0 \equiv 1$$
$$u_{2^j+i} = a_{j,i} 1_{A_{j+1,2i}} + b_{j,i} 1_{A_{j+1,2i+1}}$$

definidas para $0 \leq j < J$, e $i = 0, ..., 2^j - 1$, donde $a_{j,i} = 2^{\frac{j}{2}}$ y $b_{j,i} = -2^{\frac{j}{2}}$.

Sigue del teorema [6.1] que $\{u_j\}_{j=0}^{2^J-1}$ es un sistema Haar capaz de aproximar cualquier variable aleatoria en $L^2(\Omega, \sigma(S_T), \mathbb{P})$, eligiendo un J suficientemente grande.
Estamos ahora en condiciones de describir la construcción de un sistema Haar finito para un $n_T \geq 1$ arbitrario. La idea es simplemente construir un sistema Haar diádico por medio de una concatenación de varios sistemas Haar, cada uno análogo al del caso $n_T = 1$, pero este último restringido a intervalos de tiempo cada vez más pequeños. Dada una sucesión arbitraria de intervalos de tiempo $0 = t_0 < t_1 < ... < t_{n_T-1} < t_{n_T} = T$, consideramos los incrementos del *movimiento Browniano* $\sqrt{t_{i+1} - t_i} W_{t_i,t_{i+1}}$, donde las variables aleatorias $W_{t_i,t_{i+1}} \sim N(0,1)$ son independientes. Fijamos una sucesión de escalares $\{j_i = j_{t_i}\}_{i=1}^{n_T}$, vamos a definir la partición de los átomos por etapas, de acuerdo a los intervalos de tiempo $\{t_i, t_{i+1}\}$.
Para la primera etapa, $\{t_0, t_1\}$, definimos la separación binaria de los átomos inductivamente a partir de establecer, $A_{0,0} = \Omega$ y para $0 \leq j < j_1$, $i = 0, ..., 2^j - 1$, $A_{j+1,2i}$ y $A_{j+1,2i+1}$ como en el caso $n_T = 1$, reemplazando W_T con W_{t_0,t_1}.
Para la segunda etapa $\{t_1, t_2\}$, y como modelo para las subsecuentes, consideremos $0 \leq j < j_2$ y $i = 0, ..., 2^j - 1$, sea p y $0 \leq q < 2^j$ los respectivos cocientes y residuo en la división entera de i por 2^j, luego definimos inductivamente los conjuntos

$$A_{j_1+j+1,2i} = \{\omega \in A_{j_1+j,i} | c_{2q}^{j+1} < W_{t_1,t_2}(\omega) \leq c_{2q+1}^{j+1}\}$$

,

$$A_{j_1+j+1,2i+1} = A_{j_1+j,i} \cap \{\omega | c_{2q+1}^{j+1} < W_{t_1,t_2}(\omega) \leq c_{2q+2}^{j+1}\}$$

Notemos que $P(A_{j_1+1,i}) = \frac{1}{2^{j_1+1}}$ por la independencia de W_{t_0,t_1} y W_{t_1,t_2}. El proceso se continua, para una etapa genérica $\{t_k, t_{k+1}\}$, como en la etapa previa, reemplazando j_1 con $J_k = j_1 + ... + j_k$ y W_{t_1,t_2} con $W_{t_k,t_{k+1}}$. Esta inducción se debe continuar hasta $k + 1 = n_T$.
Hemos definido, una vez más, una sucesión diádica de particiones $\{\mathcal{P}_j\}_{j \geq 0}$ con $\mathcal{P}_j = \{A_{j,i}\}$, $i = 0, ..., 2^j - 1$ y en consecuencia, siguiendo el teorema [6.1], existe un sistema Haar $\{u_j\}_{j=0}^{2^J-1}$ asociado.

8. Variación del LSM mediante la utilización de sistemas H

Es momento de estudiar el algoritmo en el contexto de los sistemas H, para las discretizaciones finitas ya introducidas. Veremos que esto permitirá calcular bien las esperanzas condicionales y que permitirá utilizar el algoritmo más allá del contexto markoviano.
Partimos de un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, con una filtración $\{\mathcal{F}_t\}_{t=0}^T$. Sea así mismo $\{B_t\}_{t=0}^T$ un numéraire y $\{S_t\}_{t=0}^T$ un proceso de precios y X una variable aleatoria (e.g. una opción bermuda). Luego podemos construir una discretización de Ω con una sucesión diádica débil de particiones $\{R_j\}_{j\geq 0}$ generada a partir del proceso de precios $\{S_t\}_{t=0}^T$. Por lo tanto y por el teorema 6.1 tenemos un sistema H, $\{u_k\}_{k\geq 0}$, inducido naturalmente por dicha sucesión de particiones. Tenemos, además, que $X_n = \mathbb{E}(X|\mathcal{B}_n)$ converge casi seguramente a $X_\infty = \mathbb{E}(X|\mathcal{B}_\infty)$ en L^2 y se debe probar en cada caso en particular que $\sigma(S_T) = \sigma(X) \subseteq \mathcal{B}_\infty$ (los cual se satisface en los ejemplos citados).

Nota: Como vimos en la construcción de los sistemas Haar [6.1.4] debemos construir un $\mathcal{B}_\infty$ para todos los $\mathcal{F}_n$ $n = 0, 1, ..., T$, es decir para todos las fechas de ejercicio, mediante las concatenaciones propuestas en dicho apartado, sin embargo en los ejemplos que veremos más adelante, podemos entender que se construye un $\mathcal{B}_\infty$ para cada uno de los $\mathcal{F}_n$ $n = 0, 1, ..., T$ (i.e. para cada fecha de ejercicio), dado que suponemos que el proceso es markoviano.

Luego podemos aproximar $U_n = \mathbb{E}(Z_{\nu_{n+1}}|\mathcal{F}_n)$ $(= \mathbb{E}(Z_{\nu_{n+1}}|\mathcal{B}_\infty) = \sum_{k=0}^{\infty} \langle Z_{\nu_{n+1}}, u_k\rangle u_k)$ por los primeros $2^J - 1 = m$ términos, $\mathbb{E}(Z_{\nu_{n+1}^M}|\mathcal{B}_M) = \sum_{k=0}^{M} \langle Z_{\nu_{n+1}^M}, u_k\rangle u_k = P_{n+1}^M(Z_{\nu_{n+1}^M})$, si $\sigma(S_n) = \mathcal{F}_n \subseteq \mathcal{B}_\infty$. Así obtendríamos,

- $\nu_T^M = T$
- $\nu_k^M = k 1_{\{Z_k \geq \mathbb{E}(Z_{\nu_{n+1}^M}|\mathcal{B}_M)\}} + \nu_{k+1}^M 1_{\{Z_k < \mathbb{E}(Z_{\nu_{n+1}^M}|\mathcal{B}_M)\}}$

De este modo nos deshacemos de la condición markoviana del proceso, extendiendo su campo de acción a casos más generales. Solo resta estimar los coeficientes $\langle Z_{\nu_{n+1}}, u_k\rangle := a_k$ por algún método de simulación. Para ello, obviamente, podemos valernos del método de mínimos cuadrados a partir de la simulación de caminos, procediendo análogamete al algoritmo LSM. La convergencia del método se prueba como un caso particular de la convergencia del método LSM.
En este caso debemos ver que,

Teorema 8.1. *Asumamos que se satisface A_1 de la sección 1, luego, para $k = 1$ hasta T, tenemos que*

$$\lim_{M \longrightarrow +\infty} \mathbb{E}(Z_{\nu_k^M}|\mathcal{B}_\infty) = \mathbb{E}(Z_{\nu_k}|\mathcal{B}_\infty)$$

en L^2.

Dem: Claramente nuestro sistema H satisface la condición A_1, luego tomando, como ya se dijo, $\sum_{k=0}^{M} \langle Z_{\nu_{n+1}^M}, u_k\rangle u_k = P_{n+1}^M(Z_{\nu_{n+1}^M})$, la demostración sigue análogamente al caso visto en la sección 4, siendo $\alpha_k^M = E(Z_{\nu_{k+1}^M} u^M)$ $(u^M = (u_0, ..., u_M))$ en la notación propuesta en 4, dado que la matriz $(A_k^M)_{1\leq j,l\leq M}$ es la identidad.

