



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**OPTIMIZACIÓN DE OPCIONES AMERICAS CON
FUNCIÓN DE PAGO CONVEXA, EN TIEMPO
DISCRETO**

T E S I S

QUE PARA OBTENER EL TÍTULO DE :

A C T U A R I O

P R E S E N T A :

JUAN MORALES VELASCO



FACULTAD DE CIENCIAS
UNAM

TUTOR : DRA. GUADALUPE CARRASCO LICEA

2007



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

<p>1. Datos del alumno Juan Morales Velasco 56 76 72 50 Universidad Nacional Autónoma de México Facultad de Ciencias Actuaría 09710905-9</p>
<p>2. Datos del tutor Dra. Guadalupe Carrasco Licea</p>
<p>3. Datos del sinodal 1 Dr. Pablo Padilla Longoria</p>
<p>4. Datos del sinodal 2 M. en C. Julio César Cedillo Sánchez</p>
<p>5. Datos del sinodal 3 Dr. Juan González Hernández</p>
<p>6. Datos del sinodal 4 M. en A.P. María del Pilar Alonso Reyes</p>
<p>7. Datos del trabajo escrito Optimización de opciones americanas con función de pago convexa, en tiempo discreto. 84 p 2007</p>

Dedicatoria:

Este trabajo lo dedico a esas personas que siempre han estado conmigo confiando en mí y apoyándome, mi familia:

A mi mamá Lucero por obligarme a hacer las cosas, mis abuelitos Juan y Lidia por su cariño, amor y dedicación, y mis hermanos Carlos y Diana por ser eso por ser mis hermanos.

Agradecimientos:

Agradezco enormemente a la Dra. Guadalupe Carrasco Licea por su paciencia, consejos y apoyo durante la elaboración de este trabajo que sin ella no hubiera sido posible.

A mis sinodales Ma. del Pilar Alonso Reyes, Pablo Padilla Longoria, Juan González Hernández y Julio César Cedillo Sánchez por su tiempo y consejos durante la realización de mi tesis.

A todos los profesores de la facultad de ciencias que con su esfuerzo y dedicación sembraron una semilla de conocimientos en mí.

A mis amigos que estuvieron a mí lado apoyándome durante mis estudios y algunos hasta en la elaboración de mi tesis.

Gracias, muchas gracias a todas esas personas que han estado a mí lado sin más interés que ser un amigo más.

Contenido

1	Introducción	1
2	Conceptos básicos	4
2.1	Introducción	4
2.2	Probabilidad	4
2.2.1	Espacios de probabilidad	5
2.2.2	Variables aleatorias	7
2.2.3	Probabilidad Condicional	9
2.3	Procesos estocásticos	10
2.4	Procesos controlados de Markov	12
2.5	Problema: describir la estructura de los dominios de paro óptimo de una opción americana de compra	14
3	Teorema de paro óptimo, Shiryaev (1978)	17
3.1	Introducción	17
3.2	Teorema de paro óptimo, Shiryaev (1978).	17
4	Optimización de opciones Americanas, con función de pago convexa, a tiempo discreto	28
4.1	Modelo a tiempo discreto de paro óptimo	29
4.2	Optimización en el caso de una función de pago lineal	34
4.3	Optimización en el caso cuando la función de pago es lineal por trozos.	52
4.4	Optimización en el caso de función de pago convexa en general	56
5	Ejemplos	62
5.1	Elaboración del algoritmo para aproximar los dominios de paro óptimo	63
5.2	Ejemplos numéricos	65
6	Conclusiones	70
	Apéndices	72
A	Algoritmo	72
B	Datos simulados	76

CONTENIDO

ii

Bibliografia

83

Capítulo 1

Introducción

En el mercado financiero los productos derivados son cada vez más usados. Dentro de los productos derivados se encuentran los forwards, futuros, opciones y swaps, que tienen la característica común de estar vinculados al valor de un activo subyacente o de referencia. Este activo subyacente puede ser una acción, un índice, un tipo de cambio de divisas, un producto agrícola o mineral o cualquier otro producto que se desee utilizar como activo subyacente.

El objetivo para el cual fueron creados los productos derivados es el de ofrecer una cobertura frente al riesgo debido a la fluctuación del valor del activo subyacente en el mercado; es decir, estos instrumentos buscan disminuir o anular el riesgo producido por invertir en el activo subyacente de referencia. En la actualidad, aparte de la protección al riesgo, existen personas que buscan la especulación pura sobre los productos derivados.

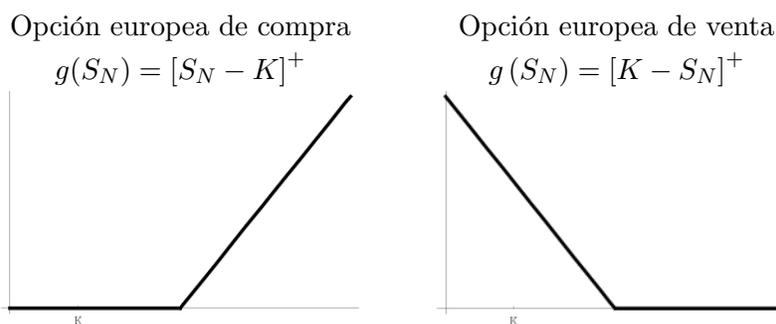
Esta tesis se enfoca al estudio de las opciones, específicamente a las opciones americanas de compra. Una opción es un contrato que da el derecho de comprar o vender un activo subyacente, más no la obligación de hacerlo, a un precio pactado. En las opciones participan dos personas: el emisor y el comprador. El emisor recibe una prima al inicio del contrato, y el comprador obtiene el derecho a comprar o vender el activo subyacente a un precio fijo denotado por K .

Una primera división entre las opciones se debe al tipo de derecho que dan al comprador: las opciones que dan el derecho de compra son llamadas *de compra*, si el derecho que brindan es el de venta son conocidas como *de venta*.

A su vez las opciones de venta y compra se dividen en europeas o americanas de acuerdo al momento en el que puede ser ejercido el derecho adquirido. Las opciones europeas sólo permiten ejercer el derecho adquirido en la fecha de expiración de la opción, fecha que será denotada por N . Por otro lado las opciones americanas pueden ser ejercidas en cualquier momento comprendido entre la fecha de inicio y la fecha de

expiración de la opción.

El ejercer o no el derecho adquirido por el comprador depende del precio del activo subyacente en el mercado. Denótese por S_t al precio del activo subyacente en el mercado al tiempo t . El poseedor de opciones europeas de compra ejercerá su derecho a comprar el activo subyacente al precio pactado K siempre y cuando S_N sea mayor o igual a K , ya que en caso contrario ejercer la opción sería comprar el activo a un precio mayor al de mercado. En el caso de tener una opción europea de venta, la decisión de ejercerla se toma cuando $S_N \leq K$ mientras que la decisión de no ejercerla es tomada cuando $S_N > K$. Las funciones de pago generadas por la decisión de ejercicio o no del poseedor de una opción europea tienen la siguiente forma



La decisión que debe tomar el poseedor de opciones europeas para maximizar su ganancia es sencilla, puesto que sólo tiene dos posibles decisiones: ejercer o no ejercer la opción en el tiempo N , y esta decisión queda completamente determinada por el precio en el mercado del activo subyacente en la fecha de vencimiento.

El poseedor de una opción americana tiene un panorama de decisiones más complejo para maximizar su ganancia, ya que éste debe elegir un tiempo de ejercicio en el intervalo comprendido entre la fecha de inicio y la fecha de expiración de la opción, intervalo que será denotado por $[0, N]$. Debido a que el precio del activo subyacente en el mercado no es una función determinista del tiempo, el poseedor de la opción no puede conocer con precisión el momento t en el cual ejercer la opción para maximizar su ganancia. Para todo tiempo $t \in [0, N]$ se tiene una función de ganancias similar a la descrita en las opciones europeas, es decir, en el caso de que una opción americana de compra sea ejercida en el tiempo t la función de ganancias es $g(S_t) = [S_t - K]^+$, mientras que en el caso de una opción americana de venta que es ejercida en el tiempo t es $g(S_t) = [K - S_t]^+$.

En esta tesis se discute el problema de como encontrar el tiempo $t \in [0, N]$ que maximice la ganancia esperada del poseedor de una opción americana de compra, lo cual se hará a través de describir los dominios de paro óptimo de la opción. Los dominios de paro óptimo son una sucesión de subconjuntos $\{\Gamma_t\}_{t \in [0, N]}$ de \mathbb{R}^+ que permiten

decidir si ejercer la opción al momento $t \in \{0, 1, \dots, N-1\}$ o esperar al tiempo $t+1$, esta decisión se hace al ver si el precio del activo subyacente S_t está en su correspondiente dominio de paro óptimo Γ_t o no.

Para analizar este problema se hará uso de las herramientas proporcionadas por la teoría de la probabilidad, procesos estocásticos y procesos controlados de Markov, mismas que serán brevemente expuestas en el capítulo uno.

En el capítulo dos se abordará la exposición y demostración del teorema de paro óptimo creado por Shiryaev en 1978 para procesos estocásticos a tiempo discreto, así como la variante de este teorema en el caso donde el proceso estocástico es de Markov.

Con las herramientas que se desprenden del teorema de paro óptimo se abordará el problema central de esta tesis en el capítulo III, "Optimización de opciones americanas con función de pago convexa". En primer lugar se encontrarán los dominios de paro óptimo en el caso de las opciones americanas, en el segundo paso se extienden estos dominios al caso donde la función de pago es convexa y lineal por trozos, por último se encontrarán los dominios de paro óptimo para funciones de pago convexas en general. Este capítulo está basado en el artículo *Optimal pricing for American type options with discrete time* [5].

Para finalizar esta tesis, en el capítulo cuatro se expondrán dos ejemplos numéricos de los dominios de paro óptimo suponiendo que el proceso del precio es una caminata aleatoria geométrica con variables Bernoulli en un primer caso y lognormal en segundo caso. Este capítulo está basado en el artículo *Algorithms and Programs for optimal Monte Carlo pricing of American options* [9].

Capítulo 2

Conceptos básicos

2.1 Introducción

En el presente capítulo se expondrán de manera breve los conceptos básicos de probabilidad, procesos estocásticos y procesos controlados de Markov; estos conceptos son necesarios para dar solución al problema de determinar el dominio de paro óptimo de una opción financiera.

La exposición está basada en los libros: *Estadística I: probabilidad* [1], *Probability* [8], *Introducción a la teoría de la probabilidad* [3], *A first course in probability* [11] y *Discrete-time Markov control processes: basic optimality criteria* [4].

2.2 Probabilidad

La probabilidad es una rama de las matemáticas que estudia el comportamiento de los experimentos aleatorios, es decir, aquellos en los que no es posible determinar el resultado que se obtendrá antes de la realización del experimento. El problema inicial es cómo asignar a cada posible resultado una medida que brinde información sobre qué tan posible es que ocurra cada resultado al realizar el experimento.

El conjunto de todos los resultados posibles de un experimento aleatorio se conoce como *espacio muestral* y será denotado por Ω . A los elementos del espacio muestral se les conoce como *eventos elementales* o *puntos muestrales*; por ejemplo, en el caso del lanzamiento de un dado su espacio muestral es el conjunto $\{1, 2, 3, 4, 5, 6\}$, donde cada número por si solo es un evento elemental de dicho espacio muestral; cuando se lanza una moneda su espacio muestral es el conjunto $\{\text{águila}, \text{sol}\}$ y sus dos únicos eventos elementales son $\{\text{águila}\}$ y $\{\text{sol}\}$; en todo experimento aleatorio existe el espacio muestral y sus correspondientes eventos elementales, aunque existen casos donde la descripción del espacio muestral y de sus eventos elementales no pueden ser descritos

de forma sencilla.

Un jugador de dados puede estar interesado en resultados más complejos al de un evento elemental, que salga un número en específico; por ejemplo que se obtenga un número par o un número mayor que 3. Cualquier evento que desee analizar el jugador, se puede crear por medio de operaciones de conjuntos de los eventos elementales, que se obtenga un número par es la unión de los eventos $\{2\}$, $\{4\}$ y $\{6\}$; que se obtenga un número mayor a 3 se construye por el complemento de $\{1, 2, 3\}$. Esta necesidad de preguntarse por eventos más complejos a los elementales se extiende a cualquier experimento aleatorio.

De esta manera, un *evento* es una característica que puede o no aparecer en los resultados de un experimento aleatorio. Al definir un evento es indispensable que no exista ambigüedad alguna de si el resultado del experimento una vez realizado posee o no la característica indicada por el evento. En este caso, todo evento es identificado con el subconjunto de Ω de todos los elementos que tienen dicha característica descrita por él. Así por ejemplo, el evento: "que salga un número par al lanzar un dado" se representa por el subconjunto $\{2, 4, 6\}$ de $\Omega = \{1, 2, 3, 4, 5, 6\}$ y el evento "obtener un número mayor que 3" por el subconjunto $\{4, 5, 6\}$.

Como se observa, el conocer el espacio muestral y los eventos elementales no es suficiente para contestar preguntas sobre un experimento aleatorio, se necesita una familia de subconjuntos del espacio muestral que contenga por lo menos los eventos no simples en los cuales se puede tener interés. La definición de esta familia se dará en conjunto con lo que es una medida de probabilidad.

2.2.1 Espacios de probabilidad

Como se ha mencionado en la teoría de la probabilidad lo que se busca es definir una medida que permita medir la incertidumbre, *medida de probabilidad o función de probabilidad*, de los posibles eventos de un espacio muestral Ω .

Una medida de probabilidad es una función que asigna a cada evento un número real entre cero y uno, esto quiere decir que el dominio de la medida de probabilidad es una colección de subconjuntos de Ω . Para que la medida de probabilidad esté bien definida, su dominio debe permitir la construcción de cualquier evento usando operaciones conjuntistas en los elementos de Ω , es por esto que el dominio de una medida de probabilidad debe ser una σ -álgebra contenida en el conjunto potencia de Ω , denotado por $\mathcal{P}(\Omega)$.

Definición 2.2.1 *Se dice que $\mathcal{F} \subset \mathcal{P}(\Omega)$ es una σ -álgebra si cumple:*

1. $\Omega \in \mathcal{F}$,

2. si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$, y
3. si $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{F}$, entonces $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$.

La σ -álgebra \mathcal{F} sobre la cual se define la medida de probabilidad es el conjunto de todos los eventos en interés a estudiar del fenómeno aleatorio. Si C es una clase de subconjuntos de Ω , se puede construir la σ -álgebra generada por C , que es denotada por $\sigma(C)$, que es la σ -álgebra más pequeña que contiene a C , es decir, si \mathcal{H} es una σ -álgebra que contiene a C , entonces $\sigma(C) \subseteq \mathcal{H}$; como $\{\emptyset, \Omega\}$ y $\mathcal{P}(\Omega)$ son σ -álgebras y $C \subset \Omega$, siempre debe existir $\sigma(C)$.

Una σ -álgebra de importancia en la probabilidad debido a su gran utilidad, es la σ -álgebra de Borel. Esta es la σ -álgebra generada por la topología de \mathbb{R} , es la misma que se genera por la familia de intervalos abiertos $\mathcal{J} = \{(a, b) : a, b \in \mathbb{R}\}$, y es denotada por $\mathcal{B}(\mathbb{R})$. La σ -álgebra de Borel contiene a todo tipo de intervalos, ya que cualquier intervalo puede obtenerse mediante complementos, uniones numerables o intersecciones numerables de intervalos abiertos. Además, contiene a cualquier subconjunto abierto de \mathbb{R} , puesto que éstos son siempre uniones numerables de intervalos abiertos, además, como $\mathcal{B}(\mathbb{R})$ es una σ -álgebra debe contener a los complementos de todos los intervalos abiertos, entonces contiene a todos los conjuntos cerrados de números reales. A cada conjunto de $\mathcal{B}(\mathbb{R})$ se le conoce como *boreliano*.

La pareja (Ω, \mathcal{F}) formada por un espacio muestral $\Omega \neq \emptyset$ y una σ -álgebra \mathcal{F} de subconjuntos de Ω , se le conoce como *espacio medible*, ya que en ella se puede definir una medida.

En el caso de un experimento aleatorio, la medida que se define sobre (Ω, \mathcal{F}) debe brindar información sobre la incertidumbre de ocurrencia de cada resultado posible del experimento. La definición de medida de probabilidad está dada por los axiomas propuestos por Kolmogorov en 1933.

Definición 2.2.2 Una terna (Ω, \mathcal{F}, P) es un espacio de probabilidad y P es una medida de probabilidad, si Ω es un conjunto no vacío, \mathcal{F} una σ -álgebra de subconjuntos de Ω y P una función de \mathcal{F} en \mathbb{R} , que cumplen con:

1. $P(A) \geq 0$ para toda $A \in \mathcal{F}$,
2. si la colección $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{F}$ está formada por conjuntos ajenos, entonces

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n), \text{ y}$$

3. $P(\Omega) = 1$.

El matemático Henri Léon Lebesgue se planteó el problema de extender el concepto de longitud a una familia tan grande como sea posible de subconjuntos de números reales, es decir, comenzar asignando a cada intervalo su longitud y después extender esa función lo más posible. Esta extensión se puede hacer hasta abarcar la σ -álgebra generada por los intervalos, es decir, la σ -álgebra de Borel. Pero los borelianos, si bien son una familia muy grande, tienen la limitación de que hay conjuntos de $\mathcal{P}(\mathbb{R})$ de medida cero que no son borelianos. En términos de probabilidad esto significa que hay subconjuntos de conjuntos con probabilidad cero para los cuales no está definida la probabilidad, lo cual es un problema; por ello, es conveniente considerar una familia de subconjuntos de \mathbb{R} aun más grande que los borelianos.

Definición 2.2.3 *Se dice que un conjunto $A \subset \mathbb{R}$ es Lebesgue-medible, si pertenece a la σ -álgebra generada por los borelianos unión los conjuntos de medida cero.*

De esta manera, todo conjunto Lebesgue-medible se puede escribir como la unión de un boreliano y un conjunto de medida cero. La medida de Lebesgue, denotada por λ , en el $(0, 1)$ es la única medida de probabilidad definida sobre los conjuntos Lebesgue-medibles contenidos en el $(0, 1)$ que asigna a cada intervalo su longitud. De aquí en adelante cuando se hable de la σ -álgebra de Borel se considerará a los borelianos completados, es decir, a la σ -álgebra de los Lebesgue-medibles.

En general, este método se conoce como la completación de una medida; se dice que un espacio de probabilidad (Ω, \mathcal{F}, P) es completo si para toda $A \in \mathcal{F}$ tal que $P(A) = 0$, se cumple que todo conjunto $B \subset A$ está en \mathcal{F} , es decir que, \mathcal{F} contiene a todos los subconjuntos de conjuntos de medida cero en \mathcal{F} .

2.2.2 Variables aleatorias

El espacio muestral Ω puede tener elementos que no son números reales, entonces, la medida de probabilidad está definida en un dominio \mathcal{F} que no es una clase de subconjuntos de \mathbb{R} , lo cual puede dificultar el manejo de la σ -álgebra y de la medida de probabilidad definida en ella, por ello resulta útil poder transformar un espacio de medida (Ω, \mathcal{F}) en un espacio de medida real $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ o un subconjunto de él. Esta transformación debe hacerse de manera que se respete la estructura de la σ -álgebra \mathcal{F} construida en Ω , este cambio se hace a través de *funciones medibles*.

Una función $f : (\Omega, \mathcal{F}) \rightarrow (\Omega^*, \mathcal{F}^*)$, donde (Ω, \mathcal{F}) y $(\Omega^*, \mathcal{F}^*)$ son espacios medibles, se dice que es $\mathcal{F} - \mathcal{F}^*$ -medible, si $f^{-1}(A) \in \mathcal{F}$ para toda $A \in \mathcal{F}^*$. Si existe una función f , $\mathcal{F} - \mathcal{F}^*$ -medible, entonces se puede trabajar en el espacio $(\Omega^*, \mathcal{F}^*)$ en lugar de (Ω, \mathcal{F}) , debido a que existe una función que permite exportar la estructura del espacio (Ω, \mathcal{F}) en $(\Omega^*, \mathcal{F}^*)$.

Dentro de las funciones medibles son de especial importancia las funciones de imagen real, $f : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, conocidas como Borel medibles.

Una *variable aleatoria* X es una función Borel-medible definida sobre (Ω, \mathcal{F}) . La construcción de variables aleatorias permite analizar un tipo particular de eventos de un experimento aleatorio. Por ejemplo, al lanzar dos dados regulares, la variable aleatoria definida por $X(a, b) = a + b$ donde $(a, b) \in \{(x, y) \mid x, y \in \{1, 2, 3, 4, 5, 6\}\} = \Omega$ permite analizar el comportamiento global de las distintas probabilidades que corresponden a los valores que puede tomar la suma.

Dada cualquier variable aleatoria X , es posible definir su función de distribución $F_X : \mathbb{R} \rightarrow \mathbb{R}$ por medio de:

$$F_X(x) = P[X \leq x].$$

Esta función condensa toda la información probabilística de la variable aleatoria X y con base en ella se puede calcular cualquier probabilidad que involucre sólo a X .

En un experimento aleatorio no se puede responder con certeza a la pregunta: ¿cuál es el resultado que tendrá la variable aleatoria X en la realización del experimento? Pero sí se puede calcular el valor promedio ponderado por la probabilidad de cada uno de los valores de X . El concepto de *valor promedio o valor esperado* de una variable aleatoria es de importancia, ya que resume la información que se tiene sobre la incertidumbre de X en un único valor. El valor esperado de una variable aleatoria es un índice de referencia de los valores promedios de la variable.

Definición 2.2.4 *La esperanza de una variable aleatoria X definida en un espacio de probabilidad (Ω, \mathcal{F}, P) , está dada por*

$$E[X] = \int_{\Omega} X dP$$

siempre y cuando la integral exista, es decir, $E[X^+] = \int_{\Omega} \max\{0, X\} dP < \infty$ o $E[X^-] = \int_{\Omega} \max\{0, -X\} dP < \infty$.

Como cualquier integral, la esperanza de una variable aleatoria es lineal:

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E X_i,$$

donde a_i es una constante y X_i una variable aleatoria para toda i .

Una medida de la dispersión de los valores de una variable aleatoria X respecto a su valor esperado es la *varianza* de la variable aleatoria.