Para el caso del teorema 4.2, debemos asumir que $\mathbb{P}(\alpha_k^M \cdot u^M = Z_k) = 0$. Luego su prueba, mediante los lemas, se satisface claramente, solo se utiliza la base $\{u_k\}_k$ en lugar de la general.
A continuación y para finalizar presentaremos algunos ejemplos del algoritmo con las variaciónes propuestas.
Consideraremos un put americano, sobre una acción que no paga dividendos, con fechas de ejercicio son $n = 1, 2, 3 = T$ y con precio de ejercicio $K = 1.1$. Consideraremos además que el valor inicial $S_0(\omega_i) = 1$ para todo i, la tasa de interés es $r = 0.06$ (esto es un factor de descuento de 0.94176), y una volatilidad $\sigma = 0.2$. Simularemos para el ejemplo 32 caminos a partir del modelo $S_t(\omega_i) = \exp(\upsilon(t) + \sigma\sqrt{t}W_t(\omega_i))$, $i = 0, ..., 31$ donde, como ya vimos $\upsilon = (r - \frac{\sigma^2}{2}) = 0.04$ y $W_t \sim N(0, 1)$.

Nota: Un problema que se sucita en la discretización de del espacio de probabilidad es que si discretizamos hasta el nivel j, en por ejemplo $n = T - 1$, entonces para todos los $\mathcal{A}_i = \{\omega \in A_{j,i}\}$ con $i = 0, ..., 2^j - 1$ fijo, tenemos que sus imagenes $u_k(\mathcal{A}_i) = cte$ $k = 0, ..., 2^j - 1$. Luego aquellos $\omega \in \mathcal{A}_i$ tales que $Z_T(\omega) = 0$ generan iguales pares de datos $(Z_T(\omega), u_0(\omega), ..., u_{2^j-1}(\omega))$, lo que claramente reduce la cantidad con la que contamos para hacer la regresión (obviamente el número de pares de datos distintos debe ser mayor a 2^j). Esto puede llevar a imposibilitar la regresión. Por lo general una solución a este problema, es conducir la dsicretización de manera tal de separar los ω tales que $Z_T(\omega) = 0$ en diferentes átomos.

La matriz de los precios de la acción se presentan en la siguiente matriz,

Camino	$t=0$	$t=1$	$t=2$	$t=3$
1	1	1.239431735	1.151117681	1.107954857
2	1	0.963765885	1.546608101	1.554043126
3	1	1.188591174	0.830836669	1.636554879
4	1	1.162961975	1.093734264	0.9439432
5	1	0.674287849	0.743197891	0.999939983
6	1	0.843997466	0.823247342	2.478933061
7	1	1.012770549	0.775973897	0.458194479
8	1	0.648006847	0.99205271	1.15895959
9	1	1.078831166	1.165266202	1.238873382
10	1	0.857417254	1.220202866	1.080214854
11	1	0.964718277	0.692860477	1.239453139
12	1	1.217366624	1.378077838	1.147139996
13	1	0.938230213	1.50186576	1.389317616
14	1	1.46374033	1.630034439	1.166755869
15	1	0.840318288	1.054509323	0.878179583
16	1	1.294451793	0.862015398	1.472694543
17	1	1.273689527	2.342111145	3.695959296
18	1	1.101470198	1.002795989	0.908222222
19	1	0.935712051	1.435399039	0.567472924
20	1	0.976997415	0.744753837	1.322222335
21	1	0.757696198	0.553456931	1.829289847
22	1	0.924020918	0.913146879	1.217697318
23	1	1.150091348	1.431197622	1.389052823
24	1	1.032979953	0.858408469	1.414627613
25	1	1.024465644	1.051418038	1.987701107
26	1	1.186986407	0.763748026	1.22987041
27	1	0.98181533	1.293864372	1.240886613
28	1	1.07950251	0.979830556	0.789205455
29	1	0.894926748	1.026977427	1.514561638
30	1	1.049777384	2.09150931	1.426667623
31	1	0.961302018	0.826116062	1.297371016
32	1	1.036591308	0.703603997	1.35177373

Tabla 1

A continuación construimos una matriz que contenga los límites de los átomos (análogo al ejemplo de un sistema H para el modelo Black-Scholes), donde los a_i^j fueron elegidos deliberadamente de manera de evitar el problema mencionado en la nota anterior.

a_i^j	c_i^j	j	i
0	$-\infty$	0	0
∞	∞	0	1
0	$-\infty$	1	0
0.84	-0.849725036524578	1	1
∞	∞	1	2
0	$-\infty$	2	0
0.75	-1.1768767713691	2	1
0.84	-0.849725036524578	2	2
0.985	-0.390039475804105	2	3
∞	∞	2	4
0	$-\infty$	3	0
0.7	-1.37604203599488	3	1
0.75	-1.1768767713691	3	2
0.825	-0.901739981536408	3	3
0.84	-0.849725036524578	3	4
0.885	-0.699077743320117	3	5
0.985	-0.390039475804105	3	6
1.035	-0.24710179663498	3	7
∞	∞	3	8
0	$-\infty$	4	0
0.62	-1.72638065329737	4	1
0.7	-1.37604203599488	4	2
0.72	-1.29471971899473	4	3
0.75	-1.1768767713691	4	4
0.8	-0.990570108609716	4	5
0.825	-0.901739981536408	4	6
0.83	-0.884297322148016	4	7
0.84	-0.849725036524578	4	8
0.86	-0.781798341454871	4	9
0.885	-0.699077743320117	4	10
0.94	-0.525029066478048	4	11
0.985	-0.390039475804105	4	12
1.015	-0.303430439358141	4	13
1.035	-0.24710179663498	4	14
1.07	-0.151096466966959	4	15
∞	∞	4	16

Tabla 2

Podemos, ahora, calcular $\mathbb{P}(A_{j,i})$ y luego los coeficientes $a_{j,i}$ y $b_{j,i}$. Una vez que poseemos dichos coeficientes, poseemos también las funciones $\{u_k\}_{k=0}^{15}$ y estamos en condiciones de realizar la regresión de Y (el vector de las ganancias de la opción en $T = 3$, descontados a $t = 2$) con las primeras 16 funciones del sistema Haar construido. Así construimos la matriz de decisión siguiente (los caminos considerados son aquellos para los cuales la opción se encuentra "in the money" en $t = 2$):

Camino	Y	X	$\hat{Y}$
3	0	0.269163331	$-3.15151e^{-17}$
4	0.146968052	0.006265736	0.146968052
5	0.094232521	0.356802109	0.047116261
6	0	0.276752658	$5.55112e^{-17}$
7	0.604426768	0.324026103	0.302213384
8	0	0.10794729	0.09030432
11	0	0.407139523	$-4.17048e^{-17}$
15	0.208901596	0.045490677	0.104450798
16	0	0.237984602	$-1.19777e^{-17}$
18	0.18060864	0.097204011	0.09030432
20	0	0.3552461636	0.047116261
21	0	0.546543069	$-3.80871e^{-17}$
22	0	0.186853121	$-2.77556e^{-17}$
24	0	0.241591531	$-1.63312e^{-17}$
25	0	0.048581962	0.104450798
26	0	0.336251974	0.302213384
28	0.292693871	0.120169444	0.292693871
29	0	0.073022573	$-8.32667e^{-17}$
31	0	0.273883938	$-2.01316e^{-17}$
32	0	0.396396003	0

Tabla 3

Tenemos así la estrategia de tiempos de parada para $t = 2$. Realizando el análisis, para $t = 1$, obtenemos la siguiente matriz de flujos de fondos futuros de la opción de acuerdo a nuestra estrategia de tiempos de parada,

Camino	$t = 1$	$t = 2$	$t = 3$
1	0	0	0
2	0.136234115	0	0
3	0	0.269163331	0
4	0	0	0.1560568
5	0.425712151	0	0
6	0.256002534	0	0
7	0.087229451	0	0
8	0.451993153	0	0
9	0	0	0
10	0.242582746	0	0
11	0.135281723	0	0
12	0	0	0
13	0.161769787	0	0
14	0	0	0
15	0.259681712	0	0
16	0	0.237984602	0
17	0	0	0
18	0	0.097204011	0
19	0	0	0.532527076
20	0.123002585	0	0
21	0.342303802	0	0
22	0	0.186853121	0
23	0	0	0
24	0	0.241591531	0
25	0	0	0
26	0	0.3362519746	0
27	0.11818467	0	0
28	0	0	0.310794545
29	0.205073252	0	0
30	0.050222616	0	0
31	0.138697982	0	0
32	0	0.396396003	0

Tabla 4

Descontando estos flujos a $t = 0$ y promediándolos obtenemos el valor de la opción. En nuestro caso tenemos que $\hat{U}^{16}_{32,0} = 0.167249453$. Calculando el valor de la opción con el algoritmo LSM, utilizando, en la regresión de las esperanzas condicionales, una constante y las primeras 15 potencias de X, logramos $\hat{U}^{16}_{32,0} = 0.168469223$.

Ahora y para finalizar, veremos el ejemplo LSM introducido en la sección 4, pero resuelto mediante los sistemas Haar propuestos. Para eso partimos nuevamente de la matriz de caminos,

Caminos	$t = 0$	$t = 1$	$t = 2$	$t = 3$
1	1.00	1.09	1.08	1.34
2	1.00	1.16	1.26	1.54
3	1.00	1.22	1.07	1.03
4	1.00	0.93	0.97	0.92
5	1.00	1.11	1.56	1.52
6	1.00	0.76	0.77	0.90
7	1.00	0.92	0.84	1.01
8	1.00	0.88	1.22	1.34

El primer paso, claramente es ver los flujos de fondos en $t = 3$,

Flujos de fondos en $t = 3$

Caminos	$t = 1$	$t = 2$	$t = 3$
1	-	-	0.00
2	-	-	0.00
3	-	-	0.07
4	-	-	0.18
5	-	-	0.00
6	-	-	0.20
7	-	-	0.09
8	-	-	0.00

A partir de estos y considerando los caminos en $t = 2$ que se encuentran "*in the money*" generamos nuestros $A_{j,i}$

a_i^j	a_i^j	j	i
0	$-\infty$	0	0
∞	∞	0	1
0	$-\infty$	1	0
1	-0.346410162	1	1
∞	∞	1	2
0	$-\infty$	2	0
0.8	-0.990570109	2	1
1	-0.346410162	2	2
1.075	-0.137638394	2	3
∞	∞	2	4

Luego debemos realizar la regresión de Y contra las primera 4 funciones Haar y obtenemos la siguiente tabla de decisión,

Decisión óptima de continuar en $t = 2$

Caminos	Ejercer	Continuar
1	0.02	7.28583E-17
2	-	-
3	0.03	0.0659232
4	0.13	0.1271376
5	-	-
6	0.33	0.188352
7	0.26	0.1271376
8	-	-

La comparación de ambos valores implica que es óptimo ejercer la opción en $t = 2$ para los caminos uno, cuatro, seis y siete.
Procediendo análogamente para $t = 1$ obtenemos las siguientes tablas,

Flujos de fondos en $t = 2$

Caminos	$t = 1$	$t = 2$	$t = 3$
1	-	0.02	0.00
2	-	0.00	0.00
3	-	0.00	0.07
4	-	0.13	0.00
5	-	0.00	0.00
6	-	0.33	0.00
7	-	0.26	0.00
8	-	0.00	0.00

A partir de la cual, nuevamente, generamos las $A_{j,i}$ convenientes para nuestros caminos "*in the money*",

a_i^j	a_i^j	j	i
0	$-\infty$	0	0
∞	∞	0	1
0	$-\infty$	1	0
0.9	-0.650559771898794	1	1
∞	∞	1	2
0	$-\infty$	2	0
0.8	-0.990570108609716	2	1
0.9	-0.650559771898794	2	2
1	-0.346410161513775	2	3
∞	∞	2	4

De este modo obtenemos la siguiente matriz de tiempos de parada,

Flujos de fondos en $t = 1$

Caminos	$t = 1$	$t = 2$	$t = 3$
1	0.01	0.00	0.00
2	0.00	0.00	0.00
3	0.00	0.00	0.07
4	0.17	0.00	0.00
5	0.00	0.00	0.00
6	0.34	0.00	0.00
7	0.00	0.26	0.00
8	0.22	0.00	0.00

Lo único que resta es valuar la opción, esto se logra, como ya vimos, descontando cada flujo de fondos de la opción hasta $t = 0$ y calculando el promedio de todos los caminos. Haciendo esto obtenemos un valor de 0.1232459 para el put americano. Recordemos que el valor que nos arroja el LSM para este ejemplo es 0.1144. Cabe aclarar, que dadas las características de este ejemplo (solo cinco caminos en cada tiempo con los cuales hacer la regresión), solo nos permite utilizar cuatro funciones Haar, lo que limíta la precisión del método.

9. Conclusiones

En este trabajo hemos introducido una variación sobre el algoritmo LSM, propuesto por Longstaff y Schwartz, la cual nos permite la utilizacíon del mismo en contextos más generales, tales como procesos no markovianos, lo que reviste, sin lugar a dudas, una cualidad deseable del algoritmo. La clave en esta variación es la discretización del espacio de probabilidad subyacente asociada al proceso de precios relevante, que a su vez nos permite construir un sistema H (o Haar), donde expandir las esperanzas condicionales. Este método nos brinda un resultado eficiente (en cuanto está probada su convergencia), aunque no presenta una simplificación de la matemática, ni tampoco modifica necesariamente la velocidad computacional. De este modo hemos alcanzado los objetivos que nos hemos propuesto al encarar este trabajo.

10. Anexo I

Comenzaremos con la definición de algunos conceptos fundamentales para desarrollar la teoría de probailidad.

Definición 10.1. *Una medida* $\mathbb{P}$ *en un espacio medible* $(\Omega, \mathcal{F})$*, se llama medida de probabilidad si*

$$\mathbb{P}(\Omega) = 1$$

$(\Omega, \mathcal{F}, \mathbb{P})$ es llamado *espacio de probabilidad.* Los conjuntos A $\in \mathcal{F}$ son llamados *eventos.*

Para describir un experimento aleatorio matemáticamente, comenzamos con el espacio muestral Ω, el conjunto de todos los posibles resultados. Cada punto ω, punto muestral de Ω, representa un posible resultado de realizar el experimento aleatorio. Para un conjunto $A \subseteq \Omega$ de puntos ω, queremos conocer la probabilidad $\mathbb{P}(A)$. Naturalmente queremos que,

(1) $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1.$
(2) $\mathbb{P}(A) \geq 0 \ \forall \ A \subset \Omega.$
(3) Si $A_1, A_2, ..., A_n$ son disjuntos dos a dos y $A_i \subset \Omega \ \forall i = 1, 2, ..., n$, $\mathbb{P}(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathbb{P}(A_i)$*
(4) Si $B \subseteq A$ y $\mathbb{P}(A) = 0$, entonces $\mathbb{P}(B) = 0$ (completitud)

* Luego, esta condición, la reemplazaremos por: Si $A_1, A_2, ...$ son disjuntos dos a dos y $A_i \subset \Omega \ \forall i = 1, 2, ...$, $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

Por lo tanto la clase $\mathcal{F}$ de subconjuntos de Ω cuyas probabilidades $\mathbb{P}(A)$ están definidas, debería ser cerrada ante uniones disjuntas numerables, complementos y contener al conjunto vacio $\emptyset$ y a todo el espacio Ω. Por lo tanto $\mathcal{F}$ debería ser una σ-álgebra y $\mathbb{P}$ debería definirse de acuerdo a la definición 10.1.
Suele se útil la cuantificación de los resultados cualitativos de un espacio muestral (e.g. el lanzamiento de una moneda puede ser 'cara' o 'cruz'; un producto puede ser 'defectuoso' o 'no defectuoso', etc.) y, mediante el empleo de medidas numéricas, estudiar luego su comportamiento aleatorio. El concepto de variable aleatoria nos da una herramienta para relacionar cualquier resultado con una medida cuantitativa.