Definición 2.2.5 *La varianza, denotada por $\sigma^2(X)$, de una variable aleatoria X está definida como*

$$\sigma^2(X) = \begin{cases} E[X - E(X)]^2 = E[X^2] - (E[X])^2 & \text{si } E[X^2] < \infty \\ +\infty & \text{si } E[X^2] = +\infty \end{cases}$$

2.2.3 Probabilidad Condicional

Cuando se tiene información de que ha ocurrido cierto evento $B \subset \mathcal{F}$, y con $P(B) > 0$, pero el resultado exacto se desconoce, esta información altera la probabilidad de que ocurra cualquier otro evento $A \subset \mathcal{F}$, es decir, la medida de probabilidad debe reflejar la información con la que se cuenta. En este caso, el espacio muestral Ω se ve reducido a B como se muestra en la siguiente definición:

Definición 2.2.6 Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Si $B \in \mathcal{F}$ tal que $P(B) > 0$, la probabilidad condicional dada la ocurrencia de B , denotada por P_B o $P(\cdot | B)$, se define como

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \quad \text{para toda } A \in \mathcal{F}.$$

La función $P_B(A)$, inducida por P y B para cada $A \in \mathcal{F}$, es una medida de probabilidad.

De forma análoga, cuando se tiene información sobre la ocurrencia de un evento B con probabilidad positiva, se define la esperanza condicional de X dado el evento $B \in \mathcal{F}$ como

$$E_B[X] = E[X|B] = \int_{\Omega} X dP_B = \frac{1}{P(B)} \int_B X dP$$

La *esperanza condicional* es una herramienta muy útil, ya que análogamente a la esperanza, es un índice del promedio ponderado por la probabilidad condicional a un evento con medida positiva.

Análogamente se puede preguntar uno, cómo se ve afectada la probabilidad de una variable aleatoria X si la condicionamos a la ocurrencia de una variable aleatoria Y . Para contestar esto, considere un espacio de probabilidad (Ω, \mathcal{F}, P) donde están definidas las variables aleatorias X y Y . Si se desea condicionar a la variable X con Y , es porque Y debe estar relacionada de alguna forma con el comportamiento de X , en otro caso es inútil el condicionamiento puesto que no afectará en ningún sentido.

Para definir la probabilidad condicional de X dada Y se debe considerar a la σ -álgebra \mathcal{G} generada por Y que debe estar contenida en \mathcal{F} .

Definición 2.2.7 Sea X una variable aleatoria integrable definida en (Ω, \mathcal{F}, P) y sea $\mathcal{G} \subset \mathcal{F}$ una sub- σ -álgebra. La esperanza condicional de X dada \mathcal{G} es cualquier función $Y : (\Omega, \mathcal{G}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ Borel medible e integrable, tal que

$$\int_E Y dP = \int_E X dP \quad \text{para todo } E \in \mathcal{G}.$$

En tal caso se dice $Y = E[X|\mathcal{G}]$.

A partir de la definición de la esperanza condicional, se extrae el concepto de probabilidad condicional dada una variable aleatoria Y o su respectiva σ -álgebra, puesto que

$$P(X \in A|Y) = P(X \in A|\mathcal{G}) = P_{\mathcal{G}}(A) = P(A|\mathcal{G}) = E[A|\mathcal{G}], \text{ para toda } A \in \mathcal{F}.$$

2.3 Procesos estocásticos

En ciertos fenómenos aleatorios como el comportamiento del precio de una acción, el número de individuos en la fila de un banco, la cantidad de agua en una presa, etc., se tiene que para cada momento en el tiempo t existe una variable aleatoria X_t que describe al fenómeno, a esta sucesión de variables aleatorias se les conoce como *proceso estocástico*.

Formalmente un proceso estocástico es una sucesión de variables aleatorias $\{X_t\}_T$ ordenadas en el tiempo, esto quiere decir que para cada $t \in T \subset \mathbb{R}^+$ X_t es una variable aleatoria definida en un espacio de probabilidad (Ω, \mathcal{F}, P) . El parámetro t comúnmente es interpretado como el tiempo, por lo que a X_t se le conoce como el estado del proceso al tiempo t .

Al conjunto T se le conoce como el índice del proceso. Cuando T es un conjunto numerable se dice que el proceso estocástico es a tiempo discreto, en cambio si T es un intervalo de la recta real se dice que el proceso es a tiempo continuo.

Un concepto de importancia dentro de los procesos estocásticos es la *probabilidad de transición del tiempo s al tiempo t* , ya que esta brinda información sobre la incertidumbre de que el proceso al tiempo t se encuentre en el evento $A \in \mathcal{F}$, dado que al tiempo $s < t$ se encontró en el estado $X_s \in \mathcal{F}$. La probabilidad de transición de s a t está denotada por

$$P(X_t \in A|X_s \in B) \text{ para } t < s \in \mathbb{R}^+ \text{ y } A, B \in \mathcal{F}.$$

Cuando el proceso estocástico es a tiempo discreto, también conocido como *cadena*, es más cómodo utilizar la probabilidad de transición a un paso, de n a $n+1$, a la cual se le suele llamar simplemente probabilidad de transición,

$$P(X_{n+1} \in A|X_n \in B), \text{ para toda } n \in \mathbb{N} \text{ y } A, B \in \mathcal{F}.$$

Como se ha visto un proceso estocástico $\{X_t\}_T$ se ha definido en un espacio de probabilidad (Ω, \mathcal{F}, P) , en este caso se hace referencia a un *proceso estocástico homogéneo* puesto que la medida de probabilidad es P para toda $t \in T$, esto quiere decir, $P(X_t \in A) = P(X_s \in A)$ para toda $t, s \in T$ y $A \in \mathcal{F}$. Cuando la medida de

probabilidad depende del tiempo se habla de un proceso estocástico no homogéneo, es decir, $P(X_t \in A) \neq P(X_s \in A)$ para algunas $t, s \in T$ y $A \in \mathcal{F}$.

Una *filtración de σ -álgebras* en Ω , es una sucesión de σ -álgebras $\{\mathcal{F}_t\}_{t \in T}$ en Ω , tales que $\mathcal{F}_s \subset \mathcal{F}_t$ para toda $s < t \in T$.

En un proceso estocástico $\{X_t\}_{t \in T}$ es importante hablar de su *filtración natural*, ya que esta es la sucesión creciente de σ -álgebras $\mathcal{F}_s = \sigma\{X_t : t \leq s\}$, esto quiere decir, que la filtración \mathcal{F}_s contiene toda la información de lo que puede suceder en el proceso hasta el tiempo s . Al conjunto $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P)$ se le conoce como espacio de probabilidad filtrado.

Dado un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in T}, P)$ se define la variable aleatoria *tiempo de paro* con respecto a \mathcal{F} , denotada por τ , como cualquier función que toma valores en T y cumple con: (i) $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$, y (ii) $P(\tau(\omega) < \infty) = 1$. El nombre de tiempo de paro se debe a la propiedad (i) $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$, que quiere decir, para saber si la variable es menor que t solamente se requiere la información anterior al tiempo t , en otras palabras, la variable aleatoria para la información del proceso al tiempo t .

El problema analizado en esta tesis, encontrar el dominio de paro óptimo, fue tratado en el caso donde el proceso que describe al precio del activo subyacente es un *proceso de Markov*. Un proceso estocástico es de Markov si tiene la característica de que la información sobre la evolución del proceso está resumida en el último tiempo observado. En este caso la probabilidad de transición del tiempo n al $n+1$ depende sólo del estado X_n y no de toda la historia de los estados anteriores, a esta característica se le conoce como la propiedad de Markov.

Definición 2.3.1 *Un proceso estocástico se dice que es de Markov si*

$$P(X_t | X_{t_k}, X_{t_{k-1}}, \dots, X_{t_0}) = P(X_t | X_{t_k}), \text{ para toda } k \in \mathbb{N} \text{ y } t_k > t_{k-1} > \dots > t_0 \in T$$

Esta definición depende de la existencia de la medida de probabilidad $P(\cdot|\cdot)$, la cual está inducida por la medida de probabilidad $P(\cdot)$ del espacio de probabilidad (Ω, \mathcal{F}, P) , en el cual está definida la cadena de Markov. A la familia de medidas de probabilidad que son generadas al condicionar una variable X por otra variable Y se le conoce como kernel estocástico de X dado Y , en términos más formales

Definición 2.3.2 *El kernel estocástico de X en Y son las funciones $P(\cdot|\cdot)$ tales que*

1. $P(\cdot|y)$ es una medida de probabilidad definida en $(X, \sigma(X))$ para cada $y \in \sigma(Y)$ fija, y
2. $P(B|\cdot)$ es una función medible en $(Y, \sigma(Y))$ para cada $B \in \sigma(X)$.

Se dice que un proceso de Markov es homogéneo cuando es un proceso estocástico homogéneo como se describió anteriormente, pero usando la propiedad de Markov se puede definir como:

Definición 2.3.3 *Un proceso de Markov es homogéneo en el tiempo si*

$$P(X_t|X_s) = P(X_m|X_0) \text{ para toda } t > s, m \in T \text{ tales que } t - s = m.$$

Otro tipo de procesos estocásticos importantes son los conocidos como *martingalas*, *submartingalas* y *supermartingalas*. La característica que los distingue es que guardan una relación en esperanza con su pasado. Una martingala es un proceso estocástico, en el cual se espera que siempre se comporte de forma análoga al último pasado conocido, esto quiere decir que la esperanza del proceso dada la historia es igual al último valor observado, en otras palabras que si se desea hacer una estimación del futuro no hay mejor aproximación que el último valor conocido. Una supermartingala se comporta en sentido decreciente, es decir que el valor esperado del futuro dado el presente será menor al valor presente, y la submartingala tiene un sentido creciente, entonces la esperanza del futuro dado el último pasado conocido será mayor a él.

Definición 2.3.4 *El proceso estocástico $\{X_t\}_{t \in T}$, se dice que*

1. *es martingala si $E[X_t|X_s] = X_s$*
 2. *es submartingala si $E[X_t|X_s] > X_s$*
 3. *es supermartingala si $E[X_t|X_s] < X_s$*
- para toda $s < t \in T$.*

2.4 Procesos controlados de Markov

En un *problema de control óptimo* se tiene un *sistema dinámico*, el cual está regulado o influenciado en su vida por ciertas variables llamadas de *control*, *reguladoras* o *acciones*. Las variables de control pueden ser aplicadas en cualquier momento de la vida del sistema y son elegidas de acuerdo a reglas definidas previamente, estas reglas son conocidas como *políticas de control*, al aplicar un control al sistema se tiene un costo. Para evaluar la respuesta del sistema a las políticas de control usadas debe existir una función llamada *índice de funcionamiento*, la cual está definida sobre el conjunto de las posibles políticas de control. El problema que se presenta aquí es como elegir una política de control, por lo general lo que se busca es obtener la política de control que maximice o minimice el índice de funcionamiento del sistema.

De acuerdo a la naturaleza del sistema, los problemas de control óptimo se clasifican en determinísticos o estocásticos, en tiempo discreto o continuo; una segunda clasificación se hace de acuerdo a la duración del sistema, cuando el sistema tiene una vida limitada se habla de un problema con horizonte finito e infinito en el caso contrario.

El problema de describir el dominio de paro óptimo de una opción americana será tratado a tiempo discreto, por lo que se tiene un problema de control óptimo estocástico a tiempo discreto y con horizonte finito. Como el proceso que describe a la evolución del precio se supone cadena de Markov se tiene un *proceso controlado de Markov*; por un proceso controlado de Markov se entiende un problema de control óptimo donde el sistema dinámico está descrito por un proceso de Markov.

Definición 2.4.1 *Un proceso controlado de Markov es una quintupla $(X, A, \{A(x) | x \in X\}, P, c)$, donde*

- (a) *X es un espacio de Borel, llamado espacio de estados y está generado por los posibles estados del proceso;*
- (b) *A un espacio de Borel, y es el conjunto de todos los controles o acciones permitidas;*
- (c) *$\{A(x) | x \in X\}$ una familia de subconjuntos medibles $A(x)$ de A , donde $A(x)$ denota al conjunto de todos los controles o acciones factibles del sistema cuando este encuentra en el estado $x \in X$, y con la propiedad de que el conjunto de parejas estado-acciones factibles*

$$\mathbb{K} := \{(x, a) | x \in X, a \in A(x)\},$$

es un subconjunto medible de $X \times A$;

- (d) *P es el kernel estocástico de X dado \mathbb{K} , la ley de transición del sistema; y*
- (e) *$c : \mathbb{K} \rightarrow \mathbb{R}$ una función medible, llamada la función de costo por estado.*

En lugar de considerar una función de costo por estado en algunos problemas es necesario considerarla como una función de recompensa por estado $r : \mathbb{K} \rightarrow \mathbb{R}$.

El proceso controlado de Markov que se acaba de definir es un sistema dinámico estocástico, de aquí en adelante solamente se considera el caso discreto, esto quiere decir que sólo es observado en los tiempos $t = 0, 1, \dots$. Al tiempo t el sistema se encuentra en el estado $x_t \in X$ y se debe aplicar un control $a_t \in A(x_t)$, por lo que ocurren dos cosas en el sistema: (i) el sistema genera un costo $c(x_t, a_t)$, y (ii) el sistema

se moverá al siguiente estado $x_{t+1} \in X$ de acuerdo a la ley de transición de X dado (x_t, a_t) , es decir,

$$P(B|x_t, a_t) := \text{Prob}(x_{t+1} \in B|x_t, a_t), B \subset X,$$

una vez que el sistema se encuentre en el estado $x_{t+1} \in X$ se aplicará un nuevo control $a_{t+1} \in A(x_{t+1})$, y en ese momento el proceso volverá a repetirse. En los puntos (i) y (ii) se ve la relación que existe con los procesos de Markov, por ello son conocidos como procesos controlados de Markov, ya que tanto el costo en que incurre al aplicar una acción y la ley de transición sólo depende de la última información del sistema.

Las políticas de control pueden ser aleatorias o deterministas, para formalizar la idea de una política de control se debe primero hablar de la historia del proceso controlado de Markov.

Para un proceso controlado de Markov el espacio de historias admisibles al tiempo t , denotado por H_t , se define como:

$$\begin{aligned} H_0 & : = X, \text{ y} \\ H_t & : = \mathbb{K}^t \times X = \mathbb{K} \times H_{t-1} \text{ para cada } t = 1, 2, \dots, \end{aligned}$$

donde a cada elemento $h_t \in H_t$ se le conoce como una t -historia admisible y es un vector de forma $h_t = (x_0, a_0, \dots, x_{t-1}, a_{t-1}, x_t)$, donde $(x_i, a_i) \in \mathbb{K}$ para toda $i = 0, 1, \dots, t-1$ y $x_t \in X$.

Definición 2.4.2 *Una política de control aleatoria es una serie $\pi = \{\pi_t, t = 0, 1, \dots\}$ de kernels estocásticos π_t del conjunto de controles A dado que H_t satisface la condición*

$$\pi_t(A(x_t)|h_t) = 1 \text{ para toda } h_t \in H_t, t = 0, 1, \dots$$

La definición de una política de control aleatoria se ha dado en términos de una serie de kernels estocásticos, pero esta serie también se puede interpretar como una serie de variables aleatorias que toman valores en A , llamadas controles o acciones, tales que para toda historia $h_t, t = 0, 1, \dots$, la distribución de a_t está dada por $\pi(\cdot|h_t)$, que por la condición de la política de control la distribución debe estar concentrada en $A(x_t)$ el conjunto de acciones factibles cuando el proceso se encuentra en el estado x_t .

2.5 Problema: describir la estructura de los dominios de paro óptimo de una opción americana de compra

Como se mencionó en la introducción, en esta tesis se analizará el problema que se le presenta al poseedor de opciones americanas de compra, es decir, maximizar sus

ganancias al elegir el tiempo de ejercicio óptimo, aunque el comprador puede elegir cualquier tiempo en el intervalo de ejercicio $[0, N]$ el problema sólo será analizado a tiempo discreto, esto quiere decir que sus únicos posibles tiempos de ejercicio son $\{0, 1, \dots, N\}$.

Una opción americana de compra tienen una función de pago convexa $g(S_t) = [S_t - K]^+$ para cada tiempo $t \in [0, N]$. La información de tener una función de pago convexa será combinada con el teorema de paro óptimo de Shiryaev (1978) para dar la descripción de los dominios de paro óptimo de dicha opción de compra. Después de analizar este caso se extenderá la descripción de los dominios de paro óptimo a problemas de optimización con función de pago convexa. La extensión del análisis se hará a través de aproximar la función de pago por medio funciones convexas y lineales por trozos.

Para poder utilizar el teorema de paro óptimo de Shiryaev el proceso que describe el comportamiento del activo subyacente debe ser de Markov, denótese por $\{S_n\}_{n \in \mathbb{N}}$ el proceso del precio del activo. Entonces los posibles flujos del poseedor de opciones americanas de compra son $\{g_0, g_1, \dots, g_N\}$, donde $g_n = g_n(S_n)$ para toda $n \in \{0, 1, \dots, N\}$.

Si el comprador elige el tiempo $\tau \in \{0, 1, \dots, N\}$ como tiempo de ejercicio la opción termina y obtiene un único flujo igual a g_τ ; la variable τ es un tiempo de paro definido en la filtración natural del proceso $\{S_n\}_{n \in \mathbb{N}}$.

Entonces el comprador está buscando una política de control $\Pi = \{\pi_n\}_{n \leq N}$ que maximice sus ganancias, para aclarar la idea de esta política de control es conveniente describir su único flujo en función del tiempo de paro y todos los posibles flujos,

$$\begin{aligned} g_\tau &= g_0 + (g_1 - g_0) + \dots + (g_{\tau-1} - g_\tau) = \\ g_0 + \sum_{k=1}^{\tau} \Delta g_k &= g_0 + \sum_{k=1}^N I\{k \leq \tau\} \Delta g_k = \\ &= g_0 + \sum_{k=1}^N \pi_k \Delta g_k \end{aligned}$$

donde $\Delta g_k = g_k - g_{k-1}$, $k \in \{1, 2, \dots, N\}$.

Entonces si el comprador decide elegir al tiempo $n \in \{0, 1, \dots, N\}$ como tiempo de ejercicio es equivalente a decir que ha ejercido la política de control $\Pi = \{\pi_t\}_{n \leq N}$, donde $\pi_t = 0$ si $t \in \{0, 1, \dots, n-1, n+1, \dots, N\}$ y $\pi_n = 1$.

El valor de cada flujo g_n es aleatorio, por esto la elección del tiempo de paro que maximice las ganancias, es decir, la elección del tiempo de paro óptimo, denotado por τ_{opt} , se hará a través de buscar el máximo de los valores esperados de los flujos en valor presente, lo cual permitirá maximizar las ganancias esperadas, lo que es equivalente a

maximizar la función

$$\phi_g(\tau) = E[e^{-R\tau} g_\tau(S_\tau)],$$

sobre los tiempos de paro $\tau \leq N$, donde R_τ es la tasa libre de riesgo en el intervalo $[0, \tau]$. Al maximizar se está buscando el tiempo de paro óptimo τ_{opt} que por definición debe cumplir con

$$\phi(\tau_{opt}) = \sup_{\tau \leq N} \phi(\tau). \quad (2.5.1)$$

De analizar este problema se obtendrá una sucesión de subconjuntos $\{\Gamma_n\}_{n \leq N}$ de \mathbb{R} , llamados dominios de paro óptimo, y con ellos se podrá dar respuesta a la pregunta ¿el tiempo paro n es óptimo?, la respuesta será afirmativa si el precio del activo subyacente S_n se encuentra dentro de su respectivo dominio de paro óptimo Γ_n , en caso contrario el tiempo n no será óptimo.

Capítulo 3

Teorema de paro óptimo, Shiryaev (1978)

3.1 Introducción

Dentro de la teoría de optimización estocástica se encuentra el teorema de paro óptimo desarrollado por Shiryaev en 1978 [6]. El tiempo de paro óptimo que se obtiene por medio de aplicar este resultado asegura que se obtendrá el supremo del valor esperado del proceso estocástico a optimizar, el cual debe ser discreto y con horizonte finito.

En el desarrollo de dicho teorema se usará un nuevo proceso $\{\gamma_n\}_{0 \leq n \leq N}$, este es definido por una función recursiva del proceso estocástico a optimizar, el cual está denotado por $\{f_n\}_{0 \leq n \leq N}$. Al finalizar la demostración del teorema se verá que se puede trabajar con el proceso $\{\gamma_n\}_{0 \leq n \leq N}$ para obtener un conjunto de valores $X \subset \Omega$ del dominio de $\{f_n\}_{0 \leq n \leq N}$, con los que se puede decidir si el tiempo $\tau \in \mathbb{N}$ es óptimo o no, ya que si $f_\tau \in X$ se tiene que τ es un tiempo de paro óptimo.

En primer lugar se verá la forma general del teorema, y después se verá como cambia este resultado cuando se está optimizando un proceso de Markov.

3.2 Teorema de paro óptimo, Shiryaev (1978).

Para poder enunciar el teorema se debe suponer que se tiene un proceso estocástico $f = \{f_n\}_{0 \leq n \leq N}$ que se desea optimizar en el sentido que se describió en la introducción del capítulo. El proceso $f = \{f_n\}_{0 \leq n \leq N}$ debe estar definido en un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{0 \leq n \leq N}, \mathbb{P})$ y ser homogéneo, donde $(\mathcal{F}_n)_{0 \leq n \leq N}$ es la filtración natural de proceso, esto quiere decir que $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ y $\mathcal{F}_N = \mathcal{F}$ y cada f_n es medible respecto a \mathcal{F}_n . También es necesario suponer que el proceso es integrable, que es análogo a pedir que $E|f_n| < \infty$ para toda $n \leq N$.

El problema que se presenta es: ¿como elegir un momento $\tau \in \{0, 1, \dots, N\}$ tal que maximiza el valor esperado del proceso?. Para responder esta pregunta se resolverán los siguientes dos puntos:

- i. encontrar el valor de la función de rendimiento

$$V_n = \sup_{\tau \in \mathfrak{M}_n} E[f_\tau]$$

donde el supremo es tomado sobre la clase de tiempos de paro $\mathfrak{M}_n = \{\tau : n \leq \tau \leq N\}$, y

- ii. dar el tiempo de paro óptimo para cada función V_n .

Como se tiene un problema de horizonte finito y discreto es posible utilizar recursión hacia atrás, esta propiedad será clave para la solución del problema, ya que la recursión hacia atrás se vuelve la principal herramienta en la búsqueda de la solución de V_n , y su correspondiente tiempo de paro óptimo; como se observa en la definición del siguiente proceso.

Defínase al proceso $\gamma = \{\gamma_n\}_{0 \leq n \leq N}$ recursivamente por medio de

$$\begin{aligned} \gamma_N &= f_N, \\ \gamma_n &= \max(f_n, E[\gamma_{n+1} | \mathcal{F}_n]), \quad 0 \leq n < N, \end{aligned}$$

también se debe definir a la variable aleatoria τ_n , llamada tiempo de paro, por

$$\tau_n = \min \{n \leq i \leq N : f_i = \gamma_i\}, \quad 0 \leq n \leq N.$$

Con base en estas definiciones a continuación se enuncia el teorema de paro óptimo de Shiryaev.

Teorema 3.2.1 *El proceso estocástico $\gamma = \{\gamma_n\}_{0 \leq n \leq N}$ y el tiempo de paro τ_n , $0 \leq n \leq N$, tienen las siguientes propiedades:*

- a) $\tau_n \in \mathfrak{M}_n$;
- b) $E[f_{\tau_n} | \mathcal{F}_n] = \gamma_n$;
- c) $E[f_\tau | \mathcal{F}_n] \leq E[f_{\tau_n} | \mathcal{F}_n] = \gamma_n$ para cada $\tau \in \mathfrak{M}_n$;
- d) $\gamma_n = \text{ess sup}_{\tau \in \mathfrak{M}_n} E[f_\tau | \mathcal{F}_n]$, en particular $\gamma_0 = \sup_{\tau \in \mathfrak{M}_0} E[f_\tau | f_0] = E[f_{\tau_0} | f_0]$;
- e) $V_n = E[\gamma_n]$.

Demostración. La propiedad a) claramente se cumple por definición de la clase de tiempos de paro $\mathfrak{M}_n = \{n \leq \tau \leq N\}$ y que $\tau_n = \min\{n \leq i \leq N : f_i = \gamma_i\} \neq \emptyset$, ya que $f_N = \gamma_N$.

Para la demostración de las propiedades b) y c) se procederá por de inducción hacia atrás.

i) Sea $n = N$, por lo que $\tau \in \mathfrak{M}_N = \{\tau : N \leq \tau \leq N\} = N$ y $\tau_N = \min\{N \leq i \leq N : f_i = \gamma_i\} = N$ entonces

$$\begin{aligned} E[f_\tau | \mathcal{F}_N] &= E[f_{\tau_N} | \mathcal{F}_N] = E[f_N | \mathcal{F}_N] = f_N = \gamma_N, \\ E[f_{\tau_N} | \mathcal{F}_N] &= \gamma_N, \end{aligned}$$

por lo que se cumplen las propiedades b y c cuando $n = N$.

ii) Suponer que para cada $n \in \{N, N-1, \dots, k\}$ se cumplen

$$\begin{aligned} \text{b) } E[f_{\tau_n} | \mathcal{F}_n] &= \gamma_n, \text{ y} \\ \text{c) } E[f_\tau | \mathcal{F}_n] &\leq E[f_{\tau_n} | \mathcal{F}_n] = \gamma_n \text{ para cada } \tau \in \mathfrak{M}_n \end{aligned}$$

iii) Demostrar para $n = k-1$. Sea $\tau \in \mathfrak{M}_{k-1} = \{k-1 \leq \tau \leq N\}$ y $A \in \mathcal{F}_{k-1}$, y defínase la variable $\bar{\tau} = \max(\tau, k)$, observe que $\bar{\tau} \in \mathfrak{M}_k = \{k \leq \tau \leq N\}$; por definición de las variables tiempo de paro se tiene que $\{\tau \leq k-1\} \in \mathcal{F}_{k-1}$, entonces $\{\tau \geq k\} \in \mathcal{F}_{k-1}$ por ser su complemento. Con esto en mente se puede ver

$$\begin{aligned} E[IAf_\tau] &= E[I_{A \cap \{\{\tau \geq k-1\} \cup \{\tau < k-1\}\}} f_\tau] \\ &= E[I_{A \cap \{\tau \geq k-1\}} f_\tau] + E[I_{A \cap \{\tau < k-1\}} f_\tau] \end{aligned}$$

como $\tau \in \mathfrak{M}_{k-1}$, se tiene que $I_{A \cap \{\tau < k-1\}} f_\tau = 0$

$$\begin{aligned} E[IAf_\tau] &= E[I_{A \cap \{\tau \geq k-1\}} f_\tau] \\ &= E[I_{A \cap \{\tau = k-1\}} f_\tau] + E[I_{A \cap \{\tau \geq k\}} f_\tau] \end{aligned}$$

en el segundo sumando del lado derecho $\tau \geq k$ y $\bar{\tau} = \max\{k, \tau\} = \tau$, además como $\mathcal{F}_{k-1} \subset \mathcal{F}_k$ se tiene que $E[f_\tau | \mathcal{F}_{k-1}] = f_\tau$ y $E\{E(f_{\bar{\tau}} | \mathcal{F}_k) | \mathcal{F}_{k-1}\} = E[f_\tau | \mathcal{F}_{k-1}]$; en el primer sumando debido a la intersección de A con $\{\tau = k-1\}$ se tiene que $f_\tau = f_{k-1}$. Entonces se sigue

$$\begin{aligned} &= E[I_{A \cap \{\tau = k-1\}} f_{k-1}] + E[I_{A \cap \{\tau \geq k\}} E[f_\tau | \mathcal{F}_{k-1}]] \quad (3.2.1) \\ &= E[I_{A \cap \{\tau = k-1\}} f_{k-1}] + E[I_{A \cap \{\tau \geq k\}} E\{E(f_{\bar{\tau}} | \mathcal{F}_k) | \mathcal{F}_{k-1}\}] \end{aligned}$$

por hipótesis de inducción se sabe que $E[f_\tau | \mathcal{F}_k] \leq E[f_{\tau_k} | \mathcal{F}_k] = \gamma_k$ para cada $\tau \in \mathfrak{M}_n$, en particular para $\bar{\tau}$

$$\leq E[I_{A \cap \{\tau = k-1\}} f_{k-1}] + E[I_{A \cap \{\tau \geq k\}} E\{\gamma_k | \mathcal{F}_{k-1}\}]$$

por definición $\gamma_{k-1} = \max(f_{k-1}, E[\gamma_k | \mathcal{F}_{k-1}])$, entonces

$$\leq E[I_A \gamma_{k-1}]$$

Como esto se vale para toda $A \in \mathcal{F}_{k-1}$, entonces

$$E[f_\tau | \mathcal{F}_{k-1}] \leq \gamma_{k-1}, \text{ para toda } \tau \in \mathfrak{M}_{k-1}$$

con esto queda demostrada la propiedad c).

Ahora se demostrara la propiedad b), las igualdades anteriores fueron construidas para $\tau \in \mathfrak{M}_{k-1}$, en particular $\tau_{k-1} \in \mathfrak{M}_{k-1}$, por lo que se pueden reescribir las igualdades 3.2.1 de la siguiente manera

$$\begin{aligned} E[I_A f_{\tau_{k-1}}] &= E[I_{A \cap \{\tau_{k-1}=k-1\}} f_{k-1}] + E[I_{A \cap \{\tau_{k-1} \geq k\}} f_{\tau_{k-1}}] \\ &= E[I_{A \cap \{\tau_{k-1}=k-1\}} f_{k-1}] + E[I_{A \cap \{\tau_{k-1} \geq k\}} E\{E(f_{\tau_{k-1}} | \mathcal{F}_k) | \mathcal{F}_{k-1}\}] \end{aligned}$$

si $\tau_{k-1} = k - 1$ se tiene que $\gamma_{k-1} = f_{k-1}$, y si $\tau \geq k$ sucede que $E(f_{\tau_{k-1}} | \mathcal{F}_k) = \gamma_k$; entonces

$$\begin{aligned} &= E[I_{A \cap \{\tau_{k-1}=k-1\}} \gamma_{k-1}] + E[I_{A \cap \{\tau_{k-1} \geq k\}} E\{\gamma_k | \mathcal{F}_{k-1}\}] \\ &= E[I_{A \cap \{\tau_{k-1}=k-1\}} \gamma_{k-1} + I_{A \cap \{\tau_{k-1} \geq k\}} E\{\gamma_k | \mathcal{F}_{k-1}\}] \\ &= E[I_A \gamma_{k-1}] \end{aligned}$$

por lo tanto

$$E[f_{\tau_{k-1}} | \mathcal{F}_{k-1}] = \gamma_{k-1}$$

con esto se concluye la demostración de las propiedades b) y c).

Por la propiedad c se tiene que

$$E[f_\tau | \mathcal{F}_n] \leq E[f_{\tau_n} | \mathcal{F}_n] = \gamma_n \text{ para cada } \tau \in \mathfrak{M}_n$$

entonces, la propiedad d) es consecuencia de c)

$$\gamma_n = \text{ess sup}_{\tau \in \mathfrak{M}_n} E[f_\tau | \mathcal{F}_n], \text{ y en particular } \gamma_0 = \sup_{\tau \in \mathfrak{M}_0} E[f_\tau] = E[f_{\tau_0}].$$

Para concluir la demostración sólo falta ver que la propiedad e) es cierta, pero también es consecuencia de la propiedad c), ya que

$$\begin{aligned} E[f_\tau | \mathcal{F}_n] &\leq E[f_{\tau_n} | \mathcal{F}_n] = \gamma_n \text{ para cada } \tau \in \mathfrak{M}_n \\ E[E[f_\tau | \mathcal{F}_n]] &\leq E[E[f_{\tau_n} | \mathcal{F}_n]] = E[\gamma_n] \text{ para cada } \tau \in \mathfrak{M}_n \\ E[f_\tau] &\leq E[f_{\tau_n}] = E[\gamma_n] \text{ para cada } \tau \in \mathfrak{M}_n \\ V_n &= \sup_{\tau \in \mathfrak{M}_n} E[f_\tau] = E[\gamma_n]. \end{aligned}$$

■

En este teorema se enuncian las propiedades del proceso $\gamma = \{\gamma_n\}_{0 \leq n \leq N}$, así como de la variable tiempo paro τ_n , $n \in \mathbb{N}$. Aunque no se ha nombrado como tal al tiempo de paro óptimo del proceso $f = \{f_n\}_{0 \leq n \leq N}$, este ya ha quedado implícito en estas propiedades, para dejarlo claro se enuncia el siguiente Corolario.

Corolario 3.2.2 *La variable*

$$\tau_0 = \min \{0 \leq i \leq N : f_i = \gamma_i\}$$

es el tiempo de paro óptimo dentro de la clase \mathfrak{M}_0 , esto quiere decir que

$$\sup_{\tau \in \mathfrak{M}_0} E[f_\tau] = E[f_{\tau_0}] (= \gamma_0).$$

Demostración. Es un caso particular de la propiedad c), cuando $n = 0$. ■

Como se observa en este Corolario, para encontrar el tiempo de paro óptimo al proceso $f = \{f_n\}_{0 \leq n \leq N}$, este debe ser comparado con el proceso $\gamma = \{\gamma_n\}_{n \leq N}$. Para la elaboración del proceso $\gamma_N, \gamma_{N-1}, \dots, \gamma_0$ es importante que se haya fijado $\gamma_N = f_N$, ya que esto permite inicializar la función recursiva $\gamma_n = \max(f_n, E[\gamma_{n+1} | \mathcal{F}_n])$.

Para concluir con el estudio del teorema de paro óptimo de Shiryaev 3.2.1 se enuncian a continuación un corolario con dos propiedades más del proceso $\gamma = \{\gamma_n\}_{n \leq N}$.

Corolario 3.2.3 *El proceso $\gamma = \{\gamma_n\}_{n \leq N}$ es una supermartingala. Además, γ es la menor supermartingala mayor que el proceso $f = \{f_n\}_{n \leq N}$, en el siguiente sentido, si $\tilde{\gamma} = \{\tilde{\gamma}_n\}_{n \leq N}$ es una supermartingala y $\tilde{\gamma}_n \geq f_n$, para toda $n \leq N$, entonces $\gamma_n \leq \tilde{\gamma}_n$ para toda $n \leq N$.*

Demostración. Por definición de $\gamma = \{\gamma_n\}_{n \leq N}$ se sabe

$$\gamma_n = \max\{f_n, E[\gamma_{n+1} | \mathcal{F}_n]\} \geq E[\gamma_{n+1} | \mathcal{F}_n] \text{ para toda } n \leq N$$

por lo que es supermartingala.

Por medio de inducción hacia atrás se demostrará que es la menor supermartingala

i) Sea $n = N$, entonces $\tilde{\gamma}_N \geq f_N = \gamma_N$; por lo tanto $\tilde{\gamma}_N \geq \gamma_N$

ii) Supóngase que para cada $n \in \{N, N-1, \dots, k\}$ se cumple que $\tilde{\gamma}_n \geq \gamma_n$.

ii) Si $n = k-1$, se tiene que

$$\tilde{\gamma}_{k-1} \geq E[\tilde{\gamma}_k | \mathcal{F}_k] \quad \text{y} \quad \tilde{\gamma}_{k-1} \geq f_{k-1}$$

entonces, $\tilde{\gamma}_{k-1}$ es mayor que el máximo de ambos

$$\tilde{\gamma}_{k-1} \geq \max(f_{k-1}, E[\tilde{\gamma}_k | \mathcal{F}_k]),$$

y como $\tilde{\gamma}_k \geq \gamma_k$, se concluye

$$\tilde{\gamma}_{k-1} \geq \max(f_{k-1}, E[\gamma_k | \mathcal{F}_k]) = \gamma_{k-1}.$$

■

Con esto se concluye el análisis de las propiedades de la variable de tiempo de paro τ_n , $n \in \mathbb{N}$, y del proceso de rendimiento $\gamma = (\gamma_n)_{n=0}^N$ en el modelo general, es decir, cuando se los flujos son modelados por un proceso estocástico sin ninguna restricción o característica especial, a excepción de que este sea integrable.

Cuando el tiempo de paro óptimo se busca en flujos que son modelados por un proceso de Markov los resultados anteriores pueden ser modificados de forma que simplifiquen los resultados obtenidos, esto se debe a la propiedad Markoviana de sólo tener memoria del pasado inmediato. Aunque en capítulos posteriores no se hará uso de estos cambios, se presenta a continuación la variación del teorema de paro óptimo cuando se optimiza un proceso de Markov.

Para enunciar esta variación, también, se debe suponer que se trabaja en un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, P)$, en el cual está definido el proceso de Markov $(X_n)_{n \in \mathbb{N}}$ el cual se supone homogéneo y discreto.

Para simplificar la notación definamos a T como el operador esperanza de transición en un paso, esto quiere decir, $Tf(x_1) = E_{x_0}[f(x_1)] = E[f(x_1)|x_0]$, para una función medible $f(x)$ tal que $E[|f(x)|] < \infty$, $x \in E$.

En lugar de optimizar al proceso $(f_n)_{n \leq N}$, se optimizará una transformación de este, la transformación está inducida por una función $g(x)$ aplicada al proceso $(X_n)_{n \in \mathbb{N}}$. Entonces para optimizar al proceso $(g(x_n))_{n \leq N}$, se debe suponer $E_x[\sup g^-(X_n)] < \infty$, $x \in E$.

La función V_N se ve transforma en

$$s(x) = \sup_{\tau \in \mathfrak{M}_0} E_x[g(x_\tau)].$$

Como el proceso es de Markov la función γ_n sólo depende del valor x_n ,

$$\begin{aligned} \gamma_n(x_n) &= \max\{g(x_n), E_{x_n}[g(x_{n+1}) | \mathcal{F}_n]\} \\ &= \max\{g(x_n), E_{x_n}[g(x_{n+1}) | x_n]\} \\ &= \max\{g(x_n), T(g(x_{n+1}))\}. \end{aligned}$$

A continuación se enuncia la versión Markoviana del teorema de paro óptimo de Shiryaev.

Teorema 3.2.4 *Sea $g = g(x)$ una función Borel-medible, tal que $E[g^-(x_k)] < \infty$ para cada $x_k \in E$ y $k \leq N$. Defínase la funcional*

$$s_N(x) = \sup_{\tau \in \mathfrak{M}_0} E_x[g(x_\tau)],$$

donde $\mathfrak{M}_0 = \{\tau : 0 \leq \tau \leq N\}$.

Sea

$$Qg(x) = \max\{g(x), Tg(x)\}$$

y

$$\tau_0 = \min\{0 \leq m \leq N : s_{N-m}(x_m) = g(x_m)\}.$$

Entonces

- a) $s_N(x) = Q^N g(x)$;
- b) $s_N(x) = \max(g(x), Ts_{N-1}(x))$, donde $s_0(x) = g(x)$;
- c) el tiempo de Markov τ_0 es óptimo en la clase \mathfrak{M}_0 , $E_x[g(x_{\tau_0})] = s_N$, $x \in E$;
- d) la serie $\gamma = (\gamma_m, \mathcal{F}_m)_{m \leq N}$ con $\gamma_m = s_{N-m}(x_m)$ es una supermartingala para cada $N \geq 0$.

Demostración. Este teorema es un caso particular del teorema 3.2.1; para ver la implicación se debe estudiar primero la estructura del operador $Qg(x)$, así como de las iteraciones $Q^n g(x)$, $n \in \mathbb{N}$.

Por definición

$$Q^n g(x) = Q(Q^{n-1}g(x)) = \max\{Q^{n-1}g(x), TQ^{n-1}g(x)\}, n = 1, \dots, N,$$

donde $Q^0 g(x) = g(x)$. Se demostrará por inducción

$$Q^n g(x) = \max\{g(x), TQ^{n-1}g(x)\}, n = 1, \dots, N.$$

i) Para $n = 1$, esto es valido por definición

$$Qg(x) = \max\{g(x), Tg(x)\}.$$

ii) Supóngase que para $n = k$ se cumple

$$Q^k g(x) = \max\{g(x), TQ^{k-1}g(x)\}.$$

iii) Demostrar para $n = k + 1$ se cumple. En primer lugar se tiene que

$$\begin{aligned} Q^{k+1}g(x) &= \max\{Q^k g(x), TQ^k g(x)\} \\ &= \max\left\{\max[g(x), TQ^{k-1}g(x)], T \max[g(x), TQ^{k-1}g(x)]\right\}, \end{aligned}$$

por contención

$$TQ^{k-1}g(x) \leq \max[g(x), TQ^{k-1}g(x)],$$

entonces, al aplicar el operador T , lo cual es integrar y preserva las igualdades se tiene

$$TQ^{k-1}g(x) = T \left[TQ^{k-1}g(x) \right] \leq T \max \left[g(x), TQ^{k-1}g(x) \right]$$

por lo tanto

$$\begin{aligned} Q^{k+1}g(x) &= \max \left\{ g(x), T \max \left[g(x), TQ^{k-1}g(x) \right] \right\} \\ &= \max \left\{ g(x), TQ^k g(x) \right\}. \end{aligned}$$

Con esto queda demostrado que

$$Q^n g(x) = \max \left\{ g(x), TQ^{n-1}g(x) \right\}, \quad n = 1, \dots, N.$$

Por la propiedad (d) del teorema 3.2.1 se tiene que

$$s_N(x) = \sup_{\tau \in \mathfrak{M}_0} E_x [g(x_\tau)] = \gamma_0,$$

pero en este caso

$$\begin{aligned} \gamma_N &= g(x), \\ \gamma_{N-1} &= \max \{g(x), Tg(x)\} = Qg(x), \\ &\dots \\ \gamma_0 &= \max \{g(x), T\gamma_1\} = \max \{g(x), TQ^{N-1}g(x)\} = Q^N g(x), \end{aligned}$$

por lo que se concluye que la propiedad a) es valida, esto quiere decir,

$$s_N(x) = Q^N g(x).$$

Por el mismo argumento se observa que para toda $n = 1, \dots, N$

$$s_n(x) = Q^n g(x).$$

La propiedad b) es valida, puesto que

$$\begin{aligned} s_N(x) &= Q^N g(x) = \max \{g(x), TQ^{N-1}g(x)\} \\ &= \max \{g(x), Ts_{N-1}\} \end{aligned}$$

El hecho de que el tiempo de Markov τ_0 sea óptimo en la clase \mathfrak{M}_0 , se debe a que τ_0 es tal que

$$s_N(x) = \sup_{\tau \in \mathfrak{M}_0} E_x [g(x_\tau)] = E_x [g(x_{\tau_0})]$$

de acuerdo a lo demostrado en el teorema 3.2.1 propiedad (d).

Por último, la propiedad d) es la conclusión a la que se ha llegado en el Corolario 3.2.3. ■

Para dejar más clara la utilización de estos teoremas, así como su alcance en la solución de los problemas de optimización a continuación se presenta un ejemplo.

Ejemplo 3.2.5 *Un hombre entra a un juego de apuestas, donde se llevarán acabo N juegos y el jugador se llevará una bolsa que depende de cuantos juegos haya ganado. En la bolsa por cada juego que gane se depositará una cantidad M y por cada juego perdido se retira la misma cantidad M de la bolsa. El jugador puede elegir retirarse después de realizarse cada uno de los N juegos.*

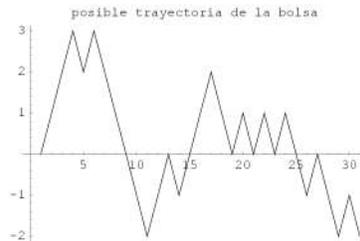
El jugador desea saber el tiempo de paro óptimo que maximizara sus ganancias suponiendo que tiene la misma probabilidad de ganar que perder cada juego.

Sean $\{\varepsilon_i\}_{i \in \mathbb{N}}$ un conjunto de variables aleatorias Bernoulli tales que

$$P(\varepsilon_i = M) = P(\varepsilon_i = -M) = \frac{1}{2},$$

y considere al proceso estocástico $\{X_n\}_{n \in \mathbb{N}}$, construido como $X_n = x + \varepsilon_1 + \dots + \varepsilon_n$ para cada $n \in \mathbb{N}$ y $x \in \{0, \pm 1, \pm 2, \dots\}$.

El proceso $\{X_n\}_{n \in \mathbb{N}}$ describe el movimiento de la bolsa de premio del jugador, y debe llevar una trayectoria similar a la siguiente gráfica.