Definición 10.2. *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad. Una variable aleatoria (vector aleatorio),* v.a.*,* X *es una función* $\mathrm{X} : \Omega \to \mathbb{R}(\mathbb{R}^k)$*, tal que* $\mathrm{X}^{-1}(B) = \{\omega \in \Omega : \mathrm{X}(\omega) \in B\} \in \mathcal{F}$ *para todos los Borelianos* $B \in \mathcal{B}(\mathbb{R})(\mathcal{B}(\mathbb{R}^k))$*, o sea, tal que sea medible (Borel).*

En particular, tenemos para una *v.a.* X que $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$ para todo $x \in \mathbb{R}$. Por lo que se define la *función de distribución* F_X de X como

$$F_X(x) := \mathbb{P}(\{\omega : \ X(\omega) \leq x\})$$

La menor σ-álgebra que contiene a todos los conjuntos $\{\omega : \ X(\omega) \leq x\}$ para todo x (equivalentemente $\{\omega : \ X(\omega) < x\}$, $\{\omega : \ X(\omega) \geq x\}$, $\{\omega : \ X(\omega) > x\}$) se llama σ-álgebra *generada* por X, y se escribe $\sigma(X)$, por lo tanto,

$$X \text{ es } \mathcal{F}\text{-medible (i.e. es una } \textit{variable aleatoria}\text{) sii } \sigma(X) \subseteq \mathcal{F}$$

Definición 10.3. *La* esperanza*,* $\mathbb{E}$*, de una* v.a. X *en* $(\Omega, \mathcal{F}, \mathbb{P})$ *se define como,*

$$\mathbb{E}(\mathrm{X}) := \int_{\Omega} \mathrm{X} \ \mathrm{d}\ \mathbb{P} = \int_{\Omega} \mathrm{X}(\omega) \mathrm{d}\ \mathbb{P}(\omega)$$

Definición 10.4. *La* varianza*,* $\mathbb{V}\mathrm{ar}$*, de una* v.a. X *en* $(\Omega, \mathcal{F}, \mathbb{P})$ *se define como,*

$$\mathbb{V}\mathrm{ar}(\mathrm{X}) := \mathbb{E}[(\ \mathrm{X} - \mathbb{E}(\mathrm{X}\ ))^2] = \mathbb{E}(\mathrm{X}\ ^2) - (\mathbb{E}(\mathrm{X}))^2$$

Definición 10.5. *Las* v.a. X_1*,* X_2*, ...,* X_n *son independientes si para cualquier* $A_i \in \mathcal{B}$*,* $i = 1, 2, ..., n$ *tenemos*

$$\mathbb{P}(\bigcap_{i=1}^{n} \{\mathrm{X}_i \in A_i\}) = \prod_{i=1}^{n} \{\mathrm{X}_i \in A_i\}$$

Teorema 10.1. *Si* X_1*,* X_2*, ...,* X_n *son independientes y* $\mathbb{E}(|\mathrm{X}_i|) < \infty$*, i= 1, 2,...,n, entonces tenemos que*

$$\mathbb{E}(\prod_{i=1}^{n} \mathrm{X}_i) = \prod_{i=1}^{n} \mathbb{E}(\mathrm{X}_i)$$

Dadas dos *medidas* $\mathbb{P}$ y $\mathbb{Q}$ definidas sobre una misma σ-álgebra $\mathcal{F}$, decimos que $\mathbb{P}$ es *absolutamente continua* respecto a $\mathbb{Q}$, y notamos $\mathbb{P} << \mathbb{Q}$, si para todo $E \in \mathcal{F}$, con $\mathbb{Q}(E) = 0$ tenemos $\mathbb{P}(E) = 0$

Teorema 10.2. *(*Radon-Nikodým*)*$\mathbb{P} << \mathbb{Q}$ *si y solo si existe una función* f_0*,* $\mathcal{F}$*-medible tal que,*

$$\mathbb{P}(E) = \int_E \mathrm{f}_0 \ \mathrm{d}\ \mathbb{Q}\ ,\ \forall E \in \mathcal{F}$$

Puesto que $\mathbb{P}(E) = \int_E \ d\,\mathbb{P}$, tenemos que $\int_E \ d\,\mathbb{P} = \int_E f_0 \ d\,\mathbb{Q}$, para todo $E \in \mathcal{F}$. Por analogía con la regla de la cadena del cálculo ordinario, escribimos $f_0 \equiv d\,\mathbb{P}/d\,\mathbb{Q}$, por lo que nos queda

$$\int_E \ d\,\mathbb{P} = \int_E \frac{d\,\mathbb{P}}{d\,\mathbb{Q}}\ d\,\mathbb{Q}, \forall E \in \mathcal{F}$$

La función medible (*v.a.*) $(d\,\mathbb{P}/d\,\mathbb{Q})$ se llama *derivada de Radon-Nikodým* de $\mathbb{P}$ respecto de $\mathbb{Q}$.
Si tenemos que $\mathbb{P} << \mathbb{Q}$ y también que $\mathbb{Q} << \mathbb{P}$, decimos que $\mathbb{P}$ y $\mathbb{Q}$ son medidas *equivalentes*, y lo escribimos $\mathbb{P} \sim \mathbb{Q}$. Entonces ambas derivadas $d\,\mathbb{P}/d\,\mathbb{Q}$ y $d\,\mathbb{Q}/d\,\mathbb{P}$, existen y tenemos

$$\frac{d\,\mathbb{P}}{d\,\mathbb{Q}} = 1/\frac{d\,\mathbb{Q}}{d\,\mathbb{P}}$$

Si $\mathbb{P} \sim \mathbb{Q}$, entonces $\mathbb{P}(E) = 0$ sii $\mathbb{Q}(E) = 0$: i.e. $\mathbb{P}$ y $\mathbb{Q}$ tienen los mismos *conjuntos nulos.*

10.1. **Esperanza Condicional.** Supongamos que X es una *v.a.*, cuya esperanza existe (i.e. $\mathbb{E}(|X|) < \infty$). Entonces $\mathbb{E}(X)$, la esperanza de X, es un escalar no aleatorio. El operador $\mathbb{E}$ 'promedia' la aleatoriedad de X y nos da su media (un promedio, ponderado de acuerdo a su probabilidad en el caso discreto, del valor de X).
Suele suceder que dispongamos de infomación parcial acerca de la *v.a.* X. En ese caso, nos puede ser útil 'promediar' la aleatoriedad restante. Esto es una esperanza *condicionada* por nuestra información parcial, y se denomina *esperanza condicional.*

En el caso *discreto*, basandonos en la conocida fórmula

$$\mathbb{P}(A|B) := \mathbb{P}(A \cap B)/\mathbb{P}(B)$$

Si $\mathbb{P}(B) > 0$, tenemos que si X toma los valores x_1, ..., x_m con probabilidades $f_1(x_i) > 0$, Y toma los valores y_1, ..., y_n con probabilidades $f_2(y_j) > 0$ y (X, Y) toma los valores (x_i, y_j) con probabilidades $f(x_i, y_j) > 0$, entonces

(1) $f_1(x_i) = \sum_{j=1}^{n} f(x_i, y_j)$, $f_2(y_j) = \sum_{j=1}^{m} f(x_i, y_j)$

(2) $\mathbb{P}(Y = y_j | X = x_i) = \frac{\mathbb{P}(X=x_i, Y=y_j)}{\mathbb{P}(X=x_i)} = \frac{f(x_i,y_j)}{f_1(x_i)} = \frac{f(x_i,y_j)}{\sum_{j=1}^{n} f(x_i,y_j)}$

Esta es la distribución de Y dado $X = x_i$, y lo notamos de la siguiente manera,

$$f_{Y|X}(y_j|x_i) = \frac{f(x_i, y_j)}{f_1(x_i)} = \frac{f(x_i, y_j)}{\sum_{j=1}^{n} f(x_i, y_j)}$$

Su *esperanza condicional* es

$$\mathbb{E}(Y|X = x_i) = \sum_j y_j f_{Y|X}(y_j|x_i) = \frac{\sum_j y_j f(x_i, y_j)}{\sum_j f(x_i, y_j)}$$

En el caso *denso*, si (X, Y) tiene densidad $f(x, y)$, X tiene densidad $f_1(x) := \int_{-\infty}^{\infty} f(x, y)\,dy$, Y tiene densidad $f_2(y) := \int_{-\infty}^{\infty} f(x, y)\,dx$, definimos la *densidad condicional* de Y dado $X = x$ como,

$$f_{Y|X}(y|x) := \frac{f(x, y)}{f_1(x)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y)\,dy}$$

su esperanza es

$$\mathbb{E}(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,dy = \frac{\int_{-\infty}^{\infty} y f(x, y)\,dy}{\int_{-\infty}^{\infty} f(x, y)\,dy}$$

Recordemos que un *espacio de probabilidad* está dado por $(\Omega, \mathcal{F}, \mathbb{P})$. Supongamos que $\mathcal{G}$ es una sub-σ-álgebra de $\mathcal{F}$, $\mathcal{G} \subset \mathcal{F}$ (la σ-álgebra $\mathcal{F}$ representa '*saber todo*', mientras que $\mathcal{G}$ representa '*saber algo*'). Supongamos que Y es una *v.a.* no negativa con esperanza $\mathbb{E}(Y) < \infty$. Luego la función

$$\mathbb{Q}(B) := \int_B Y d\,\mathbb{P} \ (B \in \mathcal{G})$$

es no negativa (pues Y lo es), σ-aditiva (fácil de ver) y definida sobre la σ-álgebra $\mathcal{G}$, por lo tanto es una *medida* sobre $\mathcal{G}$. Si $\mathbb{P}(B) = 0$, entonces $\mathbb{Q}(B) = 0$, por lo tanto $\mathbb{Q} << \mathbb{P}$. Por el teorema de *Radon-Nikodým*, existe la *derivada de Radon-Nikodým* de $\mathbb{Q}$ respecto de $\mathbb{P}$ en $\mathcal{G}$, que es $\mathcal{G}$-medible. Siguiendo a Kolmogorov, llamamos a esta derivada la esperanza condicional de Y dado (o condicionada por) $\mathcal{G}$, $\mathbb{E}(Y|\mathcal{G})$: que es $\mathcal{G}$-medible, integrable y satsiface

$$\int_B Y d\,\mathbb{P} = \int_B \mathbb{E}(Y|\mathcal{G})\, d\,\mathbb{P}\ \forall B \in \mathcal{G}$$

En el caso general, en el que Y es una *v.a.* cuya esperanza existe ($\mathbb{E}(|Y|) < \infty$), pero que puede tambien tomar valores negativos, descomponemos Y en su parte positiva y negativa,

$$Y = Y^{+} - Y^{-}$$

y definimos $\mathbb{E}(Y|\mathcal{G})$ por linealidad como

$$\mathbb{E}(Y|\mathcal{G}) = \mathbb{E}(Y^{+}|\mathcal{G}) - \mathbb{E}(Y^{-}|\mathcal{G})$$

Supongamos que $\mathcal{G}$ es la σ-álgebra generada por la *v.a.* X, i.e. $\mathcal{G} = \sigma(X)$($\mathcal{G}$ representa la información contenida en X). Luego $\mathbb{E}(Y|\mathcal{G}) = \mathbb{E}(Y|\sigma(X))$, que se escribe más simplemente como $\mathbb{E}(Y|X)$.

A continuación veremos las principales propiedades de la *esperanza condicional*. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad e Y una *v.a.* con $\mathbb{E}(|Y|) < \infty$.

(1) Si $Y = c$ en *c.t.p.*, $\Rightarrow \mathbb{E}(Y|\mathcal{G}) = c$ en *c.t.p.*.

(2) Si $Y \geq 0$, entonces $\mathbb{E}(Y|\mathcal{G}) \geq 0$ en *c.t.p.*.

(3) Sea X otra *v.a.* con $\mathbb{E}(|X|) < \infty$ y a, b constantes en $\mathbb{R}$, entonces

$$\mathbb{E}(aX + bY|\mathcal{G}) = a\,\mathbb{E}(X|\mathcal{G}) + b\,\mathbb{E}(Y|\mathcal{G})\ \ c.t.p.$$

(4) Sea $\mathcal{G} = \{\emptyset, \Omega\}$. Ahora $\mathcal{G}$ es la *menor* σ-álgebra (pues toda σ-álgebra de subconjuntos de Ω contiene a $\emptyset$ y Ω), y representa '*saber nada*'. Sea Y una *v.a.* sobre $\mathcal{G}$, entonces

$$\mathbb{E}(Y|\{\emptyset, \Omega\}) = \mathbb{E}(Y)$$

(5) Sea $\mathcal{G} = \mathcal{F}$. Ahora $\mathcal{G}$ es la '*mayor*' σ-álgebra y representa '*saber todo*'.

$$\mathbb{E}(Y|\mathcal{F}) = \mathbb{E}(Y)\ \ c.t.p.$$

(6) Si Y es $\mathcal{G}$-medible $\Rightarrow \mathbb{E}(Y|\mathcal{G}) = Y$ *c.t.p.*.

(7) Si Y es $\mathcal{G}$-medible, $\Rightarrow \mathbb{E}(Y\,Z\,|\mathcal{G}) = Y\,\mathbb{E}(Z\,|\mathcal{G})$ *c.t.p.*.

(8) Si $\mathcal{G}_0 \subset \mathcal{G}$, $\Rightarrow \mathbb{E}[\mathbb{E}(Y\,|\mathcal{G})|\mathcal{G}_0] = \mathbb{E}(Y\,|\mathcal{G}_0)$ *c.t.p.*.

(9) Si $\mathcal{G}_0 \subset \mathcal{G}$, $\Rightarrow \mathbb{E}[\mathbb{E}(Y\,|\mathcal{G}_0)|\mathcal{G}] = \mathbb{E}(Y\,|\mathcal{G}_0)$ *c.t.p.*.

(10) $\mathbb{E}[\mathbb{E}(Y\,|\mathcal{G})] = \mathbb{E}(Y\,)$ *c.t.p.*.

(11) Si Y es independiente de $\mathcal{G}$,

$$\mathbb{E}(Y\,|\mathcal{G}) = \mathbb{E}(Y\,)\ c.t.p..$$

Nos referimos, a continuación, a los principales modos de convergencia de variables aleatorias.

Definición 10.6. *Si* X_n, X *son variables aleatorias, decimos que* X_n *converge a* X casi seguramente

$$\mathrm{X}_n \to \mathrm{X}\ (n \to \infty)\ \text{c.s.}$$

si $\mathrm{X}_n \to \mathrm{X}$ *con* probabilidad *uno, i.e si*

$$\mathbb{P}(\{\omega : \mathrm{X}_n(\omega) \to \mathrm{X}(\omega)\ \textit{cuando}\ n \to \infty\}) = 1$$

Definición 10.7. *Si* X_n, X *son variables aleatorias, decimos que* X_n *converge a* X *en* probabilidad

$$\mathrm{X}_n \to \mathrm{X}\ (n \to \infty)\ \textit{en}\ \ \text{probabilidad}$$

si, para todo $\epsilon > 0$

$$\mathbb{P}(\{\omega : |\mathrm{X}_n(\omega) - \mathrm{X}(\omega)| > \epsilon\}) \to 0\ (n \to \infty).$$

Definición 10.8. *Si* X_n, X *son variables aleatorias en* L^p, *decimos que* X_n *converge a* X *en* L^p

$$\mathrm{X}_n \to \mathrm{X}\ (n \to \infty)\ \textit{en}\ \mathrm{L}^p$$

si,

$$\|\mathrm{X}_n - \mathrm{X}\|_p \to 0\ (n \to \infty).$$

esto es

$$\mathbb{E}(|\mathrm{X}_n - \mathrm{X}|^p) \to 0\ (n \to \infty).$$

Definición 10.9. *Si* X_n, X *son variables aleatorias, decimos que* X_n *converge a* X *en* distribución

$$\mathrm{X}_n \to \mathrm{X}\ (n \to \infty)\ \textit{en}\ \text{distribución}$$

si

$$\mathbb{P}(\{\mathrm{X}_n \le x\}) \to \mathbb{P}(\{\mathrm{X} \le x\})$$

$\forall\ x$ *en que* $\mathbb{P}(\{\mathrm{X} \le x\})$ *es continua.*

Teorema 10.3. *(*Ley débil de los Grandes Números*) Si* X_1, X_2, ... *son variables aleatorias independientes e idénticamente distribuidas con media* μ, *entonces*