Entonces, el problema que se le presenta al jugador es encontrar el tiempo de paro óptimo τ_{opt} tal que maximice el valor esperado de la bolsa, que es equivalente a resolver la función

$$s_N(x) = \sup_{\tau \in \mathfrak{M}_0} E_x [\beta^\tau X_\tau]$$

donde $\beta \in (0, 1]$ y representa el factor de valor presente libre de riesgo que debería ser considerado si el tiempo entre juegos es largo.

Para encontrar el tiempo de paro hay que dividir el problema en dos casos con respecto a los posibles valores de β .

Caso (1) *Cuando $\beta = 1$ (se desprecia el valor del dinero en el tiempo). Se debe resolver la ecuación $s_N(x) = \sup_{\tau \in \mathfrak{M}_0} E_x [X_\tau]$, donde se tiene que $g(X) = X$, por lo que hay que observar que pasa con $Q^N g(x)$, para esto observe que*

$$\begin{aligned} Qg(x) &= \max \{g(x), E_x (g(X_1))\} \\ &= \max \{x, E_x (X_1)\} \\ &= \max \{x, x + E_x (\varepsilon_1)\} \\ &= \max \{x, x\} = x = g(x), \end{aligned}$$

entonces $Q^N g(x) = g(x) = x = s_N(x)$, por lo tanto el tiempo de paro óptimo $\tau_{opt} = 0$ para toda $x \in \{0, \pm 1, \pm 2, \dots\}$.

Caso (2) Si $\beta \in (0, 1)$ (se considera el valor del dinero en el tiempo). Se tiene que resolver la ecuación $s_N(x) = \sup_{\tau \in \mathfrak{M}_0} E_x[\beta^\tau X_\tau]$, donde $g(x_n) = \beta^n x_n$, entonces

$$\begin{aligned} Qg(x) &= \max \{g(x), E_x(g(X_1))\} \\ &= \max \{x, E_x(\beta X_1)\} \\ &= \max \{x, \beta x + E_x(\varepsilon_1)\} \\ &= \max \{x, \beta x\} \\ &= \left\{ \begin{array}{l} x \text{ si } x \in \{0, 1, 2, \dots\} \\ \beta x \text{ si } x \in \{-1, -2, \dots\} \end{array} \right\} \end{aligned}$$

por lo tanto

$$\begin{aligned} s_N(x) &= Q^N g(x) \\ &= \left\{ \begin{array}{l} x \text{ si } x \in \{0, 1, 2, \dots\} \\ \beta^N x \text{ si } x \in \{-1, -2, \dots\} \end{array} \right\} \end{aligned}$$

entonces el tiempo de paro óptimo cuando $\beta \in (0, 1)$ es $\tau_{opt} = \min\{0 \leq n \leq N : x_n \in \{0, 1, 2, \dots\}\}$, si x_n nunca toma un valor del conjunto $\{0, 1, 2, \dots\}$ se tiene que el tiempo de paro óptimo es N , $\tau_{opt} = N$.

En conclusión si el jugador desprecia el valor del dinero en el tiempo puede jugar todos los partidos o decidir retirarse en cualquier momento puesto que todo tiempo es óptimo; por otro lado, si al jugador le interesa el valor del dinero en el tiempo y la bolsa inicial es cero debe retirarse en ese momento puesto que ya es óptimo, si la bolsa inicia con un valor negativo, que puede ser el costo de jugar, debe retirarse hasta que por lo menos la bolsa tenga valor cero.

Hasta el momento se ha trabajado con modelos tales que tienen un proceso estocástico homogéneo en el tiempo, pero en el siguiente capítulo se abordará el problema cuando no es necesariamente homogéneo en el tiempo; por esto es necesario saber que si se tiene un proceso no homogéneo en el tiempo es simple crear uno homogéneo a partir de él.

Proposición 3.2.6 Sea $\{X_n\}_{n \in \mathbb{N}}$ un proceso de Markov no homogéneo, donde cada variable aleatoria X_n , $n \in \mathbb{N}$, está definida en el espacio de probabilidad $(\Omega_n, \mathcal{F}_n, P_n(\cdot|\cdot))$, donde la función de transición entre los estados n y $n+1$ está dada por P_n . Entonces el proceso $Z_n = (X_n, n)$ es un proceso de Markov homogéneo en el tiempo.

Demostración. Es claro que el espacio de estados posibles del proceso $\{Z_n\}_{n \in \mathbb{N}}$ está dado por $\Omega = \cup_{n \in \mathbb{N}} (\Omega_n \times \{n\})$, en este espacio de estados existe \mathcal{F} σ -álgebra,

tal que hace medible a toda variable aleatoria Z_n , para toda $n \in \mathbb{N}$, de hecho $\mathcal{F} = \sigma(\cup_{n \in \mathbb{N}} \mathcal{F}_n)$.

Entonces se debe demostrar que

$$P[Z_{n+1} \in A | Z_n = (X_n, n)] = P[Z_1 \in A | Z_0 = (X_n, n)]$$

para toda $A \in \mathcal{F}$ para comprobar que el proceso $\{Z_n\}_{n \in \mathbb{N}}$ es homogéneo.

Sea $k \in \mathbb{N}$ fija y $A \in \mathcal{F}$, al calcular $P[Z_{k+1} \in A | Z_k = (X_n, n)]$ observe que si $Z_k = (X_n, n)$ por construcción $Z_{k+1} = (X_{n+1}, n+1)$, entonces A debe de ser de la forma $A = A_{n+1} \times \{n+1\}$, donde $A_{n+1} \in \Omega_{n+1}$, por lo tanto

$$\begin{aligned} P[Z_{k+1} \in A | Z_k = (X_n, n)] &= P[(X_{n+1}, n+1) \in A_{n+1} \times \{n+1\} | Z_k = (X_n, n)] \\ &= P[X_{n+1} \in A_{n+1} | X_n] = P_n(A | X_n), \end{aligned}$$

como esto es para cualquier $k \in \mathbb{N}$, se concluye que el proceso $\{Z_n\}_{n \in \mathbb{N}}$ es homogéneo en el tiempo. ■

Por medio de este resultado es claro que si se trabaja con un proceso de Markov el cual no sea homogéneo en el tiempo existe la transformación exhibida en la proposición anterior para transformarlo a homogéneo.

Con ayuda del teorema de paro óptimo de Shiryaev, así como de la proposición anterior, en el siguiente capítulo se expondrá como obtener dominios de paro óptimo cuando se tiene una opción americana.

Capítulo 4

Optimización de opciones Americanas, con función de pago convexa, a tiempo discreto

Como ya se ha mencionado, una opción es un instrumento financiero tal que su precio depende de conocer el valor del activo subyacente en la fecha de ejercicio. En el caso de las opciones americanas su precio está descrito por medio de las siguientes funciones:

1. $g(S_n) = [S_n - K]^+$ en el caso de una opción americana de compra, y
2. $g(S_n) = [K - S_n]^+$ en el caso de una opción americana de venta,

donde S_n es el precio del activo subyacente al tiempo n y K el precio pactado, estas funciones son siempre positivas y convexas.

Para estudiar una opción americana se debe suponer que se conoce al proceso estocástico $\{S_n\}_{n \in \mathbb{N}}$ que describe el valor del precio del activo subyacente, ya que con él se conoce al proceso estocástico del precio de la opción americana, pues este es la transformación que produce la función $g(x)$ en el proceso $\{S_n\}_{n \in \mathbb{N}}$.

En el capítulo anterior se mostró, como para un proceso estocástico $\{S_n\}_{n \in \mathbb{N}}$ se puede construir un proceso estocástico $\{\gamma_n\}_{n \in \mathbb{N}}$, con el cual es posible saber si el tiempo actual es tiempo de paro ya que $\tau_{n_opt} = \min\{n \leq i \leq N : S_i = \gamma_i\}$, τ_{n_opt} es tiempo de paro óptimo con respecto al valor esperado del proceso.

En este capítulo se analizará el proceso estocástico $\{\gamma_n\}_{n \in \mathbb{N}}$ con el fin de obtener subconjuntos de los posibles valores que toma la opción americana de compra, en los cuales se pueda asegurar que si la opción toma un valor dentro de ese subconjunto se está en un tiempo de paro óptimo. El optimizar una opción americana es equivalente a optimizar el proceso $\{g(S_n)\}_{n \in \mathbb{N}}$, ya que este es el proceso estocástico que describe

el valor de la opción. La obtención de los subconjuntos o dominios de paro óptimo se hará a través de reflejar la información que se tiene de la función $g(x)$ en el proceso estocástico $\{\gamma_n\}_{n \in \mathbb{N}}$. De hecho el análisis se puede extender a cualquier instrumento que depende de un activo subyacente y tenga una función de pago positiva, convexa y que en el cero valga cero.

En primer lugar se enunciarán las hipótesis del problema en general, para después hacer un estudio en tres etapas del problema. En la primera etapa se aborda el caso cuando la función de pago es la estándar de una opción americana de compra, la cual es casi lineal en todo su dominio a excepción de en un punto a lo más; después se prosigue a extender los resultados al problema donde la función de pago es convexa y lineal por trozos; para finalizar con los resultados de estas dos primeras etapas se analizará el problema de funciones convexas en general.

4.1 Modelo a tiempo discreto de paro óptimo

Como ya se ha mencionado se debe suponer que se conoce al proceso estocástico que describe al precio del activo subyacente, el cual será denotado por $\{S_n\}_{n \in \mathbb{N}}$, se incluirá la posibilidad de que exista un proceso estocástico índice $\{I_n\}_{n \in \mathbb{N}}$ que controle al proceso del precio. Estos dos procesos en conjunto deben formar un proceso $\{Z_n = (S_n, I_n)\}_{n \in \mathbb{N}}$, el cual debe ser de Markov, posiblemente no homogéneo en el tiempo.

El proceso $Z_n = (S_n, I_n)_{n \in \mathbb{N}}$ tiene como espacio de estados a $Z = [0, \infty) \times Y$, y está definido en el espacio medible (Z, \mathcal{B}_Z) , donde $\mathcal{B}_Z = \sigma(\mathcal{B}_+, \mathcal{B}_Y)$, \mathcal{B}_+ la σ -álgebra de Borel de $\mathbb{R}^+ \cup \{0\}$ y \mathcal{B}_Y la σ -álgebra generada por Y , el espacio de los posibles valores del proceso índice $\{I_n\}_{n \in \mathbb{N}}$.

Como se trabaja con un proceso que es de Markov se debe tener el kernel de transiciones del proceso. Se denotará por $P_{x,y}^n(A, B)$ a la probabilidad de transición de que bajo la condición $S_n = x, I_n = y$ suceda el evento $(S_{n+1} \in A, I_{n+1} \in B)$, donde $A \in \mathcal{B}_+$ y $B \in \mathcal{B}_Y$, indistintamente se escribirá $P_n(x, y; A, B) = P_{x,y}^n\{(S_{n+1}, I_{n+1}) \in A \times B\}$. Además suponga que el estado inicial del proceso es conocido y constante, $Z_0 = (S_0, I_0)$ constante.

De acuerdo al teorema de paro óptimo de Shiryaev se debe conocer la filtración del proceso, por lo que se usará la filtración natural del proceso $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$, esto quiere decir, $\mathcal{F}_n = \sigma(Z_0, \dots, Z_n)$ para toda $n \in \mathbb{N}$. Con respecto a esta filtración se tomará la variable aleatoria tiempo de paro τ , entonces τ debe cumplir con que $\{\omega : \tau(\omega) \leq n\} \in \mathcal{F}_n$ para toda $n \in \mathbb{N}$.

Con el fin de que se pueda extender el trabajo a instrumentos diferentes a las opciones americanas las funciones de pago que relacionan el precio del activo subyacente con el precio de la opción se considerarán no homogéneas en el tiempo, denotando por

$g_n(x)$ a la función de pago en el tiempo $n \in \mathbb{N}$; las funciones de pago $g_n(x)$, $n \in \mathbb{N}$, deben ser Borel medibles y no negativas. Debido a que se trabajará con pagos en el tiempo se deben conocer las tasas libres de riesgo r_m entre los tiempos m y $m+1$, las cuales se supondrán constantes, entonces $R_m = r_0 + \dots + r_m$ es la tasa libre de riesgo hasta el tiempo m , con $r_0 = 0$.

Con estas hipótesis hechas se tiene el problema de maximizar sobre los tiempos de paro, $\tau \leq N$, el valor esperado del valor presente de la opción; esto quiere decir, encontrar el tiempo de paro que maximice la funcional

$$\phi_g(\tau) = E[e^{-R_\tau} g_\tau(S_\tau)], \quad (4.1.1)$$

donde N es la fecha de vencimiento, por definición el tiempo de paro que se busca es el tiempo de paro óptimo, denotado por τ_{opt} , que cumple con la siguiente igualdad

$$\phi_g(\tau_{opt}) = \sup_{\tau \leq N} \phi_g(\tau). \quad (4.1.2)$$

Para definir el proceso equivalente a $\{\gamma_n\}_{n \in \mathbb{N}}$ se utilizará el operador T_n , este operador actúa sobre funciones no negativas y medibles $f(x, y) : Z \rightarrow [0, +\infty)$ de la siguiente manera

$$T_n f(x, y) = E_{x,y}^n [f(S_{n+1}, I_{n+1})],$$

donde $E_{x,y}^n$ es la esperanza condicional del proceso al tiempo n , dado que al tiempo $n-1$ se encontraba en el estado $(S_n = x, I_n = y)$.

Para $x \geq 0$, $y \in Y$ se defina el proceso $\{w_n\}_{0 \leq n \leq N}$ por medio de la siguiente recursión:

$$\begin{aligned} w_0(x, y) &= g_N(x), y \\ w_k(x, y) &= \max \{g_{N-k}(x), e^{-r_{N-k}} T_{N-k} w_{k-1}(x, y)\}, \text{ para } k = 1, 2, \dots, N. \end{aligned} \quad (4.1.3)$$

Note que este proceso está definido de forma equivalente al proceso $\{\gamma_n\}_{n \in \mathbb{N}}$ utilizado en el teorema de paro óptimo de Shiryaev 3.2.1; y los cambios se deben a que esta vez no se está maximizando el proceso $\{S_n\}_{n \in \mathbb{N}}$, sino al proceso transformado por $g_n(x)$ en valor presente al tiempo $k \in \mathbb{N}$.

A partir del proceso estocástico $\{w_n\}_{0 \leq n \leq N}$ se analizarán los dominios de paro óptimo. Recordemos que según el Teorema 3.2.1 el valor supremo de la funcional cumple la siguiente igualdad $V_n = \gamma_n$ y $\gamma_n = E[f_{\tau_n} | \mathcal{F}_n]$, donde $\tau_n = \min\{n \leq i \leq N : f_i = \gamma_i\}$. Entonces para $n \leq N$ fija, el dominio de paro óptimo para el tiempo n está dado por el conjunto

$$\Gamma_n = \{(x, y) : g_n(x) = w_{N-n}(x, y)\} \subseteq \mathbb{R}^+ \times Y, \quad (4.1.4)$$

que es análogo a decir que n es óptimo puesto que $g_n(x) = w_{N-n}(x, y)$. Si la $y \in Y$ es conocida o fija se define el conjunto cruzado

$$\Gamma_n[y] = \{x : g_n(x) = w_{N-n}(x, y)\} \subseteq \mathbb{R}^+, y \in Y. \quad (4.1.5)$$

sobre estos conjuntos se centrará el estudio de este capítulo con el fin de describirlos y analizar sus características para poder dar una forma más explícita de ellos. Es claro que si $n = N$ se tiene $\Gamma_N = \mathbb{R}^+ \times Y$ y $\Gamma_N[y] = \mathbb{R}^+$, con $y \in Y$, esto quiere decir, si ha llegado la fecha de expiración cualquier valor será óptimo, esto se debe a que en la fecha de vencimiento ya no existen opciones de ganar más.

De acuerdo al Teorema 3.2.1 la variable aleatoria tiempo de paro óptimo τ_{opt} cumple con la siguiente igualdad:

$$\begin{aligned} \tau_{opt} &= \min\{0 \leq n \leq N : g_n(x) = w_{N-n}(x, y)\} \\ &= \min\{0 \leq n \leq N : S_n \in \Gamma_n[I_n]\}, y \\ \phi_g(\tau_{opt}) &= w_N(S_0, I_0). \end{aligned}$$

Por esta característica es que a los conjuntos $\Gamma_n [I_n]$ se les da el nombre de dominios de paro óptimo.

El propósito de esta tesis es explicar la estructura de los dominios de paro óptimo $\Gamma_n [I_n]$, y así aprovechar su estructura en la solución del problema de optimización.

Una forma clásica de ver el problema es la forma dinámica, es decir, cuando el proceso del precio y el proceso índice están dados en forma dinámica por medio de

$$S_n = A_n(S_n, I_n), I_n = B_n(I_{n-1}, \xi_n), n \in \mathbb{N}, \quad (4.1.6)$$

donde $(\xi_n)_{n \in \mathbb{N}}$ es una sucesión de variables aleatorias independientes, que toman valores en un espacio medible $(E, \mathcal{B}(E))$, con $A_n : Z \rightarrow \mathbb{R}^+$ y $B_n : Y \times E \rightarrow Y$ funciones medibles con respecto a su respectiva σ -álgebra. El valor de $Z_0 = (S_0, I_0)$ constante.

Entonces el proceso $Z_n = (S_n, I_n)$, $n \in \mathbb{N}$ tiene asociada la filtración natural $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$, $n \in \mathbb{N}$.

Ejemplo 4.1.1 *Un caso particular de la forma dinámica (4.1.6) es la caminata aleatoria geométrica, en este caso se tiene*

$$S_n = S_{n-1}\xi_n, I_n = \xi_n, n \in \mathbb{N},$$

donde $(\xi_n)_{n \in \mathbb{N}}$ es una sucesión de variables aleatorias no negativas e independientes.

Antes de iniciar el estudio de la estructura de los dominios de paro óptimo de acuerdo a la convexidad de la función de pago de la opción es necesario hacer algunas observaciones con respecto a la dependencia de los dominios de paro del proceso índice, así como la unicidad del tiempo de paro.

1. Relación de los dominios de paro óptimo con respecto al proceso índice $\{I_n\}_{n \in \mathbb{N}}$.

- (a) Caso donde el dominio de paro no depende del proceso índice. Esto sucede cuando la probabilidad de transición para el proceso de Markov Z_n puede ser escrita como

$$P_n(x, y, A, B) = \int_B P_n(dz) P_n(x, z, A), \quad A \in \mathcal{B}_+, \quad B \in \mathcal{B}_Y. \quad (4.1.7)$$

Esta igualdad quiere decir que la componente índice I_n es una serie de variables aleatorias independientes con distribución $P_n(B)$, y es condicionalmente independiente de la primera componente S_n en el sentido de la ecuación (4.1.7). En este caso el proceso del precio S_n también es un proceso no homogéneo de Markov con probabilidad de transición

$$P_n(x, A) = \int_Y P_n(dz) P_n(x, z, A), \quad A \in \mathcal{B}_+.$$

Cuando el proceso está dado en forma dinámica (4.1.6), la condición (4.1.7) se traduce en que el proceso $I_n = B_n(\xi_n)$, es una serie de variables aleatorias independientes que son transformaciones de las variables aleatorias ξ_n . En este caso el modelo dinámico se ve como:

$$S_n = A_n(S_{n-1}, B_n(\xi_n)) = \tilde{A}(S_{n-1}, \xi_n), \quad n \in \mathbb{N}. \quad (4.1.8)$$

Ya sea la condición (4.1.7) o (4.1.8) ambas quieren decir que la función $\omega_n(x, y) = \omega(x)$ sólo depende de x , así que los conjuntos Γ_n y $\Gamma_n[y]$ no dependen de y (proceso índice I_n). Sea el conjunto $\Gamma_n^* = \{x \in \mathbb{R}^+ : g_n(x) = \omega_{N-n}(x)\}$, entonces se puede describir a los conjuntos $\Gamma_n, \Gamma_n[y]$ como $\Gamma_n = \Gamma_n^* \times Y$ y $\Gamma_n[y] = \Gamma_n^*$, como no existe dependencia hacia el proceso índice I_n el tiempo de paro óptimo queda definido como $\tau_{opt} = \min\{0 \leq i \leq N : S_n \in \Gamma_n^*\}$.

- (b) En este caso se analiza que sucede cuando el proceso índice toma valores en $Y = Y' \times Y''$, entonces la σ -álgebra de este proceso está dada por $\mathcal{B}_Y = \mathcal{B}_{Y'} \times \mathcal{B}_{Y''}$, y para cada $n \in \mathbb{N}$ el proceso índice es un vector aleatorio $I_n = (I'_n, I''_n)$, $n \in \mathbb{N}$. Si en este caso la probabilidad de transición del proceso de Markov Z_n no depende del valor de la componente índice I''_n , es decir, la probabilidad de transición puede ser escrita como

$$P_n(x, (y', y''), A, B) = \int_B P_n(y', dz) P_n(x, y', z, A), \quad A \in \mathcal{B}_+, \quad B \in \mathcal{B}_Y, \quad (4.1.9)$$

entonces la componente I'_n es un proceso no homogéneo de Markov con espacio de estados Y' y probabilidad de transición $P_n(y', B' \times Y'')$, $B' \in Y'$. También se tiene que el proceso (S_n, I'_n) es de Markov y no homogéneo con probabilidad de transición

$$P_n((x, y'), A, B') = \int_{B' \times Y''} P_n(y', dz) P_n(x, y', z, A), \quad A \in \mathcal{B}_+, \quad B \in \mathcal{B}_Y.$$

Cuando el proceso Z_n está dado en forma dinámica (4.1.6), la condición (4.1.9) quiere decir que el proceso índice debe estar dado como

$$I_n = B_n(I'_n, \xi_n), \quad n \in \mathbb{N}, \quad (4.1.10)$$

donde $(\xi_n)_{n \in \mathbb{N}}$ son variables aleatorias independientes entre ellas. Entonces la forma dinámica del precio está dada por

$$S_n = A_n(S_{n-1}, B_n(I'_n, \xi_n)), \quad n \in \mathbb{N}. \quad (4.1.11)$$

Ya sea la condición (4.1.9) o ((4.1.10) - (4.1.11)) ambas implican que

$$\omega(x, (y', y'')) = \omega(x, y'),$$

es decir, que sólo depende de (x, y') . Entonces $\Gamma[(y', y'')] = \Gamma'[y']$, $(y', y'') \in Y$, no depende de y'' , donde $\Gamma'_n[y'] = \{x \in \mathbb{R}^+ : g_n(x) = w_{N-n}(x, y')\}$. El tiempo de paro óptimo queda definido como $\tau_{opt} = \min\{0 \leq n \leq N : S_n \in \Gamma_n[I'_n]\}$.