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{X}_i \to \mu\ (n \to \infty)\ \textit{en}\ \text{probabilidad}$$

Teorema 10.4. *(*Teorema Central del Límite*) Si* X_1*,* X_2*, ... son variables aleatorias independientes e idénticamente distribuidas con media* μ *y varianza* σ^2*, entonces, siendo* $\mathrm{N}(0,1)$ *la* distribución normal estándar*,*

$$\frac{\sqrt{n}}{\sigma}\frac{1}{n}\sum_{i=1}^{n}(\mathrm{X}_i-\mu)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{(\mathrm{X}_i-\mu)}{\sigma}\rightarrow\ \mathrm{N}(0,1)\ (n\to\infty)\ \textit{en}\ \text{distribución}$$

esto es, para todo $x\in\mathbb{R}$*,*

$$\mathbb{P}(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{(\mathrm{X}_i-\mu)}{\sigma}\leq x)\to\Phi(x):=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{1}{2}y^2}\mathrm{d}y\ \textit{cuando}\ n\to\infty$$

El acceso a información completa, precisa y actualizada es claramente esencial para la participación activa en operaciones financieras. Nos vamos a referir únicamente a la situación en la que los agentes se basan en la información de dominio público y disponible a todos, para tomar sus decisiones. Vamos a suponer así mismo, que la información, una vez conocida, no se olvida y puede ser consultada en tiempo real. De esta manera, a medida que el tiempo pasa, nueva información esta disponible. Necesitamos, por lo tanto, un lenguaje matemático para modelar este flujo de información, que se revela con el transcurso del tiempo. Así surge la idea de *filtración.*
Tomemos como tiempo inicial $t=0$. El tiempo, podemos asumir, puede transcurrir de manera discreta o continua, en el caso discreto asumimos que el tiempo evoluciona con los enteros, i.e. $t=0,1,2,3,...$(e.g. cotizaciones diarias del mercado de acciones o cada hora o minuto, etc.). Podemos plantear un tiempo final $t=T$, o podemos contar con un horizonte temporal infinito (e.g. en el caso de futuros u opciones $t=T$ representa la fecha de expiración.).
Dado que se puede pensar que las σ-álgebras representan información o conocimiento, necesitamos una sucesión de σ-álgebras $\{\mathcal{F}_n : n=0,1,2,...\}$, creciente:

$$\mathcal{F}_n\subset\mathcal{F}_{n+1}\ (n=0,1,2,...),$$

con $\mathcal{F}_n$ representando la información disponible en $t=n$.
Vamos a suponer que todas la σ-álgebras son *completas* (esto puede ser evitado y no siempre es apropiado, pero simplifica el trabajo y es suficiente para nuestros propositos).
$\mathcal{F}_0$ representa la información inicial (si no hay ninguna, $\mathcal{F}_0=\{\emptyset,\Omega\}$), mientras que,

$$\mathcal{F}_\infty:=\lim_{n\to\infty}\mathcal{F}_n$$

representa todo lo que alguna vez sabremos. Una familia $\{\mathcal{F}_n : n=0,1,2,...\}$ se llama *filtración*; un espacio de probabilidad provisto de una tal familia, $\{\Omega,\{\mathcal{F}_n\},\mathcal{F},\mathbb{P}\}$, se denomina un *espacio de probabilidad filtrado.*
Para el caso particular, en que el *espacio muestral* es discreto, $\Omega=\{\omega_1,...,\omega_n\}$ y una σ-álgebra $\mathcal{F}$ en Ω dada (que en este caso es un álgebra), podemos encontrar una partición finita $\mathcal{P}=\{A_1,...,A_t\}$ de Ω, i.e. los conjuntos A_i son disjuntos y $\bigcup_{i=1}^{t}A_i=\Omega$, correspondiente a $\mathcal{F}$. Una *filtaración* $\mathcal{F}_n$, por lo tanto corresponde a una sucesión de particiones $\mathcal{P}_n$ cada vez más finas. En el momento $t=0$ los agentes solo saben que un evento $\omega\in\Omega$ sucederá, en el momento $T<\infty$ los agentes saben que evento especifico ω^* ha acontecido. Durante el fluir del tiempo

los agentes aprenden la estructura especifica de las (σ-) álgebras $\mathcal{F}_n$, lo que significa que aprenden las correspondientes particiones $\mathcal{P}_n$. Que la información contenida en $\mathcal{F}_n$ este revelada, es equivalente a saber a que $A_i^{(n)} \in \mathcal{P}_n$ el evento ω^* pertenece. Puesto que las particiones se vuelven cada vez más finas, la información sobre ω^* es cada vez más detallada.
Lamentablemente esta interpretación no se mantiene cuando Ω es infinito.

Definición 10.10. *Un* proceso estocástico $\mathrm{X} = \{\mathrm{X}_t : t \in \mathrm{I}\}$ *es una familia de* variables aleatorias *definidas sobre un* espacio de probabilidad *común, indexado por un conjunto de índicies* I.

Usualmente, como en nuestro caso, I representa el tiempo. Aquí, $I = \{0, 1, 2, ..., T\}$ (horizonte finito) o $I = \{0, 1, 2, ...\}$ (horizonte infinito). El *proceso estocástico* $X = (X_n)_{n=0}^{\infty}$ se dice *adaptado* a la filtración $(\mathcal{F}_n)_{n=0}^{\infty}$ si

$$X_n \text{ es } \mathcal{F}_n\text{-medible.}$$

Por lo tanto, si X esta *adaptada*, conoceremos el valor de X_n en el tiempo t =n. Si

$$\mathcal{F}_n = \sigma(X_0, X_1, ..., X_n)$$

llamamos a $(\mathcal{F}_n)$ la filtración *natural* de X. Por lo que un *proceso* esta siempre adaptado a su filtración *natural*.

Definición 10.11. *Un* proceso estocástico $\mathrm{X} = (\mathrm{X}_n)$ *es llamado una* martingala *con respecto a* $((\mathcal{F}_n), \mathbb{P})$ *si*

(1) X *esta* adaptada *(a* $(\mathcal{F}_n)$*):*
(2) $\mathbb{E}(|\mathrm{X}_n|) < \infty$ *para todo n;*
(3) $\mathbb{E}[\mathrm{X}_n|\mathcal{F}_{n-1}] = \mathrm{X}_{n-1}$, c.s. $(n \geq 1)$

X es una *supermartingala* si en lugar de (3) tenemos

$$\mathbb{E}[X_n|\mathcal{F}_{n-1}] \leq X_{n-1}\, c.s. \ (n \geq 1)$$

X es una *submartingala* si en lugar de (3) tenemos

$$\mathbb{E}[X_n|\mathcal{F}_{n-1}] \geq X_{n-1}\, c.s. \ (n \geq 1)$$

Nota: Si X_n representase la fortuna de un jugador después del n-ésimo juego y $\mathcal{F}_n$ representase su información acerca del juego en ese instante, (3) nos diría que su fortuna esperada después del juego $n+1$ es la misma que su fortuna presente. Luego una *martingala* representa un *juego justo*, una *supermartingala* un *juego desfavorable* y una *submartingala* un *juego favorable*.

Notemos lo siguiente:

(a) X es una *submartingala* (*supermartingala*) si y solo si $-X$ es una *supermartingala* (*submartingala*); X es una *martingala* si y solo si es una *submartingala* y una *supermartingala*.

(b) (X_n) es una *martingala* sii $(X_n - X_0)$ es una *martingala*. Por lo tanto podemos, sin pérdida de generalidad, tomar $X_0 = 0$ cuando resulte conveniente.

(c) Si X es una *martingala*, entonces para $m < n$, utilizando las propiedades de la esperanza condicional y la definición de las martingalas, tenemos (todas las igualdades son *casi seguramente*):

$$\mathbb{E}[X_n|\mathcal{F}_m] = \mathbb{E}[\mathbb{E}(X_n|\mathcal{F}_{n-1})|\mathcal{F}_m] = \mathbb{E}[X_{n-1}|\mathcal{F}_m] = ... = \mathbb{E}[X_{m+1}|\mathcal{F}_m] = X_m$$

Veamos algunos teoremas que hacen de las martingalas una herramienta tan importante.
Llamemos a X *acotada* en L^1 si

$$\sup_n \mathbb{E}(|X_n|) < \infty.$$

Teorema 10.5. *Sea* (X_n) *una* supermartingala acotada *en* L^1. *Entonces existe una* v.a X_∞ *con* $\mathbb{E}(|\mathrm{X}_\infty|) < \infty$ *tal que*

$$\mathrm{X}_n \to \mathrm{X}_\infty \ (n \to \infty) \text{ c.s.}$$

Teorema 10.6. *Una* martingala *acotada en* L^1 *converge* c.s..

Decimos que

$$X_n \to X_\infty \text{ en } L^1$$

si

$$E(|X_n - X_\infty|) \to 0 \ (n \to \infty).$$

Teorema 10.7. *Sea* $\mathrm{X} = (\mathrm{X}_n)$ *una* martingala. *Entonces las siguientes propiedades son equivalentes:*

(1) X_n *converge en* L^1.

(2) X_n *es acotada en* L^1 *y su límite* X_∞ *satisfacen*

$$\mathrm{X}_n = \mathbb{E}[\mathrm{X}_\infty|\mathcal{F}_n]$$

(3) *Existe una* v.a. *integrable* X *con*

$$\mathrm{X}_n = \mathbb{E}[\mathrm{X}|\mathcal{F}_n]$$

Definición 10.12. *Un proceso* $\mathrm{H} = (\mathrm{H})_{n=1}^{\infty}$ *se dice* predecible *si*

$$\mathrm{H}_n \textit{ es } \mathcal{F}_{n-1}\textit{-medible, para todo } n \geq 1.$$

H_n puede verse como la cantidad de dinero que un jugador apuesta en el instante n (H_0 no esta definida, dado que no hay jugada en $n = 0$). Esta decisión se basa en los resultados del juego en los instantes anteriores, i.e. $1, 2, ..., n-1$. H_n se puede entender como una estrategia de apuestas.
Supongamos que el juego consiste en arrojar una moneda y que por cada $1 que apostamos, ganamos $1, si sale cara y perdemos $1 si sale ceca.
La ganancia en el juego n estan dadas por, $H_n(\triangle X_n) = H_n(X_n - X_{n-1})$. donde X_n se definde como la ganancia neta luego de la n-ésima jugada (si se apostase $1 por juego). La ganancia total (neta) en el instante n es

$$Y_n = \Sigma_{k=1}^n H_k(\triangle X_k) = \Sigma_{k=1}^n H_k(X_k - X_{k-1}).$$

Escribimos

$$Y = H \bullet X,\ Y_n = (H \bullet X)_n,\ \triangle Y_n = H_n(\triangle X_n)$$

($(H \bullet X)_0 = 0$), y llamamos a $H \bullet X$ la *transformada martingala* de X por H.

Teorema 10.8. *(*Teorema de descomposición de Doob*) Sea* $\mathrm{X} = (\mathrm{X}_n)$ *un proceso adaptado con cada* $\mathrm{X}_n \in \mathrm{L}^1$. *Luego* X *se puede escribir (de forma única) como,*

$$\mathrm{X} = \mathrm{X}_0 + \mathrm{M} + \mathrm{A} :\ \ \mathrm{X}_n = \mathrm{X}_0 + \mathrm{M}_n + \mathrm{A}_n\ \forall n$$

con M *una* martingala *nula en cero y* A *un proceso predecible nulo en cero. Si* X *es además una* submartingala*,* A *es creciente:* $\mathrm{A}_n \leq \mathrm{A}_{n+1}$ *para todo* n*,* c.s..

Pensemos, ahora, en un *juego de apuestas*, o una serie de inversiones especulativas, en tiempo discreto. No hay ninguna jugada en $n = 0$; si las hay en $n = 1, 2, ...$ y

$$\triangle X_n := X_n - X_{n-1}$$

representa nuestra ganancia neta, por unidad apostada en la jugada n. Por lo que si X_n es una *martingala*, el juego es *justo*.

Teorema 10.9. (a) *Si* C *es un proceso no negativo, acotado y predecible, y* X *es una* supermartingala*,* $\mathrm{C} \bullet \mathrm{X}$ *es una* supermartingala *nula en cero.*

(b) *Si* C *es un proceso acotado y predecible y* X *es una* martingala*,* $\mathrm{C} \bullet \mathrm{X}$ *es una* martingala *nula en cero.*

Lema 10.1. *Una sucesión adaptada de* v.a. *reales e integrables* (M_n) *es una* martingala *sii para cualquier sucesión acotada y predecible* (H_n)*,*

$$\mathbb{E}(\sum_{k=1}^n \mathrm{H}_k(\triangle \mathrm{M}_k)) = 0\ (n = 1, 2, ...).$$

Definición 10.13. *Una variable aleatoria* Υ *que toma los valores* $\{0, 1, 2, ...; +\infty\}$ *se denomina* tiempo de parada *si*

$$\{\Upsilon \leq n\} = \{\omega : \Upsilon(\omega) \leq n\} \in \mathcal{F}_n\ \forall n \leq \infty.$$

equivalentemente

$$\{\Upsilon = n\} \in \mathcal{F}_n,\ n \leq \infty\ o\ \{\Upsilon \geq n\} \in \mathcal{F}_n,\ n \leq \infty.$$

Se puede pensar en Υ como en el instante en que se decide abandonar un juego de apuestas: que se abandone o no el juego en n depende únicamente de la historia hasta e inclusive el tiempo n, no el futuro.

Teorema 10.10. *(*Teorema de la parada opcional de Doob*) Sea* Υ *un* tiempo de parada*,* $X = (X_n)$ *una supermartingala y asumamos que se verifica alguna de las siguientes condiciones:*

(1) Υ *es acotada (*$\exists n \in \mathbb{N} : \Upsilon(\omega) \leq n,\ \forall \omega \in \Omega$*).(tiempo acotado).*

(2) $X = (X_n)$ *es acotada (*$\exists k : |X_n(\omega)| \leq k,\ \forall n, \omega$*).*

(3) $\mathbb{E}(\Upsilon) < \infty$ *y* $(X_n - X_{n-1})$*es acotada (*$\exists k : |X_n(\omega) - X_{n-1}(\omega)| \leq k,\ \forall n, \omega$*).*

entonces X_Υ *es integrable y*

$$\mathbb{E}(X_\Upsilon) \leq \mathbb{E}(X_0)$$

si X *es una martingala, entonces*

$$\mathbb{E}(X_\Upsilon) = \mathbb{E}(X_0)$$

11. Anexo II

Vamos a introducir el modelo desarrollado a comienzos de 1970 por Fischer Black, Myron Scholes y Robert Merton para la valuación de opciones sobre acciones, conocido como el modelo *Black-Scholes*

11.1. **Supuestos sobre la evolución de los precios de las acciones.**

Vamos a establecer algunos supuestos sobre como se da la evolución de los precios de las acciones a lo largo del tiempo. Dado el precio de una acción hoy, cuál es la distribución de probabilidad del mismo dentro de un día o una semana? El supuesto que subyace al modelo *Black-Scholes* es que, suponiendo que las acciones no pagan dividendos, los precios siguen un *camino aleatorio* (*random walk*). Luego el logaritmo de los cambios porcentuales del precio de una acción en un periodo de tiempo corto se distribuyen normalmente. Definimos:

$$\begin{cases} \mu: & \text{el retorno esperado por la acción, en un periodo corto de tiempo,} \\ \sigma: & \text{Volatilidad del precio de la acción.} \end{cases}$$

La media del cambio porcentual en el tiempo δt es $\mu\delta t$. La desviación estándar del cambio porcentual es $\sigma\sqrt{\delta t}$. El supuesto subyacente al modelo *Black-Scholes* es por lo tanto:

$$\ln\frac{\delta S}{S} \sim N(\mu\delta t, \sigma\sqrt{\delta t})$$

donde δS es el cambio en el precio de la acción, S, en el tiempo δt y $N(\mu,\sigma)$ denota una distribución normal con media μ y desviación estándar σ.