2. La segunda observación es respecto a la unicidad del dominio de paro óptimo y su correspondiente tiempo de paro óptimo.

Consideremos al conjunto

$$C_n = \{x \geq 0 : g_n(x) = 0\}, \quad n \in \mathbb{N},$$

y para $n = 0, 1, \dots, N$ defínase los dominios de paro reducidos, como:

$$\bar{\Gamma}_n[y] = \Gamma_n[y] \setminus C_n = \{x \in \mathbb{R}^+ : g_n(x) = \omega(x, y) > 0\}, \quad y \in Y.$$

Si al tiempo $n < N$ la opción no ha sido ejercida y $S_n = x \in C_n$, entonces no tiene sentido ejercer la opción en ese momento ya que la ganancia es cero, pues $g_n(S_n) = 0$, y es conveniente esperar al siguiente momento; esto quiere decir que ejercer la opción en algún momento del intervalo $(n, N]$ no es peor a ejercerla en el tiempo n . Como se tiene la siguiente contención $\bar{\Gamma}_n[y] \subseteq \Gamma_n[y]$ de los dominios de paro, entonces

$$\bar{\tau}_{opt} = \min\{0 \leq n \leq N : S_n \in \bar{\Gamma}_n[I_n]\} \geq \tau_{opt},$$

como al dominio de paro $\Gamma_n[y]$ sólo se le han quitado los puntos en los cuales la ganancia es cero, se tiene que $\bar{\tau}_{opt}$ también maximiza la funcional (2.5.1), por lo tanto $\phi(\bar{\tau}_{opt}) = \phi(\tau_{opt})$.

Ahora tómesese una nueva colección de conjuntos $\tilde{C}_n \subseteq C_n, n \in \mathbb{N}$, y su respectivo dominio de paro $\tilde{\Gamma}_n[y] = \Gamma_n[y] \setminus \tilde{C}_n, n = 0, 1, \dots, N - 1, \tilde{\Gamma}_N = \mathbb{R}^+$; entonces se define el nuevo tiempo de paro óptimo como

$$\tilde{\tau}_{opt} = \min \left\{ 0 \leq n \leq N : S_n \in \tilde{\Gamma}_n[I_n] \right\}.$$

Entonces por construcción se tiene que $\bar{\Gamma}_n[y] \subseteq \tilde{\Gamma}_n[y] \subseteq \Gamma_n[y], y \in Y$, entonces $\bar{\tau}_{opt} \leq \tilde{\tau}_{opt} \leq \tau_{opt}$. Por la misma razón de que sólo se han eliminados puntos del dominio de paro óptimo donde la ganancia es cero, se concluye que

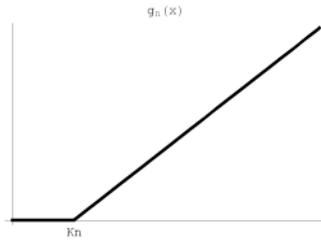
$$\phi(\bar{\tau}_{opt}) = \phi(\tilde{\tau}_{opt}) = \phi(\tau_{opt}).$$

Esto es lo que se entiende como la unicidad en el tiempo de paro óptimo, aunque no se puede asegurar que sean los mismo tiempos, si se asegura que el valor máximo de la funcional siempre alcanza el mismo valor.

Con las definiciones enunciadas hasta el momento y estas observaciones se puede proseguir con el estudio de la estructura de los dominios de paro. En primer lugar se analizará el caso donde la función de pago es lineal en el tiempo; partiendo de este primer análisis se procederá a extenderlo al caso donde la función de pago sea lineal por trozos, para concluir con el caso donde la función de pago es convexa en general.

4.2 Optimización en el caso de una función de pago lineal

Este análisis corresponde al caso donde se tiene a la función de pagos estándar de una opción americana, dicha función de pago debe ser de la siguiente forma



$$g_n(x) = a_n [x - K_n]_+ = \begin{cases} a_n (x - K_n), & \text{si } x \geq K_n \\ 0, & \text{si } 0 \leq x \leq K_n \end{cases}, \quad (4.2.12)$$

donde $a_n > 0$ y $K_n > 0, n \in \mathbb{N}$, son las escalas de los precios y los precios pactados respectivamente en cada momento n ; como la función de pago depende del tiempo es

un modelo no homogéneo, cuando esta no dependen del tiempo se tiene la función de pago estándar de una opción americana $g_n(x) = a[x - K]_+$.

Como se ha mencionado para encontrar el tiempo de paro óptimo debemos analizar el dominio del paro óptimo $\Gamma_n[I_n]$. La existencia de este dominio está sustentada en el teorema de paro óptimo de Shiryaev y su unicidad fue discutida en la sección anterior.

De acuerdo a la observación 2 de la sección anterior se puede empezar por reducir el dominio de paro óptimo $\Gamma_n[I_n] = \{x \in \mathbb{R}^+ : g_n(x) = \omega(x, y)\}$ quitando los puntos donde la función $g_n(x) = 0$, esto sucede cuando $x \leq K_n$ y si $x > K_n$ se tiene $g(x) > 0$. Entonces $C_n = [0, K_n] = \{x \in \mathbb{R}^+ : g_n(x) = 0\}$, por comodidad se desea trabajar en un intervalo cerrado, por lo que es conveniente tomar a $C_n^* = [0, K_n)$. El dominio de paro óptimo reducido queda como:

$$\bar{\Gamma}_n[y] = \Gamma_n[y] \setminus C_n^* = \{x \geq K_n : g_n(x) = \omega_n(x, y)\}.$$

A continuación se presentarán las hipótesis necesarias del proceso $\{Z_n\}_{n \in \mathbb{N}}$ para poder hacer el análisis de la estructura de los dominios de paro óptimo.

A: Para toda $n \in \mathbb{N}$ existe un kernel de transición $P_n(y, B)$, $y \in Y$, $B \in \mathcal{B}_Y$, y $P_n(x, y, z, A)$, $x \in \mathbb{R}^+$, $y, z \in Y$, $A \in \mathcal{B}_+$ tal que para toda $x \in \mathbb{R}^+$, $y \in Y$, $A \in \mathcal{B}_+$, $B \in \mathcal{B}_Y$ se tiene que

$$P_n(x, y, A, B) = \int_B P_n(y, dz) P_n(x, y, z, A),$$

entonces la componente índice I_n es un proceso de Markov no homogéneo con kernel de transición $P_n(y, B)$. La función $P_n(x, y, z, A)$ es el kernel de la probabilidad de transición del precio S_{n+1} con respecto a las variable S_n , I_n , y I_{n+1} .

Cuando el precio está dado de forma dinámica (4.1.6) se sigue manteniendo valida esta condición, ya que (4.1.7) y (4.1.9) son casos particulares de **A**.

B: Para toda $n \in \mathbb{N}$, $y, z \in Y$, $s \geq 0$ y $0 \leq x' \leq x'' < \infty$:

$$P_n(x', y, z, [s, \infty)) \leq P_n(x'', y, z, [s, \infty)).$$

Con esto se pide que el proceso $\{S_n\}_{n \in \mathbb{N}}$ sea monótono creciente con respecto al nivel del precio dado. Si el precio está dado en forma dinámica se puede reescribir como:

B': Para toda $n \in \mathbb{N}$ y $y \in Y$ la función $A_n(x, y)$ es monótona no decreciente en $x \geq 0$.

Esta condición es necesaria, ya que sería ilógico pensar que la probabilidad de que el precio alcance un precio alto va ser mayor mientras se mas pequeño sea el precio actual.

4. Optimización de opciones Americanas, con función de pago convexa, a tiempo discreto 36

La siguiente condición pide que la distribución del proceso del precio $P_n(x, y, z, A)$ sea convexa estocásticamente con respecto al nivel del precio.

Para enunciar esta propiedad usamos la función de distribución del proceso

$$F_{n,x,y,z}(s) = P_n(x, y, z, [0, s]), \quad \text{para toda } s \geq 0,$$

y la función cuantil

$$F_{n,x,y,z}^{-1}(u) = \sup\{s \geq 0 : F_{n,x,y,z}(s) \leq u\}, \quad 0 \leq u \leq 1.$$

Sea ρ una variable aleatoria uniforme en el $[0, 1]$, entonces se tiene que

$$P(F_{n,x,y,z}^{-1}(\rho) \leq s) = F_{n,x,y,z}(s), \quad s \geq 0.$$

Con estas especificaciones la propiedad de la convexidad estocástica es la siguiente.

C: Para toda $n \in \mathbb{N}, y \in Y, s \geq 0, B \in \mathcal{B}_Y$ y $0 \leq x' \leq x'' < \infty$:

$$P_n\left(\frac{x' + x''}{2}, y, z, [s, \infty)\right) \leq P\left\{\frac{F_{n,x',y,z}^{-1}(\rho) + F_{n,x'',y,z}^{-1}(\rho)}{2} \geq s\right\}.$$

En el caso que el proceso del precio es dado en forma dinámica la condición **C** se puede escribir como:

C': Para toda $n \in \mathbb{N}$ y $y \in Y$ la función $A_n(x, y)$ es convexa para toda $x \geq 0$.

A lo largo del análisis se trabajará con una familia específica de funciones, esta familia será denotada por \mathbf{V} . Una función $f(x, y) : Z \rightarrow [0, +\infty]$ está en \mathbf{V} si cumple con:

- i) ser medibles,
- ii) no decreciente en $x \geq 0$, para toda $y \in Y$, y
- iii) convexa en $x \geq 0$, para toda $y \in Y$.

Observe que si $f(x, y) \in \mathbf{V}$ y sólo toma valores en $[0, \infty)$, entonces $f(x, y)$ es una función continua en $x \geq 0$ para toda $y \in Y$, esto se deduce porque es no decreciente y convexa con respecto a $x \geq 0$.

Ahora se verá que el espacio de funciones \mathbf{V} es invariante con respecto al operador T_n , es decir que si $f(x, y) \in \mathbf{V}$ entonces $T_n f(x, y) \in \mathbf{V}$.

Lema 4.2.1 *Bajo las condiciones A-C. Si una función $f(x, y)$ está en \mathbf{V} entonces para toda $n \in \mathbb{N}$ la función $T_n f(x, y)$ está en \mathbf{V} .*

Demostración. Sea $n \in \mathbb{N}$ fija y $f(x, y) \in \mathbf{V}$, entonces $f(x, y) \geq 0$ para toda $(x, y) \in Z$ y medible con respecto a \mathcal{B}_Z . Al ser T_n el operador esperanza condicional se tiene que $T_n f(x, y) = E_{x,y}^n [f(S_{n+1}, I_{n+1})]$ es una función no negativa y medible con respecto a \mathcal{B}_Z .

Ahora se verá que $T_n f(x, y)$ es no decreciente en $x \geq 0$ para toda $y \in Y$. Gracias a la condición **A** se tiene que,

$$\begin{aligned} T_n f(x, y) &= \int_Y \int_{\mathbb{R}^+} f(u, z) P_n(x, y, du, dz) \\ &= \int_Y \int_{\mathbb{R}^+} f(u, z) P_n(y, dz) P_n(x, y, z, du) = \int_Y P_n(y, dz) \int_{\mathbb{R}^+} f(u, z) P_n(x, y, z, du), \end{aligned}$$

por lo que será suficiente con mostrar que la función

$$f_n(x, y, z) = \int_{\mathbb{R}^+} f(u, z) P_n(x, y, z, du)$$

es no decreciente en $x \geq 0$ para toda $y \in Y$.

Considere los conjuntos $A_{k,h,z} = \{u \geq 0 : kh \leq f(u, z) < (k+1)h\}$ para $k = 0, 1, \dots$ y $A_{\infty,z} = \{u \geq 0 : f(u, z) = \infty\}$. Como $f(x, y)$ es no decreciente, el conjunto $A_{\infty,z}$ debe de ser vacío o un intervalo de forma (a, ∞) o $[a, \infty)$, de la condición **B** se tiene que $P_n(x, y, z, A_{\infty,z})$ es no decreciente en $x \geq 0$, entonces, si existe $x' \geq 0$ tal que $P_n(x', y, z, A_{\infty,z}) > 0$ el conjunto $A_{\infty,z}$ es no vacío, y para toda $x'' \geq x'$ se tiene

$$\begin{aligned} f_n(x'', y, z) &= \int_{\mathbb{R}^+} f(u, z) P_n(x, y, z, du) \\ &= \int_{\mathbb{R}^+ \setminus A_{\infty,z}} f(u, z) P_n(x, y, z, du) + \int_{A_{\infty,z}} f(u, z) P_n(x, y, z, du) \\ &= \int_{\mathbb{R}^+ \setminus A_{\infty,z}} f(u, z) P_n(x, y, z, du) + \int_{A_{\infty,z}} \infty P_n(x, y, z, du) = \infty, \end{aligned}$$

debido a que $f_n(x'', y, z) = \infty$ cuando existe $x' \geq 0$ tal que $P_n(x', y, z, A_{\infty,z}) > 0$, la monotonía sólo debe ser verificada para las $x \geq 0$ tales que $P_n(x, y, z, A_{\infty,z}) = 0$, sea $C_{y,z} \subset \mathbb{R}^+$ el conjunto de las $x \geq 0$ con esta última propiedad.

Para $h > 0$ fija, defínase la función $f^{(h)}(u, z)$ como sigue:

$$\begin{aligned} f^{(h)}(u, z) &= \left\{ \begin{array}{ll} kh & \text{si } u \in A_{k,h,z} \quad k = 0, 1, \dots \\ \infty & \text{si } u \in A_{\infty,z} \end{array} \right\}, y \\ f_n^{(h)}(x, y, z) &= \int_{\mathbb{R}^+} f^{(h)}(u, z) P_n(x, y, z, du). \end{aligned}$$

Por construcción

$$\begin{aligned}
 f^{(h)}(u, z) &\leq f(u, z) \leq f^{(h)}(u, z) + h \\
 \Rightarrow \int_{\mathbb{R}^+} f^{(h)}(u, z) P_n(x, y, z, du) &\leq \int_{\mathbb{R}^+} f(u, z) P_n(x, y, z, du) \leq \\
 &\int_{\mathbb{R}^+} [f^{(h)}(u, z) + h] P_n(x, y, z, du) \\
 \Rightarrow \int_{\mathbb{R}^+} f^{(h)}(u, z) P_n(x, y, z, du) &\leq f_n(x, y, z) \leq \\
 \int_{\mathbb{R}^+} f^{(h)}(u, z) P_n(x, y, z, du) + \int_{\mathbb{R}^+} h P_n(x, y, z, du) \\
 \Rightarrow \sum_{k=0}^{\infty} \int_{A_{k,h,z}} kh P_n(x, y, z, du) + \int_{A_{\infty,z}} \infty P_n(x, y, z, du) &\leq f_n(x, y, z) \leq \\
 \sum_{k=0}^{\infty} \int_{A_{k,h,z}} kh P_n(x, y, z, du) + \int_{A_{\infty,z}} \infty P_n(x, y, z, du) + h
 \end{aligned}$$

como $P_n(x, y, z, w) = 0$ si $x \in A_{\infty,z}$ entonces $\int_{A_{\infty,z}} \infty P_n(x, y, z, du) = 0$; al tomar el límite sobre h cuando tiende a cero, se tiene

$$\begin{aligned}
 \lim_{h \rightarrow 0} f_n^{(h)}(x, y, z) &\leq \lim_{h \rightarrow 0} f_n(x, y, z) \leq \lim_{h \rightarrow 0} \left\{ f_n^{(h)}(x, y, z) + h \right\} \\
 \Rightarrow \lim_{h \rightarrow 0} f_n^{(h)}(x, y, z) &\leq f_n(x, y, z) \leq \lim_{h \rightarrow 0} f_n^{(h)}(x, y, z) + \lim_{h \rightarrow 0} h \\
 \Rightarrow \lim_{h \rightarrow 0} f_n^{(h)}(x, y, z) &= f_n(x, y, z), \quad x \geq 0, y, z \in Y.
 \end{aligned}$$

Así que es equivalente mostrar que $f_n^{(h)}(x, y, z)$ es no decreciente para las $x \in C_{y,z}$ y para toda $y, z \in Y$, con $h > 0$. Como $f(u, z)$ es no decreciente para $u \geq 0$ los conjuntos $A_{k,h,z}$, $k = 0, 1, \dots$, deben ser intervalos disjuntos tales que $B_{k,h,z} = \bigcup_{r \geq k} A_{r,h,z}$ es el vacío o un intervalo de la forma $(b_{k,h,z}, \infty)$ o $[b_{k,h,z}, \infty)$. La función $f_n^{(h)}(x, y, z)$ se puede ver como

$$\begin{aligned}
 f_n^{(h)}(x, y, z) &= \int_{\mathbb{R}^+} hk P_n(x, y, z, du) \\
 &= h \sum_{k=0}^{\infty} k P_n(x, y, z, A_{k,h,z}) = h \left\{ \sum_{k=1}^{\infty} P_n(x, y, z, A_{k,h,z}) + \sum_{k=2}^{\infty} P_n(x, y, z, A_{k,h,z}) + \dots \right\} \\
 &= h \sum_{k \geq 1} P_n(x, y, z, B_{k,h,z}),
 \end{aligned}$$

de acuerdo a la condición **B** la $P_n(x, y, z, [a, \infty])$ es monótona no decreciente para $x \geq 0$, entonces $f_n^{(h)}(x, y, z)$ al ser una combinación lineal positiva de funciones monótonas no decrecientes también lo es.

4. Optimización de opciones Americanas, con función de pago convexa, a tiempo discreto 39

Ahora se probará que $T_n f(x, y)$ es convexa para $x \geq 0$, $y \in Y$. Para esto se usarán las siguientes variables aleatorias

$$\begin{aligned} S' &= F_{n,x',y,z}^{-1}(\rho) & S'' &= F_{n,x'',y,z}^{-1}(\rho), \\ \tilde{S} &= F_{n,v,y,z}^{-1}(\rho) & \text{donde } v &= \frac{x' + x''}{2}, \end{aligned}$$

donde ρ es una variable aleatoria con distribución uniforme en el intervalo $[0, 1]$.

Por hipótesis $f(x, y)$ es convexa en $x \geq 0$, entonces,

$$\frac{f(x', y) + f(x'', y)}{2} \geq f\left(\frac{x' + x''}{2}, y\right),$$

para $x', x'' \geq 0$, como aplicar el operador T es tomar una esperanza, lo que preserva igualdades, se tiene que,

$$\begin{aligned} \frac{T_n f(x', y) + T_n f(x'', y)}{2} &= \int_Y P_n(y, dz) E \left[\frac{f(S', y) + f(S'', z)}{2} \right] \geq \\ &\int_Y P_n(y, dz) E \left[f\left(\frac{S' + S''}{2}, z\right) \right], \end{aligned}$$

por la condición **C** $\frac{S' + S''}{2}$ es estocásticamente mayor a \tilde{S} , entonces,

$$E \left[f\left(\frac{S' + S''}{2}, z\right) \right] \geq E \left[f(\tilde{S}, z) \right]$$

por lo tanto $T_n f(x, y)$ es convexa, puesto que

$$\frac{T_n f(x', y) + T_n f(x'', y)}{2} \geq \int P_n(y, dz) E [f(S, z)] = T_n f\left(\frac{x' + x''}{2}, y\right).$$

■

La última hipótesis que se incluirá es sobre el comportamiento del rendimiento de la opción; está dada en la siguiente condición, que quiere decir que cuando el precio del activo subyacente crece mucho lo más conveniente es ejercer la opción inmediatamente, ya que el rendimiento esperado se vuelve menor al que dan los coeficientes y por lo general los coeficientes deben ser crecientes en el tiempo por lo que $\frac{a_n}{a_{n+1}} e^{r_n} < 1$ en la mayoría de los casos.

D: Para toda $n \in \mathbb{N}$ y $y \in Y$:

$$\lim_{x \rightarrow \infty} \frac{1}{x} E_{x,y}^n [S_{n+1}] < \frac{a_n}{a_{n+1}} e^{r_n}.$$

En el caso donde el proceso del precio está dado en forma dinámica, la condición **D**, se puede reescribir como:

D': Para toda $n \in \mathbb{N}$ y $y \in Y$:

$$\lim_{x \rightarrow \infty} \frac{1}{x} E [A_{n+1}(x, B_{n+1}(y, \xi_{n+1}))] < \frac{a_n}{a_{n+1}} e^{rn}.$$

Por las condiciones **A-D** se tiene que $E_{x,y}^n [S_{n+1}] < \infty$ para toda $n \in \mathbb{N}$ y $x \geq 0$, $y \in Y$. El límite de la condición **D** siempre existe bajo las condiciones **A-C**, ya que por el Lema 4.2.1 la función $h_n(x, y) = E_{x,y}^n [S_{n+1}]$ está dentro del espacio **V** para toda $n \in \mathbb{N}$, entonces es continua y por consiguiente el límite existe.

En el siguiente lema se analizarán propiedades de la función $w_n(x, y)$, la cual describe las ganancias.

Lema 4.2.2 *Bajo las condiciones A-D. Para toda $n = 0, 1, \dots, N$:*

- a. $w_n(x, y) < \infty$, $x \geq 0$, $y \in Y$.
- b. $w_n(x, y)$ es continua en $x \geq 0$ para toda $y \in Y$.
- c. $w_n(x, y) \in V$.

Una observación importante que se desprende de este lema es: que el valor supremo de la funcional (4.1.1) es finito, puesto que como se tiene

$$\phi_g(\tau_{opt}) = w_N(S_0, I_0) < \infty,$$

La demostración de este lema se hará en forma conjunta con la demostración del siguiente teorema, ya que tienen una gran relación. En el siguiente teorema se inicia el análisis de la estructura del dominio de paro óptimo, y la primera característica de la estructura del dominio de paro óptimo, es que este debe ser un intervalo.