La distribución Lognormal Se puede demostrar que suponer que el precio de las acciones siguen un *camino aleatorio* implica que, en cualquier momento futuro, el precio de las acciones presentan una distribución *lognormal*. Una variable con una distribución lognormal tiene la propiedad de que su logaritmo natural se distribuye *normalmente*. Por lo tanto, bajo los supuestos del modelo *Black-Scholes* para el precio de las acciones, implica que $\ln(S_T)$ se distribuye normalmente, donde S_T es el precio de la acción en T. Se demuestra así mismo que la media y la desviación estándar de $\ln(S_T)$ son respectivamente,

$$\ln(S_0) + (\mu - \frac{\sigma^2}{2})\,T$$

y

$$\sigma\sqrt{T}$$

donde S_0 es el precio actual de la acción. Podemos escribir este resultado como,

$$\ln(S_T) \sim N(\ln(S_0) + (\mu - \frac{\sigma^2}{2})\,T, \sigma\sqrt{T})$$

Nota: Sea X una variable aleatoria cuyo logaritmo, Y, se distribuye normalmente con media μ y varianza σ^2; esto es $Y = \ln(X)$ y donde

$$f(Y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{Y-\mu}{\sigma})^2}$$

es la función de densidad de Y, entonces X se denomina variable aleatoria *lognormal*.

Con su esperanza y su varianza dadas por:

$$\mathbb{E}(X) = e^{\mu + \frac{\sigma^2}{2}}$$

y

$$\mathbb{V}ar(X) = e^{\sigma^2 + 2\mu}(e^{\sigma^2} - 1)$$

respectivamente.

El valor esperado o media de S_T, $E(S_T)$, esta dado por,

$$E(S_T) = S_0 e^{\mu T + \frac{\sigma^2 T}{2}}$$

lo que coincide con la definición de μ como el retorno esperado. La varianza de S_T, $var(S_T)$, se puede ver que esta dada por

$$var(S_T) = e^{\sigma^2 T}(S_0)^2 e^{2\mu T}(e^{\sigma^2 T} - 1)$$

11.2. **Como estimar la volatilidad σ.** Una manera de hacerlo, es mediante el uso de series históricas, es decir una serie de precios de la acción a lo largo del tiempo. El precio de una acción es, usualmente, observado a intervalos fijos de tiempo (e.g. cada día, cada semana). Vamos a definir lo siguiente,

$$\begin{cases} n+1: & \text{número de observaciones,} \\ S_i: & \text{el precio de la acción al finalizar el } i\text{-ésimo intervalo de tiempo } (i\text{=0, 1, 2,...,}n), \\ \tau: & \text{duración de cada intervalo de tiempo, expresado en años.} \end{cases}$$

y sea

$$u_i = \ln(\frac{S_i}{S_{i-1}})$$

Una estimación, s, de la desviación estándar de u_i esta dada por

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(u_i - \bar{u})^2}$$

donde $\bar{u}$ es la media de los u_i. Sabemos que la desviación estándar de u_i es $\sigma\sqrt{\tau}$. Por lo tanto la variable s es una estimación de $\sigma\sqrt{\tau}$. De donde se desprende que la misma σ puede ser estimada por $\hat{\sigma}$, donde

$$\hat{\sigma} = \frac{s}{\sqrt{\tau}}$$

El error estándar de esta estimación es aproximadamente $\frac{\hat{\sigma}}{\sqrt{2n}}$. Elegir un n correcto no es fácil. Mayor cantidad de datos generalmente llevan a más precisión, pero σ cambia a lo largo del tiempo y datos muy viejos pueden resultar irrelevantes para predecir el futuro.

11.3. **Supuestos del modelo *Black-Scholes*.** A la hora de derivar la fórmula de fijación del precio de opciones, Black y Scholes establecieron los siguiente supuestos:

- El comportamiento del precio de las acciones corresponde al modelo *log-normal* (desarrollado anteriormente) con μ y σ constantes.
- No existen costos de transacción o impuestos. Todos las *seguridades* son perfectamente divisibles.
- No se pagan dividendos por la acción durante la vida de la opción.
- No hay oportunidades de arbitraje.
- La negociación de las *seguridades* es continua.
- Los inversores pueden prestar o pedir prestado a la misma tasa de interés libre de riesgo.
- La tasa de interés libre de riesgo de corto plazo, r, es constante (expresada con capitalización instantánea).

El análisis de *Black-Scholes/Merton* considera un portafolio que no implica riesgo alguno, consistente de una posición en la opción y una posición en la acción subyacente. En la ausencia de oportunidades de arbitraje, el retorno del portafolio debe ser la tasa libre de reisgo r. Esto deviene en una ecuación diferencial que debe ser satisfecha por la opción. La razón por la cual se puede construir un portafolio libre de riesgo (en el corto plazo) es que los precios de la acción y de la opción son afectados por la misma fuente de incertidumbre, el movimiento del precio de la acción (i.e. están perfectamente correlacionados). Lo que significa que cuando un portafolio de estas características es creado, lo que se pierde con el precio de la acción se gana con el de la opción y viceversa.
Las fórmulas para la valuación de calls y puts europeos sobre acciones que no pagan dividendos son,

$$c = S_0 N(d_1) - K e^{-rT} N(d_2)$$
$$p = K e^{-rT} N(-d_2) - S_0 N(-d_1)$$

donde

$$d_1 = \frac{\ln(\frac{S_0}{K}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(\frac{S_0}{K}) + (r - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

La función $N(x)$ es la distribución acumulativa normal estándar. La notación es conocida, K es el precio de ejercicio, T es el tiempo hasta la fecha de ejercicio, etc. En teoría las fórmulas de *Black-Scholes* son correctas únicamente si la tasa de interés de corto plazo, r, es constante.

Bibliografía

[1] HULL, John C..2002. *Fundamentals of Futures and Options Markets.* 4 ed.–New Jersey: Prentice Hall. 502p.

[2] GOODMAN, V., STAMPFLI, J. 2001. *The Mathematics of finance: Modeling and Hedging.* United States of America: Brooks/Cole. 259p.

[3] BINGHAM, N.H., RUIDIGER, K. 1998. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives.* Great Britain: Springer Finance. 293p.

[4] GILBERT, G.W. 1994. *Wavelets and other orthogonal Systems with applications.*U.S.A: CRC Press.248p.

[5] CATUOGNO,P.J., FERRANDO,S.E., GONZALEZ,A.L. 2004. *Haar Wavelets Systems for hedging Financial Derivatives.* Relatorio de Pesquisa 28/05 IMECC-UNICAMP.

[6] LONGSTAFF, F.A., SCHWARTZ, E.S. 2001.*Valuing American options by simulation: a simple least-squares approach.* Rev. Financial Stud.14,113-148.

[7] CLEMENT, E., LAMBERTON, D., PROTTER, P. 2002 *An analysis of a least squares regression method for American option pricing.* Finance and Stochastics 6 (4), 449-471.

[8] GUNDY, F.R. 1996 *Martingale theory and pointwise convergence of certain orthogonal systems.* Trans. Amer. Math. Soc., 124, 228-248.

[9] KOLMOGOROV, A.N., FOMIN, S.V. 1970 *Introductory real analysis.* United States of America: Dover Publications 403p.

[10] SHIRYAYEV, A.N. 1984 *Probability* New York : Springer Verlag.

[11] CARRIERE J. F., 1996, *Valuation of early-exercice price of options using simulation and non-parametric regression.* Insurance: Mathematics and Economics 19, 19-30.

[12] TSITSIKLIS, J. N., VAN ROY, B. 1999. *Optimal stopping of Markov Processes: Hilbert space theory, aproximation algorithms, and an application to pricing high-dimensional financial derivatives.*IEEE Trasactions on Automatic Control 44 (10), 1840-1851.