Teorema 4.2.3 *Bajo las condiciones A-D (en el caso de que el proceso de precios este dado en forma dinámica las condiciones B-D pueden ser sustituidas por B'-D'). Se tiene que:*

1. para la función de pagos (4.2.12) y todo $n = 0, 1, \dots, N - 1$ existe una única raíz $d_n^*(I_n)$ para la ecuación

$$g_n(x) = e^{-rn} T_n w_{N-n-1}(x, I_n), \quad x \geq K_n, \quad (4.2.13)$$

más aun, $g_n(x) > e^{-rn} T_n w_{N-n-1}(x, I_n)$ para las $x > d_n^*(I_n)$ y $g_n(x) < e^{-rn} T_n w_{N-n-1}(x, I_n)$ para $K_n \leq x < d_n^*(I_n)$, si $K_n < d_n^*(I_n)$;

2. el dominio de paro óptimo tiene la siguiente forma

$$\Gamma_n[I_n] = [d_n^*(I_n), \infty) \subseteq [K_n, \infty),$$

por eso, el tiempo de paro óptimo que maximiza a la funcional $\phi_g(\tau)$ está dado por $\tau_{opt} = \min\{0 \leq n \leq N : S_n \geq d_n^*(I_n)\}$, donde $d_N^*(I_N) = 0$.

Para poder demostrar el Lema 4.2.2 y Teorema 4.2.3 se introduce una nueva función, en la cual se estudian en conjunto las dos funciones que intervienen en la definición de la función de ganancias $w_n(x, y)$, $n = 0, 1, \dots, N$, tal función está definida como:

$$\varphi_n(x, y) = e^{-r_n} T_n w_{N-n-1}(x, I_n) - g_n(x), \quad n = 0, 1, \dots, N-1, \quad (4.2.14)$$

y algunas de sus características son descritas en el siguiente lema.

Lema 4.2.4 *Bajo las condiciones A-D. Para toda $n = 0, 1, \dots, N-1$:*

- a. $\varphi_n(x, y) < \infty$, $x \geq 0$, $y \in Y$,
- b. $\varphi_n(x, y)$ es continua en $x \geq 0$ para toda $y \in Y$,
- c. $\varphi_n(x, y)$ es no negativa, no decreciente y convexa si $x \in [0, K_n]$ para toda $y \in Y$,
- d. $\varphi_n(x, y)$ es monótona decreciente y convexa si $x \in [K_n, \infty)$ para toda $y \in Y$, y,
- e. existe el $\lim_{x \rightarrow \infty} \frac{1}{x} \varphi_n(x, y) < 0$ para toda $y \in Y$.

A continuación se demostrarán los lemas 4.2.2 y 4.2.4, y el Teorema 4.2.3.

Demostración. Primero se mostrará la propiedad (c) del Lema 4.2.2, es decir que, $w_n(x, y) \in V$ para toda $n = 0, 1, \dots, N$.

Esto se hará por medio de inducción.

- i) Si $n = 0$, se tiene que $w_0(x, y) = g_N = a_N[x - K_N]^+$, la cual es medible, no decreciente y convexa para $x \geq 0$ y toda $y \in Y$; por lo tanto $w_0(x, y) \in V$; y por el lema 4.2.1 la función $T_{N-1}w_0(x, y) \in V$, como $w_1(x, y) = \max\{g_{N-1}(x), e^{-r_{N-1}} T_{N-1}w_0(x, y)\}$ está en V , ya que es el máximo de dos funciones en V .
- ii) Hipótesis de inducción, suponer que $w_k(x, y) \in V$.
- iii) Sea $n = k + 1$, como $w_{k+1}(x, y) = \max\{g_{N-k-1}(x), e^{-r_{N-k-1}} T_{N-k-1}w_k(x, y)\}$, por el Lema 4.2.1 se sabe que $T_{N-k-1}w_k(x, y) \in V$, entonces $w_{k+1}(x, y) \in V$ ya que es el máximo de dos funciones en V .

La propiedad (b) del Lema 4.2.2 es una consecuencia de la propiedad (c) que se acaba de demostrar; ya que $w_n(x, y)$ es convexa y no decreciente, entonces $w_n(x, y)$ es continua para toda $n = 0, 1, \dots, N$.

Ahora se verá la propiedad (c) del Lema 4.2.4; por el Lema 4.2.1 $e^{-rn} T_n w_{N-n}(x, y) \in V$, para toda $n = 0, 1, \dots, N$, y como $g_n(x) = a_n [x - K_n]_+ = 0$ si $x \in [0, K_n]$, entonces $\varphi_n(x, y) = e^{-rn} T_n w_{N-n}(x, y) \in V$ si $x \in [0, K_n]$ para toda $y \in Y$. Por lo tanto $\varphi_n(x, y)$ es convexa, no decreciente y no negativa si $x \in [0, K_n]$ para toda $y \in Y$.

Si $x \in [K_n, \infty]$ se tiene que $-g_n(x) = -a_n(x - K_n)$, entonces es convexa en $x \in [K_n, \infty]$ por ser lineal; por lo tanto $\varphi_n(x, y)$ es convexa en $x \in [0, K_n]$ para toda $y \in Y$ por ser la suma de dos funciones convexas. Entonces $\varphi_n(x, y)$ es convexa en \mathbb{R}^+ , $y \in Y$.

Para probar el resto de las propiedades de $w_n(x, y)$, $n = 1, 2, \dots, N$ y $\varphi_n(x, y)$, $n = 0, 1, \dots, N - 1$ declaradas en los lemas 4.2.2 y 4.2.4 se hará uso de inducción hacia atrás.

Las propiedades restantes son:

Lema 4.2.2

(a') $w_n(x, y) < \infty$, $x \geq 0$, $y \in Y$.

Lema 4.2.4

(a) $\varphi_n(x, y) < \infty$, $x \geq 0$, $y \in Y$.

(b) $\varphi_n(x, y)$ es continua en $x \geq 0$ para toda $y \in Y$.

(c) $\varphi_n(x, y)$ es monótona decreciente y convexa si $x \in [K_n, \infty)$ para toda $y \in Y$, y

(d) existe el $\lim_{x \rightarrow \infty} \frac{1}{x} \varphi_n(x, y) < 0$ para toda $y \in Y$.

i) Para $n = N - 1$ y $y \in Y$,

$$\begin{aligned} \frac{1}{x} \varphi_{N-1}(x, y) &= \frac{1}{x} e^{-rN-1} T_{N-1} w_0(x, y) - \frac{1}{x} g_{N-1}(x) = \\ &= \frac{1}{x} e^{-rN-1} E_{x,y}^{N-1} [a_N [x - K_N]_+] - \frac{1}{x} a_{N-1} [x - K_{N-1}]_+ \end{aligned}$$

si $x \geq \max\{K_N, K_{N-1}\}$,

$$\frac{1}{x} \varphi_{N-1}(x, y) = \frac{a_N}{x} e^{-rN-1} (E_{x,y}^{N-1} [x] - K_N) - \frac{1}{x} a_{N-1} (x - K_{N-1})$$

como $K_n \geq 0$,

$$\frac{1}{x} \varphi_{N-1}(x, y) \leq a_N e^{-rN-1} \frac{1}{x} E_{x,y}^{N-1} S_N - a_{N-1} \left(1 - \frac{K_{N-1}}{x}\right),$$

al tomar el límite cuando x tiende a ∞ y por la condición **D**, se concluye que

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{1}{x} \varphi_{N-1}(x, y) &\leq a_N e^{-r_{N-1}} \lim_{x \rightarrow \infty} \frac{1}{x} E_{x,y}^{N-1} S_N - a_{N-1} \lim_{x \rightarrow \infty} \left(1 - \frac{K_{N-1}}{x}\right) \\ &< a_N e^{-r_{N-1}} \frac{a_{N-1}}{a_N} e^{r_{N-1}} - a_{N-1} = 0. \end{aligned} \quad (4.2.15)$$

Entonces $\varphi_{N-1}(x, y) < 0$ para valores grandes de x , es decir para $x > \varepsilon_{N-1}(y)$, $y \in Y$, por lo tanto $T_{N-1}w_0(x, y) < \infty$ para $x > \varepsilon_{N-1}(y)$, pero por el Lema 4.2.1 se sabía que $T_{N-1}w_0(x, y)$ es continua y no decreciente, entonces $T_{N-1}w_0(x, y)$ es continua y finita para $x \geq 0$, por lo tanto $w_1(x, y) = \max\{g_{N-1}(x), e^{-r_{N-1}} T_{N-1}w_0(x, y)\}$ y $\varphi_{N-1}(x, y) = e^{-r_{N-1}} T_{N-1}w_0(x, y) - g_{N-1}(x)$ son funciones finitas y continuas para $x \geq 0$ y toda $y \in Y$.

Ahora se verá que $\varphi_{N-1}(x, y)$ es monótona decreciente si $x \in [K_{N-1}, \infty)$, $y \in Y$

$$\begin{aligned} \varphi_{N-1}(K_{N-1}, y) &= e^{-r_{N-1}} T_{N-1}w_0(K_{N-1}, y) - g_{N-1}(K_{N-1}) \\ &= e^{-r_{N-1}} T_{N-1}w_0(K_{N-1}, y) \geq 0, y \\ \varphi_{N-1}(x, y) &\rightarrow l_{N-1} < 0 \text{ si } x \rightarrow \infty, \end{aligned}$$

junto con la convexidad de $\varphi_{N-1}(x, y)$ en $[K_{N-1}, \infty]$ es imposible la existencia de puntos $K_{N-1} \leq x_1 < x_2$ tales que $\varphi_{N-1}(x_1, y) = \varphi_{N-1}(x_2, y)$. Para demostrar esto supongamos que existen tales puntos, pero también siempre existe $x_3 > x_2$ tal que $\varphi_{N-1}(x_3, y) < \varphi_{N-1}(x_2, y)$, podemos hacer a $x_2 = px_1 + (1-p)x_3$, con $p = \frac{x_3 - x_2}{x_3 - x_1}$, entonces

$$\begin{aligned} (1-p)\varphi_{N-1}(x_3, y) &< (1-p)\varphi_{N-1}(x_2, y) \\ \Rightarrow p\varphi_{N-1}(x_2, y) + (1-p)\varphi_{N-1}(x_3, y) &< \varphi_{N-1}(x_2, y) \\ \Rightarrow p\varphi_{N-1}(x_1, y) + (1-p)\varphi_{N-1}(x_3, y) &< \varphi_{N-1}(x_2, y) ! \end{aligned}$$

contradice la convexidad de $\varphi_{N-1}(x, y)$, por lo tanto no existen tales puntos x_1 y x_2 . Por lo tanto $\varphi_{N-1}(x, y)$ es estrictamente decreciente.

Con esto probamos que $w_1(x, y)$ y $\varphi_{N-1}(x, y)$ poseen las propiedades.

- ii) Por paso inductivo se supone que $w_k(x, y)$ y $\varphi_{N-k}(x, y)$ cumplen las propiedades.
- iii) Falta probar que $w_{k+1}(x, y)$ y $\varphi_{N-k-1}(x, y)$ las cumplen.

Para mostrar esto se debe conseguir una relación análoga a la (4.2.15) para la función $\varphi_{N-k-1}(x, y)$. Debido a la hipótesis de inducción el $\lim_{x \rightarrow \infty} \frac{1}{x} \varphi_{N-k}(x, y) < 0$ existe, $y \in Y$, por lo que $\varphi_{N-k}(x, y) < 0$ para $x > \varepsilon_{N-k}(y)$, $y \in Y$, entonces

$$\begin{aligned} e^{-r_{N-k}} T_{N-k} w_{k-1}(x, y) &< g_{N-k}(x) \text{ para } x > \varepsilon_{N-k}(y), y \in Y \\ \Rightarrow w_k(x, y) &= g_{N-k}(x) \text{ para toda } x > \varepsilon_{N-k}(y) \end{aligned}$$

por definición $w_k(x, y) \geq g_{N-k}(x)$, para toda $x \in \mathbb{R}^+$; con estas ideas en mente se puede escribir $w_k(x, y) = g_{N-k}(x) + (w_k(x, y) - g_{N-k}(x))$, quedando como la suma de dos funciones no negativas. La segunda función $(w_k(x, y) - g_{N-k}(x))$ se hace cero si $x > \varepsilon_{N-k}(y)$, y por hipótesis de inducción ambas funciones son continuas. Como el operador T_{N-k-1} es lineal se tiene

$$T_{N-k-1}w_k(x, y) = T_{N-k-1}g_{N-k}(x) + T_{N-k-1}(w_k(x, y) - g_{N-k}(x)).$$

Como $(w_k(x, y) - g_{N-k}(x))$ es continua entonces es acotada en el intervalo $[0, \varepsilon_{N-k}(y)]$. Ahora defínase $G_k(y) = \sup_{0 \leq x \leq \varepsilon_{N-k}(y)} (w_k(x, y) - g_{N-k}(x))$. Entonces se tiene que

$$\begin{aligned} T_{N-k-1}(w_k(x, y) - g_{N-k}(x)) &= \\ \int_Y P_{N-k-1}(y, dz) \int_0^{\varepsilon_{N-k}(y)} (w_k(x, y) - g_{N-k}(x)) P_{N-k-1}(x, y, z, du) &\leq \\ \int_Y P_{N-k-1}(y, dz) \int_0^{\varepsilon_{N-k}(y)} G_k(y) P_{N-k-1}(x, y, z, du) &\leq G_k(y), \end{aligned}$$

por lo tanto

$$\begin{aligned} 0 \leq \overline{\lim}_{x \rightarrow \infty} \frac{1}{x} T_{N-k-1}(w_k(x, y) - g_{N-k}(x)) &\leq G_k(y) \overline{\lim}_{x \rightarrow \infty} \frac{1}{x} = 0 \\ \Rightarrow \overline{\lim}_{x \rightarrow \infty} \frac{1}{x} T_{N-k-1}(w_k(x, y) - g_{N-k}(x)) &= 0, \end{aligned}$$

De la última desigualdad se concluye que los siguientes límites son iguales (finitos o infinitos simultáneamente):

$$\lim_{x \rightarrow \infty} \frac{1}{x} T_{N-k-1}w_k(x, y) = \lim_{x \rightarrow \infty} \frac{1}{x} T_{N-k-1}g_{N-k}(x) \quad (4.2.16)$$

Utilizando la propiedad **D** se obtiene

$$\begin{aligned}
 & \lim_{x \rightarrow \infty} \frac{1}{x} \varphi_{N-k-1}(x, y) = \tag{4.2.17} \\
 & e^{-r_{N-k-1}} \lim_{x \rightarrow \infty} \left[\frac{1}{x} T_{N-k-1} w_k(x, y) - a_{N-k-1} [x - K_{N-k-1}]_+ \right] = \\
 & e^{-r_{N-k-1}} \lim_{x \rightarrow \infty} \frac{1}{x} T_{N-k-1} w_k(x, y) - a_{N-k-1} = \\
 & e^{-r_{N-k-1}} \lim_{x \rightarrow \infty} \frac{1}{x} T_{N-k-1} g_{N-k}(x) - a_{N-k-1} = \\
 & e^{-r_{N-k-1}} \lim_{x \rightarrow \infty} \frac{1}{x} T_{N-k-1} a_{N-k} [x - K_{N-k}]_+ - a_{N-k-1} = \\
 & e^{-r_{N-k-1}} a_{N-k} \lim_{x \rightarrow \infty} \frac{1}{x} T_{N-k-1} [x - K_{N-k}]_+ - a_{N-k-1} \leq \\
 & e^{-r_{N-k-1}} a_{N-k} \lim_{x \rightarrow \infty} \frac{1}{x} E_{x,y}^{N-k-1} S_{N-k} - a_{N-k-1} < \\
 & e^{-r_{N-k-1}} a_{N-k} \frac{a_{N-k-1}}{a_{N-k}} e^{r_{N-k-1}} - a_{N-k-1} = 0
 \end{aligned}$$

Por lo que $\varphi_{N-k-1}(x, y) < 0$ para valores grandes de $x > \varepsilon_{N-k-1}(y)$, $y \in Y$. Esto quiere decir que $T_{N-k-1} w_k(x, y) < \infty$ para $x > \varepsilon_{N-k-1}(y)$ y por el lema 4.2.1 $T_{N-k-1} w_k(x, y)$ no decrece para $x \geq 0$ y es continua, entonces $T_{N-k-1} w_k(x, y) < \infty$ para $x \geq 0$, por lo tanto se puede concluir que $w_{k+1}(x, y) = \max\{g_{N-k}(x), T_{N-k-1} w_k(x, y)\}$ y $\varphi_{N-k-1}(x, y) = T_{N-k-1} w_k(x, y) - g_{N-k-1}(x)$ son funciones finitas y continuas para $x \geq 0$ y toda $y \in Y$.

Ahora para ver que $\varphi_{N-k-1}(x, y)$ es monótona decreciente si $x \in [K_{N-k-1}, \infty)$, $y \in Y$. Se tiene que

$$\begin{aligned}
 \varphi_{N-k-1}(K_{N-k-1}, y) &= T_{N-k-1} w_k(K_{N-k-1}, y) - g_{N-k-1}(K_{N-k-1}) = \\
 T_{N-k-1} w_k(K_{N-k-1}, y) &\geq 0 \text{ y} \\
 \varphi_{N-k-1}(x, y) &\rightarrow l_{N-k-1} < 0 \text{ si } x \rightarrow \infty
 \end{aligned}$$

por la propiedad de convexidad de $\varphi_{N-k-1}(x, y)$ en $[K_{N-k-1}, \infty)$ es imposible la existencia de puntos $K_{N-k-1} \leq x_1 < x_2$ tales que $\varphi_{N-k-1}(x_1, y) = \varphi_{N-k-1}(x_2, y)$. Supóngase que existen tales puntos, entonces también siempre existe $x_3 > x_2$ tal que $\varphi_{N-k-1}(x_3, y) < \varphi_{N-k-1}(x_2, y)$, y podemos hacer que $x_2 = px_1 + (1-p)x_3$, con $p = \frac{x_3 - x_2}{x_3 - x_1}$, entonces

$$\begin{aligned}
 & (1-p)\varphi_{N-k-1}(x_3, y) < (1-p)\varphi_{N-k-1}(x_2, y) \\
 & \Rightarrow p\varphi_{N-k-1}(x_2, y) + (1-p)\varphi_{N-k-1}(x_3, y) < \varphi_{N-k-1}(x_2, y) \\
 & \Rightarrow p\varphi_{N-k-1}(x_1, y) + (1-p)\varphi_{N-k-1}(x_3, y) < \varphi_{N-k-1}(x_2, y)!
 \end{aligned}$$

que es absurdo debido a la convexidad $\varphi_{N-k-1}(x, y)$, por lo tanto no existen los puntos x_1 y x_2 . Por lo tanto $\varphi_{N-k-1}(x, y)$ es estrictamente decreciente. y con esto se concluye la inducción.

Con esto han quedado demostrados los Lemas 4.2.2 y 4.2.4. La demostración del teorema 4.2.3 se desprende de estos Lemas; para ver que la ecuación (4.2.13)

$$g_n(x) = e^{-r_n} T_n w_{N-n-1}(x, I_n), \quad x \geq K_n,$$

tiene una única raíz $d^*(I_n)$, basta con ver que existe una única $x \geq K_n$ talque

$$\varphi_n(x, I_n) = e^{-r_n} T_n w_{N-n-1}(x, I_n) - g_n(x) = 0, \quad x \geq K_n,$$

por el Lema 4.2.4 se puede concluir que la función $\varphi(x, y)$ tiene una única raíz debido a que en el intervalo $[0, K_n)$ es positiva y no decreciente, y en el intervalo $[K_n, \infty)$ se vuelve monótona decreciente y se sabe que para valores grandes de x tomará valores negativos; entonces por este cambio de signo se puede asegurar la existencia de la raíz, la unicidad se debe a la monotonía de la función. Como la raíz es única y $\varphi_{N-k-1}(x, I_n)$ es negativa para valores grandes de x y monótona decreciente entonces $g_n(x) > e^{-r_n} T_n w_{N-n-1}(x, y)$ para $x > d_n^*(y)$, como en $[K_n, \infty)$ es monótona decreciente y positiva en $K_n < x < d_n^*(y)$ se tiene que $g_n(x) < e^{-r_n} T_n w_{N-n-1}(x, y)$ para $K_n < x < d_n^*(y)$, si $K_n < d_n^*(y)$.

Debido a esto el dominio de paro óptimo tiene la siguiente transformación

$$\begin{aligned} \Gamma_n[y] &= \{x \geq K_n : g_n(x) = w_{N-n}(x, y)\} = & (4.2.18) \\ & \{x \geq K_n : g_n(x) = \max \langle g_n(x), e^{-r_n} T_n w_{N-n-1}(x, y) \rangle\} \\ & = [d_n^*(y), \infty), \quad y \in Y. \end{aligned}$$

Para concluir la demostración esto es valido para toda $y \in Y$, en particular para $y = I_n$. ■

En el Lema 4.2.2 se ve que la funcional $\phi_g(\tau_{opt}) = w_N(S_0, I_0)$ tiene una dependencia del valor inicial del precio S_0 , por este Lema se sabe que $\phi_g(\tau_{opt})$ es una función no negativa, continua, no decreciente y convexa para toda $S_0 \geq 0$ para toda $I_0 \in Y$. Por las desigualdades (4.2.16) y (4.2.17) usadas en la demostración anterior se puede concluir

$$\lim_{S_0 \rightarrow \infty} \frac{1}{S_0} \phi_g(\tau_{opt}) = \max\{a_0, e^{-r_0} a_1 \lim_{x \rightarrow \infty} \frac{1}{x} E_{x,y}^0 S_1\} < \infty$$

puesto que

$$\begin{aligned}
 \lim_{S_0 \rightarrow \infty} \frac{1}{S_0} \phi_g(\tau_{opt}) &= \lim_{S_0 \rightarrow \infty} \frac{1}{S_0} w_N(S_0, I_0) = \\
 &= \max\left\{ \lim_{S_0 \rightarrow \infty} \frac{1}{S_0} g_0(x_0), \lim_{S_0 \rightarrow \infty} \frac{1}{S_0} e^{-r_0} T_0 w_{N-1}(S_0, I_0) \right\} \\
 &= \max\left\{ \lim_{S_0 \rightarrow \infty} \frac{1}{S_0} a_0 [S_0 - K_0]^+, e^{-r_0} \lim_{S_0 \rightarrow \infty} \frac{1}{S_0} T_0 g_1(S_1) \right\} \\
 &= \max\left\{ a_0, e^{-r_0} a_1 \lim_{x \rightarrow \infty} \frac{1}{x} E_{x,y}^0 S_1 \right\}.
 \end{aligned}$$

Por esta desigualdad podemos saber que $\phi_g(\tau_{opt})$ crece a lo más como una función lineal de S_0 .

Cuando el proceso del precio no depende de ningún proceso índice, como en las observaciones (4.1.7) y (4.1.8), la ecuación (4.2.13), de la cual se calcula la raíz que delimita el dominio de paro óptimo, se transforma en

$$g_n(x) = e^{-r_n} T_n w_{N-n-1}(x), \quad x \geq K_n,$$

entonces de acuerdo con el Teorema 4.2.3 para cada $n = 0, 1, \dots, N-1$ sólo existirá una única raíz d_n^* de esta ecuación anterior y el momento de paro óptimo queda definido por $\tau_{opt} = \min\{0 \leq n \leq N : S_n \geq d_n^*\}$, con $d_N^* = 0$.

Las condiciones del Teorema 4.2.3 no excluyen la posibilidad de que el dominio de paro sea $\Gamma_n[I_n] = [K_n, \infty)$. Para que esto suceda es necesario que $P_n(S_n, y, (K_n, \infty), X) > 0$ para toda $n = 0, 1, \dots$ y $y \in Y$, ya que está da la posibilidad de que el precio incremente y sobrepase los niveles del precio pactado, con lo cual puede ser que $d_n^*(y) > K_n$, $y \in Y$.

En el siguiente teorema se analiza el caso cuando para alguna n en particular la condición **D** no se; sin la condición **D** la función $\varphi_n(x, y)$ puede tomar el valor infinito lo cual debe verse reflejado en la forma del dominio de paro óptimo $\Gamma_n[y] = \{x \geq K_n : g_n(x) = w_{N-n}(x, y)\}$ que como se vio en la demostración del Teorema 4.2.3 es lo mismo a pedir

$$\Gamma_n[y] = \{x \geq K_n : \varphi_n(x, y) \leq 0\}, \quad y \in Y.$$

de acuerdo a (4.2.18).

Teorema 4.2.5 *Suponga que se cumplen las condiciones **A-C** (en el caso donde el proceso del precio este dado en forma dinámica como en (4.1.6) las condiciones **B-C** pueden ser remplazadas por las condiciones **B'-C'**). Entonces para la función de pagos estándar dada en (4.2.12) sucede:*

1. para toda $n = 0, 1, \dots, N-1$ el dominio de paro óptimo $\Gamma_n[I_n] \subseteq [K_n, \infty)$ es un conjunto vacío, un simple punto, un intervalo finito $[d_n^*(I_n), f_n^*(I_n)]$ o un intervalo infinito $[d_n^*, \infty)$;

2. para alguna $n = 0, 1, \dots, N - 1$ fija, el dominio de paro óptimo $\Gamma_n[I_n] = \emptyset$ si y sólo si la desigualdad

$$g_n(x) < e^{-r_n} T_n w_{N-n-1}(x, I_n)$$

se cumple para toda $x \geq K_n$.

3. para alguna $n = 0, 1, \dots, N - 1$ fija, el dominio de paro óptimo $\Gamma_n[I_n]$ consiste de un simple punto $d_n^*(y) \geq K_n$ si y sólo si la ecuación (4.2.13) tiene una única raíz y la desigualdad

$$g_n(x) < e^{-r_n} T_n w_{N-n-1}(x, I_n)$$

se cumple para $x \geq K_n$ y $x \neq d_n^*(I_n)$;

4. para alguna $n = 0, 1, \dots, N - 1$ fija, el dominio de paro óptimo $\Gamma_n[I_n]$ es un intervalo finito $[d_n^*(I_n), f_n^*(I_n)] \subset [K_n, \infty)$ si y sólo si la ecuación (4.2.13) tiene como raíz mínima a $d_n^*(I_n)$ y raíz máxima $f_n^*(I_n)$, además en este caso

$$\begin{aligned} g_n(x) &< e^{-r_n} T_n w_{N-n-1}(x, I_n), \quad \text{para } x \in [K_n, \infty) \setminus [d_n^*(I_n), f_n^*(I_n)], \\ g_n(x) &\geq e^{-r_n} T_n w_{N-n-1}(x, I_n), \quad \text{para } x \in (d_n^*(I_n), f_n^*(I_n)); \end{aligned}$$

5. para alguna $n = 0, 1, \dots, N - 1$ fija, el dominio de paro óptimo $\Gamma_n[I_n]$ es un intervalo infinito $[d_n^*(I_n), \infty) \subseteq [K_n, \infty)$ si y sólo si $d_n^*(I_n)$ es la raíz mas pequeña de la ecuación (4.2.13) y

$$g_n(x) \geq e^{-r_n} T_n w_{N-n-1}(x, I_n), \quad \text{para } x > d_n^*(I_n),$$

además en este caso también

$$g_n(x) < e^{-r_n} T_n w_{N-n-1}(x, I_n), \quad \text{para } x \in [K_n, \infty) \setminus [d_n^*(I_n), \infty).$$

Demostración.

1. En la observación sobre la unicidad del dominio de paro óptimo se vio como el refinamiento de un dominio de paro óptimo debe estar incluido en el dominio de paro óptimo inicial

$$\Gamma_n^*(I_n) = \{x \geq K_n : g_n(x) = w_{N-n}(x, y)\} = [K_n, \infty),$$

entonces $\Gamma_n(I_n) \subseteq [K_n, \infty)$ y puede ser vacío, un punto, un intervalo finito o un intervalo infinito.

2. Sea n fija, si $\Gamma_n [I_n] = \emptyset$ quiere decir que $\emptyset = \{x \geq K_n : \varphi_n(x, y) \leq 0\}$, y esto sucede si y sólo si

$$g_n(x) < e^{-rn} T_n w_{N-n-1}(x, I_n) \text{ para toda } x \geq K_n.$$

3. Sea n fija, si $\Gamma_n [I_n] = d^*(I_n)$ quiere decir que $d^*(I_n) = \{x \geq K_n : \varphi_n(x, y) \leq 0\}$, lo cual sucede si y sólo si

$$g_n(x) = e^{-rn} T_n w_{N-n-1}(x, I_n) \text{ sólo en } x = d^*(I_n);$$

si sucede esta última se ve claramente que $g_n(x) < e^{-rn} T_n w_{N-n-1}(x, I_n)$ para toda $x \in [K_n, \infty) \setminus d^*(I_n)$.

4. Sea n fija, si $\Gamma_n [I_n] = [d_n^*(I_n), f_n^*(I_n)]$ quiere decir que $[d_n^*(I_n), f_n^*(I_n)] = \{x \geq K_n : \varphi_n(x, y) \leq 0\}$, lo cual sucede si y sólo si (4.2.13) tiene como raíces a $d_n^*(I_n)$ su mínima raíz y $f_n^*(I_n)$ su raíz mayor; además sucede

$$g_n(x) < e^{-rn} T_n w_{N-n-1}(x, I_n) \text{ para } x \in [K_n, \infty) \setminus [d_n^*(I_n), f_n^*(I_n)];$$

ahora por demostración al absurdo se verá que si $x \in [d_n^*(I_n), f_n^*(I_n)]$, entonces $g_n(x) \geq e^{-rn} T_n w_{N-n-1}(x, I_n)$.

Supóngase que existe $x' \in [d_n^*(I_n), f_n^*(I_n)]$ tal que $g_n(x') < e^{-rn} T_n w_{N-n-1}(x', I_n)$, pero se sabe que $g_n(d^*(I_n)) = e^{-rn} T_n w_{N-n-1}(d^*(I_n), I_n)$ y $g_n(f^*(I_n)) = e^{-rn} T_n w_{N-n-1}(f^*(I_n), I_n)$.

Sea $p = \frac{f^*(I_n) - x'}{f^*(I_n) - d^*(I_n)}$, entonces se tiene que $x = p \cdot d^*(I_n) + (1 - p) \cdot f^*(I_n)$, además $g_n(x)$ y $e^{-rn} T_n w_{N-n-1}(x, I_n)$ son funciones convexas, entonces

$$\begin{aligned} g_n(x) &< e^{-rn} T_n w_{N-n-1}(x, I_n) \\ \Rightarrow g_n(p \cdot d^* + (1 - p) \cdot f^*) &< e^{-rn} T_n w_{N-n-1}(p \cdot d^* + (1 - p) \cdot f^*, I_n) \\ &\Rightarrow p g_n(d^*) + g_n(f^*) < \\ &pe^{-rn} T_n w_{N-n-1}(d^*, I_n) + (1 - p) f^* e^{-rn} T_n w_{N-n-1}(f^*, I_n)! \end{aligned}$$

por lo tanto no existe x' . Así que

$$g_n(x) \geq e^{-rn} T_n w_{N-n-1}(x, I_n) \text{ si } x \in [d_n^*(I_n), f_n^*(I_n)].$$

5. Sea n fija, si $\Gamma_n [I_n] = [d^*(I_n), \infty)$, quiere decir que $[d^*(I_n), \infty) = \{x \geq K_n : \varphi_n(x, y) \leq 0\}$, y esto sucede si y sólo si $d^*(I_n)$ es la mínima raíz de (4.2.13) y $g_n(x) \geq e^{-rn} T_n w_{N-n-1}(x, I_n)$, para toda $x > d_n^*(I_n)$; si se tiene esto entonces $g_n(x) < e^{-rn} T_n w_{N-n-1}(x, I_n)$, para $x \in [K_n, \infty) \setminus [d_n^*(I_n), \infty)$.

■

Ahora se verán una condición necesaria para que el tiempo de paro óptimo sea igual a la fecha de vencimiento, $\tau_{opt} = N$; la condiciones quedan en termino de los parámetros a_n y K_n de la función de pago estándar (4.2.12), estos parámetros deben ser tales que

$$\frac{a_n K_n}{a_{n+1} K_{n+1}} e^{r_n} \geq 1, \quad n = 0, 1, \dots, N. \quad (4.2.19)$$

En este caso no es necesario que se cumplan las condiciones **A-C**, y la condición **D** es negada puesto que se espera que $S_n > K_n$ para toda n , por lo que la condición **D** se cambia

$$\bar{\mathbf{D}}: \frac{a_n}{a_{n+1}} e^{r_n} \leq \frac{1}{x} E_{x,y}^n S_{n+1} < \infty \text{ para } n \in \mathbb{N}, x > 0, y \in Y.$$

(anteriormente **D**: Para toda $n \in \mathbb{N}$ y $y \in Y : \lim_{x \rightarrow \infty} \frac{1}{x} E_{x,y}^n [S_{n+1}] < \frac{a_n}{a_{n+1}} e^{r_n}$.)

Es intuitivo saber que el tiempo de paro óptimo será N , puesto que esta condición quiere decir que el rendimiento esperado del precio siempre será mayor al rendimiento de las constantes $a_n K_n$.

Teorema 4.2.6 *Bajo las condiciones (4.2.19) y $\bar{\mathbf{D}}$, se tiene que $\tau_{opt} = N$*

Demostración. Considere al proceso

$$V_n = a_n e^{-R_n} K_n \left[\frac{S_n}{K_n} - 1 \right]_+ = a_n K_n e^{-R_n} g_0 \left(\frac{S_n}{K_n} \right), \quad n \in \mathbb{N},$$

donde $g_0(x) = [x - 1]_+$, $x \in \mathbb{R}^+$, la cual es una función convexa y no decreciente. Ahora se mostrará que este proceso es una submartingala.

Por la condición (4.2.19) se tiene que

$$\begin{aligned} E(V_{n+1} | \mathcal{F}_n) &= E \left[a_{n+1} K_{n+1} e^{-R_{n+1}} g_0 \left(\frac{S_{n+1}}{K_{n+1}} \right) | \mathcal{F}_n \right] = \\ & a_{n+1} K_{n+1} e^{-R_{n+1}} E \left[g_0 \left(\frac{S_{n+1}}{K_{n+1}} \right) | \mathcal{F}_n \right] \underset{\text{c.s.}}{\geq} \\ & a_{n+1} K_{n+1} e^{-R_{n+1}} g_0 \left(E \left[\frac{S_{n+1}}{K_{n+1}} | \mathcal{F}_n \right] \right) = \\ & a_{n+1} K_{n+1} e^{-R_{n+1}} g_0 \left(\frac{1}{K_{n+1}} E[S_n | \mathcal{F}_n] \right) \end{aligned}$$

la desigualdad se da gracias a que $g_0(x)$ es una función convexa y se puede aplicar la desigualdad de Jensen (la cual se da casi seguramente, c.s.).

Por la condición $\bar{\mathbf{D}}$

$$\begin{aligned} \frac{a_n}{a_{n+1}} e^{r_n} &\leq \frac{1}{S_n} E_{S_n, y}^n S_{n+1} \\ \Rightarrow S_n \frac{a_n}{a_{n+1}} e^{r_n} &\leq E_{S_n, y}^n S_{n+1} < \infty \text{ para } n \in \mathbb{N}, x > 0, y \in Y, \end{aligned}$$

y por la monotonía de la esperanza condicional

$$\begin{aligned} a_{n+1} K_{n+1} e^{-R_{n+1}} g_0 \left(\frac{1}{K_{n+1}} E[S_{n+1} | \mathcal{F}_n] \right) &\geq \\ a_{n+1} K_{n+1} e^{-R_{n+1}} g_0 \left(\frac{1}{K_{n+1}} S_n \frac{a_n}{a_{n+1}} e^{r_n} \right) &= \\ a_{n+1} K_{n+1} e^{-R_{n+1}} g_0 \left(\frac{a_n K_n}{a_{n+1} K_{n+1}} e^{r_n} \frac{S_n}{K_n} \right) \end{aligned}$$

ahora se vera que $g_0(\lambda x) \geq \lambda g_0(x)$, si $\lambda \geq 1$ y $x \geq 0$; sea $x \geq 0$ tal que $g_0(x) = 0$ y $\lambda \geq 1$, entonces $g_0(\lambda x) = [\lambda x - 1]_+ \geq 0 = \lambda 0 = \lambda g_0(x)$; por otro lado sea $y \geq 0$ tal que $g_0(y) > 0$ con la misma $\lambda \geq 1$ se tiene que $\lambda g_0(x) = \lambda x - \lambda \leq \lambda x - 1 = g_0(\lambda x)$; por lo tanto $g_0(\lambda x) \geq \lambda g_0(x)$, si $\lambda \geq 1$ y $x \geq 0$.

Sea $\lambda = \frac{a_n K_n}{a_{n+1} K_{n+1}} e^{r_n} \geq 1$ y recuerde que $R_n = r_0 + r_1 + \dots + r_{n-1}$, entonces $-R_{n+1} + r_n = -R_n$; siguiendo con las desigualdades anteriores

$$\begin{aligned} a_{n+1} K_{n+1} e^{-R_{n+1}} g_0 \left(\frac{a_n K_n}{a_{n+1} K_{n+1}} e^{r_n} \frac{S_n}{K_n} \right) &= \\ a_{n+1} K_{n+1} e^{-R_{n+1}} g_0 \left(\lambda \frac{S_n}{K_n} \right) &\geq \\ a_{n+1} K_{n+1} e^{-R_{n+1}} \lambda g_0 \left(\frac{S_n}{K_n} \right) &= \\ a_{n+1} K_{n+1} e^{-R_{n+1}} \left(\frac{a_n K_n}{a_{n+1} K_{n+1}} e^{r_n} \right) g_0 \left(\frac{S_n}{K_n} \right) &= \\ a_n K_n e^{-R_n} g_0 \left(\frac{S_n}{K_n} \right) &= V_n \end{aligned}$$

Por lo tanto V_n es una submartingala, entonces para cada momento de Markov $0 \leq \tau \leq N$

$$E e^{-R_\tau} g_\tau(S_\tau) = EV_\tau \leq EV_N = E e^{-R_N} g_N(S_N).$$

Por lo que se puede concluir que

$$\tau_{opt} = N.$$

■

Con este teorema se concluye el análisis del caso en el que la función de pagos tiene la forma (4.2.12), como se ha visto en este caso el dominio de paro óptimo puede

verse como un intervalo $[d_n^*(y), f_n^*(y))$ donde $f_n^*(y) \in \{\{x \geq d_n^*(y)\}, \infty\}$; entonces para encontrar el dominio de paro óptimo en este caso basta con encontrar las raíces $d_n^*(y)$ y $f_n^*(y)$; en el siguiente capítulo se darán ejemplos en los cuales se encuentra la raíz $d_n^*(y)$ en el caso donde se cumple la condición **D**. Cuando la hipótesis **D** no se cumple se deberá esperar a la fecha de vencimiento para ejercer la mejor opción que se tiene.

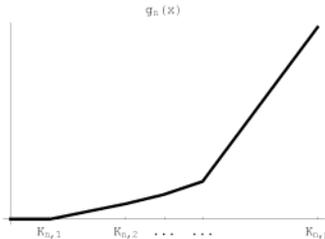
En la siguiente sección se hará uso de la estructura o forma de los dominios de paro óptimo aquí analizados para tratar de generalizar estos resultados a casos un poco más generales en cuanto a la forma de la función de pago.

4.3 Optimización en el caso cuando la función de pago es lineal por trozos.

Para generalizar el análisis de la sección anterior ahora se considerará el caso cuando la función de pago $g_n(x)$ convexa, lineal por trozos y creciente, esto quiere decir, para alguna $p \geq 2$ fija e independiente de n se deben considerar dos series de números positivos $0 < K_{n,1} < K_{n,2} < \dots < K_{n,p} < \infty$ y $0 < a_{n,1} < a_{n,2} < \dots < a_{n,p} < \infty$ conocidos y fijos, tales que la función $g_n(x)$ debe ser continua, no negativa, definida en $[0, \infty)$ talque

$$g_n(x) = \begin{cases} 0 & \text{si } x \in [0, K_{n,1}] \\ a_{n,i}[x - K_{n,i}]^+ & \text{si } x \in [K_{n,i}, K_{n,i+1}], i \in \{1, \dots, p-1\} \\ a_{n,p}[x - K_{n,p}]^+ & \text{si } x \in [K_{n,p}, \infty) \end{cases} \quad (4.3.20)$$

entonces $g_n(x)$ es lineal dentro de cada intervalo $[0, K_{n,1}], [K_{n,1}, K_{n,2}], \dots, [K_{n,p-1}, K_{n,p}]$ y $[K_{n,p}, \infty)$ con su correspondiente pendiente $a_{n,1}, \dots, a_{n,p-1}, a_{n,p}$. La grafica de dicha función para una n fija debe ser parecida a la siguiente imagen



Las condiciones **B** (**B'**) y **C** (**C'**) no requieren cambio, puesto que estas no depende de la forma de la función de pago. En **D** (**D'**) las constantes a_n deben ser remplazadas por las $a_{n,p}$.

Así como en el caso anterior donde se tenía la función de pago estándar (4.2.12) las condiciones **A**, **B** y **D** implican que $w_n(x, y) < \infty$ para toda $x \geq 0$, $y \in Y$ y

$n = 0, 1, \dots, N$, por lo tanto $\phi_g(\tau_{opt}) < \infty$, esto se demostrará dentro del siguiente teorema; en el teorema que se enuncia a continuación se darán las características que tiene los dominios de paro óptimo bajo esta función de pago (4.3.20), su demostración es una simple aplicación correcta del caso de la función de pago estándar.

Teorema 4.3.1 *Suponga que la función de pagos $g_n(x)$ es convexa y lineal por trozos, además de que se cumplen las condiciones \mathbf{A} , \mathbf{B} y \mathbf{D} (en el caso donde el precio del proceso está dado en forma dinámica por (3) las condiciones $\mathbf{B-D}$ pueden ser remplazadas por las condiciones $\mathbf{B'-D'}$). Entonces:*

1. para toda $n = 0, 1, \dots, N - 1$ e $i = 1, 2, \dots, p - 1$ la desigualdad

$$g_n(x) \geq e^{-rn} T_n w_{N-n-1}(x, I_n), \quad x \in [K_{n,i}, K_{n,i+1}],$$

tiene como solución al conjunto $J_{n,i}$, el cual puede ser vacío o un intervalo cerrado $[d_{n,i}^*(I_n), f_{n,i}^*(I_n)] \subseteq [K_{n,i}, K_{n,i+1}]$ en tal caso puede suceder que el punto izquierdo (derecho) del intervalo $J_{n,i}$ coincida con el correspondiente punto del intervalo $[K_{n,i}, K_{n,i+1}]$ o es la solución de la ecuación

$$g_n(x) = e^{-rn} T_n w_{N-n-1}(x, I_n), \quad x \in [K_{n,i}, K_{n,i+1}].$$

2. existe para toda $n = 0, 1, \dots, N - 1$ una única raíz $d_{n,p}^*(I_n)$ de la ecuación

$$g_n(x) = e^{-rn} T_n w_{N-n-1}(x, I_n), \quad x \geq K_{n,p},$$

además $g_n(x) > e^{-rn} T_n w_{N-n-1}(x, I_n)$ para $x > d_{n,p}^*(I_n)$ y $g_n(x) < e^{-rn} T_n w_{N-n-1}(x, I_n)$ para $K_{n,p} \leq x < d_{n,p}^*(I_n)$ si $K_{n,p} < d_{n,p}^*(I_n)$.

3. el dominio de paro óptimo tiene la siguiente forma

$$\Gamma_n[I_n] = J_{n,1} \cup \dots \cup J_{n,p-1} \cup [d_{n,p}^*(I_n), \infty), \quad (4.3.21)$$

y por lo tanto el momento que maximiza la funcional (4.1.1) está dado por $\tau_{opt} = \min\{0 \leq n \leq N : S_n \in \Gamma_n[I_n]\}$, donde el conjunto $\Gamma_N[I_N] = \mathbb{R}^+$.

Demostración. La función $g_n(x)$ es una función no negativa, no decreciente y convexa para $x \geq 0$ por definición, con esto se probará que $w_n(x, y)$ cumple las propiedades del Lema 4.2.2, es decir, para toda $n = 0, 1, \dots, N$: (a) $w_n(x, y) < \infty$, para toda $x \geq 0$, $y \in Y$, (b) $w_n(x, y)$ es continua, $x \geq 0$, $y \in Y$ y (c) $w_n(x, y) \in V$.

Primero se mostrará que $w_n(x, y) \in V$, es decir, $w_n(x, y)$ es medible, no decreciente y convexa para toda $x \geq 0$ y $y \in Y$, lo cual se hará por inducción.

1. Por definición $w_0(x, y) = g_N(x)$ la cual es medibles, no decreciente y convexa para toda $x \geq 0$ y $y \in Y$, entonces $w_0(x, y) \in V$.
Como $w_1(x, y) = \max\{g_{N-1}(x), e^{-R_{N-1}}T_{N-1}w_0(x, y)\}$ y $T_{N-k}w_0(x, y) \in V$ por Lema 4.2.1, entonces $w_1(x, y) \in V$ al ser el máximo de dos funciones en V .
2. Suponga que $w_n(x, y) \in V$.
3. Entonces $w_{n+1}(x, y) = \max\{g_{N-n-1}(x), e^{-R_{N-n-1}}T_{N-n-1}w_n(x, y)\} \in V$ por ser el máximo de dos funciones en V .

Así que $w_n(x, y) \in V$ para toda $n = 0, 1, \dots, N$.

Como $w_n(x, y) \in V$, entonces $w_n(x, y)$ es continua en $x \in \mathbb{R}^+$, $y \in Y$, debido a que es no decreciente y convexa; ahora sólo falta probar que $w_n(x, y) < \infty$, para toda $x \geq 0$, $y \in Y$; para esto sea $x \geq 0$ y $y \in Y$, entonces existe $i = 0, 1, \dots, p$ tal que $x \in [K_{n,i}, K_{n,i+1}]$ y como se vio en el Lema 4.2.2 si se considera a la función de pago $g_n^i(x) = a_{n,i}[x - K_{n,i}]^+$ se tiene que $w_n(x, y) < \infty$ para toda $x \geq 0$ y $y \in Y$, en particular para $x \in [K_{n,i}, K_{n,i+1}]$, entonces $w_n(x, y) < \infty$ con la función de pago $g_n(x)$.

A continuación se demostrará que $\varphi_n(x, y) = e^{-R_n}T_n w_{N-n-1}(x, y) - g_n(x)$, $n = 0, 1, \dots, N - 1$ posee propiedades similares a las enunciadas en el lema 4.2.4. Por demostrar que para toda $n = 0, 1, \dots, N - 1$:

- a. $\varphi_n(x, y) < \infty$, $x \geq 0$, $y \in Y$;
- b. $\varphi_n(x, y)$ es continua para toda $x \geq 0$, $y \in Y$;
- c. $\varphi_n(x, y)$ es no negativa, no decreciente y convexa como función de x en el intervalo $[0, K_{n,1}]$ para toda $y \in Y$;
- d. $\varphi_n(x, y)$ es convexa como función de x en el intervalo $[K_{n,i}, K_{n,i+1}]$ para toda $i = 1, \dots, p - 1$;
- e. $\varphi_n(x, y)$ es monótona decreciente y convexa como función de x en el intervalo $[k_{n,p}, \infty)$ para toda $y \in Y$;
- f. existe $\lim_{x \rightarrow \infty} \frac{1}{x} \varphi_n(x, y) < 0$ para toda $y \in Y$.

Para ver la demostración de estas propiedades sólo es necesario ver la forma de adaptar el Lema 4.2.4 en este caso; esto quiere decir que veamos a $\varphi_n(x, y) = e^{-R_n}T_n w_{N-n-1}(x, y) - g_n(x)$ como $\varphi_n^i(x, y) = e^{-R_n}T_n w_{N-n-1}^i(x, y) - g_n^i(x)$, donde $w_{N-n-1}^i(x, y)$ y $g_n^i(x)$ son las funciones de pago y el proceso de ganancias validos en el

intervalo $[K_{n,i}, K_{n,i+1}]$, pero si las consideramos a cada una para toda $x \geq 0$ se estará en las condiciones del Lema 4.2.4.

Así que $\varphi_n(x, y)^i < \infty$, $x \geq 0$, $y \in Y$, en particular $\varphi_n^i(x, y) < \infty$, $x \in [K_{n,i}, K_{n,i+1}]$, $y \in Y, i = 0, \dots, p$ con $K_{n,0} = 0$ y $K_{n,p+1} = \infty$, por lo tanto $\varphi_n(x, y) < \infty$, $x \geq 0$, $y \in Y$.

La propiedad **b** es más simple puesto que ${}_n w_{N-n-1}(x, y)$ y $g_n(x)$ son continuas, entonces como $\varphi_n(x, y)$ es la diferencia de dos funciones continuas es continua para toda $x \geq 0$, $y \in Y$.

Se sabe que $\varphi_n^0(x, y) = e^{-R_n T_n} w_{N-n-1}^0(x, y) - g_n^0(x)$, entonces $\varphi_n^0(x, y)$ es no negativa, no decreciente y convexa como función de x en el intervalo $[0, K_{n,1}]$ para toda $y \in Y$, por lo tanto $\varphi_n(x, y)$ también lo es en $[0, K_{n,1}]$ para toda $y \in Y$.

$\varphi_n^i(x, y)$ es convexa para toda $x \geq 0$, esto quiere decir que en particular para $[K_{n,i}, K_{n,i+1}]$ entonces $\varphi_n(x, y)$ es convexa como función de $x \in [K_{n,i}, K_{n,i+1}]$ para toda $i = 1, \dots, p-1$.

En el intervalo $[k_{n,p}, \infty)$ la función $\varphi_n^p(x, y) = e^{-R_n T_n} w_{N-n-1}^p(x, y) - g_n^p(x)$ es monótona decreciente y convexa como función de x entonces $\varphi_n(x, y)$ también lo es.

El límite de la propiedad **f** pide observar que pasa cuando x tiende a infinito, basta con ver como se comporta la función $\varphi_n^p(x, y)$ ya que modela los valores grandes de x , pero esta ya cumple la propiedad, por lo tanto $\varphi_n(x, y)$ también.

Para cada $g_n^i(x)$ utilizando los recursos del Teorema 4.2.3 se puede resolver el problema para cada $i = 0, 1, \dots, p$ se obtiene un intervalo $[d_n^*(y), \infty)$ que es el dominio de paro óptimo en cada caso, entonces este intervalo aportará con el conjunto $J_{n,i}(y) = [d_n^*(y), \infty) \cap [K_{n,i}, K_{n,i+1}]$ al dominio de paro óptimo del caso de la función $g_n(x)$. Es por eso que al dominio de paro óptimo se define como

$$\begin{aligned} \Gamma_n[y] &= \{x : g_n(x) = w_{N-n}(x, y)\} = \\ &\{x \geq K_{n,1} : \varphi_n(x, y) \leq 0\} = \bigcup_{i=1}^p J_{n,i}(y), \quad y \in Y \end{aligned}$$

de manera similar a cada conjunto $J_{n,i}(y)$ se le puede ver como $J_{n,i}(y) = \{K_{n,i} \leq x \leq K_{n,i+1} : \varphi_n(x, y) \leq 0\}$, $i = 1, \dots, p-1$ y $J_{n,p}(y) = \{K_{n,p} \leq x : \varphi_n(x, y) \leq 0\}$, el conjunto $J_{n,0}(y)$ es despreciado puesto que en $[0, K_{n,1})$ la función $g_n(x)$ vale cero. ■

En los casos donde las hipótesis se tienen las hipótesis de independencia de procesos índices (4.1.7) y (4.1.8) el dominio de paro óptimo también será independiente de todo proceso índice

$$\Gamma_n = \bigcup_{i=1}^p J_{n,i}.$$

Como se ha visto el resolver el problema de paro óptimo en el caso donde la función de pagos es lineal por trozos $g_n(x)$ es resolver el problema para cada función lineal implícita en $g_n(x)$, así que con el algoritmo que se presentará en el capítulo 5

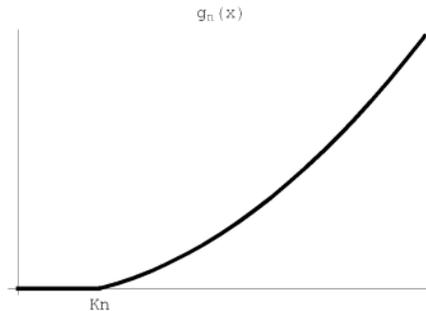
de ejemplos para funciones lineales se puede utilizar para resolver el problema de los dominios de paro óptimo con funciones de pago lineales por trozo.

4.4 Optimización en el caso de función de pago convexa en general

En esta sección se analiza que sucede con los dominios de paro óptimo si la función de pago, $g_n(x) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, no es lineal por trozos pero cumple con las siguientes condiciones.

E: Para toda $n \in \mathbb{N}$:

- (a) $g_n(0) = 0$,
- (b) $g_n(x)$ no es idénticamente cero, y
- (c) $g_n(x)$ es convexa para $x \geq 0$.



La estructura de los dominios de paro óptimo para una función convexa general pueden ser demasiado complejos, es por esto, que se busca encontrar una aproximación de los dominios de paro óptimo, esta aproximación se hará a través de la solución del problema con funciones lineal por trozos.

Debido a que $g_n(x)$ es convexa, con valor inicial cero y no es idénticamente cero, esta debe ser continua y estrictamente creciente para alguna $K_n \geq 0$, es decir, $g_n(x) = 0$ si $x \in [0, K_n]$, y $g_n(x) > 0$ y creciente si $x > K_n$.

Como $g_n(x)$ es completamente continua, para toda $x \geq 0$, se pueden hablar de sus derivadas unilaterales. Se denotará por $g'_n(x)$ a la derivada derecha de $g_n(x)$, que siempre existe porque $g_n(x)$ es continua, además $g'_n(x) = 0$ si $x \in [0, K_n)$, $g'_n(K_n) \geq 0$ y $g'_n(K_n) > 0$ si $x > K_n$, estas dos últimas características se deben a la monotonía de $g_n(x)$.

Como la función de pago $g_n(x)$ puede tener derivadas no acotadas, primer lugar se estudiara el caso cuando estas derivadas si son acotadas, es decir, que cumplen la siguiente condición.

F: Para toda $n \in \mathbb{N}$: $G_n = \sup_{x \geq 0} g'_n(x) < \infty$.

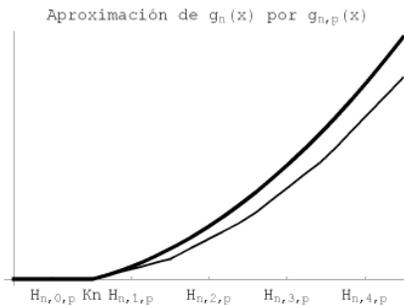
Bajo las condiciones **E** y **F** se construye una aproximación para $g_n(x)$ por medio de funciones lineales por trozos.

Sea $G = \max_{1 \leq n \leq N} G_n$, para alguna $p \in \mathbb{N}$ considere los conjuntos $H_{n,m,p} = \{x \geq 0 : \frac{m}{p}G \leq g'_n(x) < \frac{m+1}{p}G\}$ para $0 \leq m \leq p-1$ y $H_{n,p,p} = \{x \geq 0 : g'_n(x) \geq G\}$. Debido a la monotonía de la derivada de $g_n(x)$ estos conjuntos son intervalos ajenos, situados en \mathbb{R}^+ con un orden secuencial de acuerdo a el índice m , es decir, si $m' < m_j$ el intervalo $H_{n,m',p}$ quedará del lado izquierdo de $H_{n,m'',p}$, y la unión de estos conjuntos es una partición de \mathbb{R}^+ .

En cada conjunto $H_{n,m,p}$ se aproxima a la derivada $g'_n(x)$ por medio de la función $g'_{n,p}(x) = \sum_{m=0}^p (\frac{mG}{p} * I_{\{x \in H_{n,m,p}\}})$, para aproximar a la función de pago $g_n(x)$ por medio de la función

$$g_{n,p}(x) = \int_0^x g'_{n,p}(y) dy, \quad x \geq 0.$$

Se ha construido la aproximación $g_{n,p}(x)$ para la función $g_n(x)$, la función $g_{n,p}(x)$ siempre es menor igual a $g_n(x)$, además de ser lineal por trozos y convexa.



La distancia que existe entre las dos funciones esta dada por:

$$0 \leq g_n(x) - g_{n,p}(x) \leq \frac{G}{p} [x - K_n]^+, \text{ si } x \geq 0. \tag{4.4.22}$$

que se desprende de la relación que guarda $g'_n(x)$ con respecto a $g'_{n,p}(x)$ en los conjuntos $H_{n,m,p}$, ya que si $x \in H_{n,m,p}$ se tiene que $g'_n(x) - g'_{n,p}(x) \leq \frac{G}{p}$ para toda $m = 0, 1, \dots, p$, integrando la desigualdad

$$0 \leq g_n(x) - g_{n,p}(x) = \int_0^x g'_n(y) dy - \int_0^x g'_{n,p}(y) dy \leq \frac{G}{p} x,$$

pero la función $g_n(x) = 0$ si $x \in [0, K_n]$, entonces $\int_0^x g'_n(y)dy = 0 = \int_0^x g'_{n,p}(y)dy$ si $x \in [0, K_n]$, por lo tanto se puede describir la desigualdad como

$$\begin{aligned} 0 \leq g_n(x) - g_{n,p}(x) &= \int_{K_n}^x g'_n(y)dy - \int_{K_n}^x g'_{n,p}(y)dy \leq \frac{G}{p}(x - K_n), \text{ si } x \geq K_n, \text{ y} \\ 0 \leq g_n(x) - g_{n,p}(x) &\leq 0, \text{ si } x < K_n, \end{aligned}$$

para concluir que $0 \leq g_n(x) - g_{n,p}(x) \leq \frac{G}{p}[x - K_n]^+$, si $x \geq 0$.

Para la función $g_{n,p}(x)$ se pueden encontrar sus dominios de paro óptimo por medio de los teoremas 4.2.3 y 4.3.1. Para la función de pago $g_n(x)$ el encontrar sus dominios de paro es muy complicado debido a que no se tiene mucha información sobre ella, es por esto que se propone el calcular dominios de paro ε -óptimo y de hecho son los dominios de paro de la función $g_{n,p}(x)$, es decir, si τ_ε se encuentra en el dominio de paro ε -óptimo la diferencia entre el valor de $\phi_g(\tau_\varepsilon)$ y $\phi_g(\tau_{opt})$ es menor o igual a una cierta $\varepsilon > 0$.

La ε que produce el tiempo de paro ε -óptimo τ_ε está acotada por

$$\varepsilon_p = \sup_{\tau \leq N} E[e^{-R\tau} \frac{G}{p}[S_\tau - K_\tau]^+], \quad (4.4.23)$$

que será demostrada en el teorema 4.4.1.

Observe que para encontrar el valor de ε_p se pueden utilizar los Teorema 4.2.3 y 4.3.1, ya que se ha regresado al caso de que la función de pago está dada por $g_n^*(x) = G/p[x - K_n]^+$. Con respecto a las condiciones **B** (**B'**) y **C** (**C'**) no requieren de ningún cambio puesto que no dependen de la forma de $g_n(x)$, en la condición **D** (**D'**) las constantes a_n deben ser remplazadas por G/p .

Bajo las condiciones **A-D** el tamaño de ε_p puede ser tan pequeño como se desee, esto sólo depende de una elección apropiada del parámetro p .

Denótese por $\tau[p]$ al tiempo de paro óptimo para la función convexa lineal por trozos $g_{n,p}(x)$ que aproxima a la función $g_n(x)$ de acuerdo a la construcción de las aproximaciones, el valor supremo o máximo del caso $g_{n,p}(x)$ es:

$$\phi_{g_p}(\tau[p]) = E[e^{-R\tau[p]} g_{\tau[p],p}(S_{\tau[p]})] = \sup_{\tau \leq N} E[e^{-R\tau} g_{\tau,p}(S_\tau)]. \quad (4.4.24)$$

Por medio de los teoremas 4.2.3 y 4.3.1 se puede encontrar el tiempo de paro óptimo $\tau[p]$ y su respectivo dominio de paro óptimo, así como el valor máximo que toma la ecuación (4.4.24).

En resumen las condiciones necesarias para poder resolver este problema son **A** (**A'**), **B** (**B'**) y **C** (**C'**) que no requieren de ningún cambio puesto que no depende de la forma de $g_n(x)$; en **D** (**D'**) las constantes a_n pueden ser remplazadas por G/p . Las condiciones **A**, **B**, **D** y **F** quieren decir que $w_n(x, y) < \infty$ para toda $x \geq 0$, $y \in Y$ y

$n = 0, 1, \dots, N$, por lo tanto $\phi_{g_p}(\tau_{opt}) < \infty$, esto último sólo es valido para las funciones de aproximación $g_{n,p}(x)$.

Con estas observaciones se enuncia el Teorema de paro óptimo para la clase de funciones $g_n(x)$ convexas.

Teorema 4.4.1 *Bajo las condiciones A-F. El momento de Markov $\tau[p]$ es un momento de paro ε -óptimo para la optimización del problema (4.1.1)-(4.1.2) en el sentido de que*

$$\phi_g(\tau[p]) \geq \sup_{\tau \leq N} \phi_g(\tau) - \varepsilon_p.$$

Demostración. Por la desigualdad (4.4.22) se tiene que

$$0 \leq g_n(x_n) - g_{n,p}(x_n) \leq \frac{G}{p}[x - K_n]^+, \text{ para toda } n \leq N$$

entonces al tomar el valor esperado de ambos lados se llega a

$$0 \leq E [e^{-R\tau} g_n(x_n) - e^{-R\tau} g_{n,p}(x_n)] \leq E \left[e^{-R\tau} \frac{G}{p} [x_n - K_n]^+ \right]$$

de donde sabemos que $\phi_g(n) = E [e^{-R\tau} g_n(x_n)]$ y $\phi_{g_p}(n) = E [e^{-R\tau} g_{n,p}(x_n)]$, y sustituyendo esto se tiene que

$$0 \leq \phi_g(n) - \phi_{g_p}(n) \leq E \left[e^{-R\tau} \frac{G}{p} [x_n - K_n]^+ \right]$$

ahora al tomar los supremos sobre los tiempos de paro

$$0 \leq \sup_{\tau \leq N} \phi_g(n) - \sup_{\tau \leq N} \phi_{g_p}(n) \leq \sup_{\tau \leq N} \left[E \left[e^{-R\tau} \frac{G}{p} [x - K_n]^+ \right] \right]$$

donde se define a $\varepsilon_p = \sup_{\tau \leq N} \left[E \left[e^{-R\tau} \frac{G}{p} [x - K_n]^+ \right] \right]$; por lo tanto

$$0 \leq \sup_{\tau \leq N} \phi_g(\tau) - \varepsilon_p \leq \phi_g(\tau[p]).$$

■

Para concluir este análisis consideremos el caso donde la función de pagos $g_n(x)$ tiene una derivada no acotada, esto quiere decir que se tiene la siguiente condición.

$$\bar{\mathbf{F}}: \lim_{x \rightarrow \infty} g'_n(x) = +\infty, \quad n = 0, 1, \dots$$

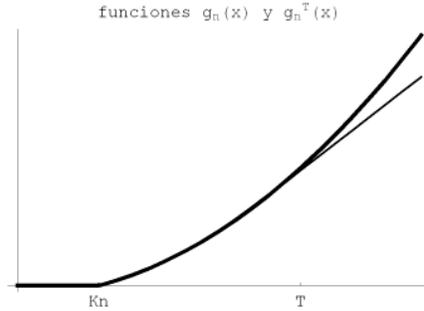
Análogo a cuando se tenían derivadas acotadas se busca poder aproximar a la función $g_n(x)$ por medio de funciones convexas y lineales por trozos.

4. Optimización de opciones Americanas, con función de pago convexa, a tiempo discreto **60**

Para poder aproximar a la función $g_n(x)$ se debe elegir $T > \max_{0 \leq n \leq N} K_n$ (el máximo de las x donde las $g_n(x)$ se vuelven estrictamente mayores que cero) y considerar la función

$$g_n^T(x) = \begin{cases} g_n(x), & \text{si } 0 \leq x \leq T \\ g_n(T) + g'_n(T) \cdot (x - T), & \text{si } x > T \end{cases}$$

La idea de construir $g_n^T(x)$ es suavizar el crecimiento de $g_n(x)$ para poder construir la aproximación por medio de funciones lineales por trozos como se hizo anteriormente. Cuando $x > T$ la función toma el valor $g_n(T) + g'_n(T) \cdot (x - T)$, es decir, que a partir de ese momento crecerá de forma lineal con una pendiente constante $g'_n(T)$.



La función $g_n^T(x)$ satisface las condiciones **E** y **F**, entonces puede ser aproximada por medio de una función $g_{n,p}^T(x)$, una aproximación lineal por trozos de acuerdo al descrito anteriormente. En este caso se tiene la constante $G = \max_{0 \leq n \leq N} g'_n(T)$.

La cota que permite medir la eficiencia de la aproximación de los tiempo de paro ε -óptimos está dada por

$$\varepsilon_{p,T} = \sum_{n=0}^N e^{-R_n} E(g_n - g_n^T)(S_n) + \sup_{\tau \leq N} E \left[e^{-R_\tau} \frac{G}{p} [S_\tau - K_\tau]_+ \right].$$

Entonces por medio de los Teoremas 4.2.3 y 4.3.1 se puede encontrar el tiempo de paro óptimo y valor máximo del segundo termino del lado derecho de la ecuación (??). Para poder utilizar los Teoremas 4.2.3 y 4.3.1 las condiciones **A(A')**, **B(B')** y **C(C')** no requieren de cambios, en la condición **D(D')** las constantes a_n deben ser remplazadas por G/p .

Como la eficiencia de la aproximación depende de $\varepsilon_{p,T}$ es deseable que $\lim_{T \rightarrow \infty} E(g_n - g_n^T)(S_n) = 0$ para toda $n = 0, 1, \dots, N$, por lo que se debe pedir una condición más.

G: $E_{x,y}^0 g_n(S_n) < \infty, (x, y) \in Z, n = 0, 1, \dots, N.$

Si las condiciones **A-G** se cumplen $\varepsilon_{p,T}$ puede ser tan pequeña como se desee, siempre que se tenga una buena elección de T y luego de p .

Denótese por $\tau[p, T]$ el momento de paro óptimo para la función de pagos convexa lineal por trozos $g_{n,p}^T(x)$, esto quiere decir que:

$$E \left[e^{-R_{\tau[p, T]}} g_{\tau[p, T], p}^T(S_{\tau[p, T]}) \right] = \sup_{\tau \leq N} E \left[e^{-R_{\tau}} g_{\tau, p}^T(S_{\tau}) \right].$$

El tiempo de paro $\tau[p, T]$ puede ser encontrado por medio de los Teoremas 4.2.3 y 4.3.1, ya que se puede obtener sus dominios de paro óptimo.

En resumen las condiciones que deben existir para que se pueda construir la aproximación para una T y p dadas son: las condiciones **B** (**B'**) y **C** (**C'**) sin ningún cambio y en **D** (**D'**) remplazar las constantes a_n por G/p .

Las condiciones **A**, **B**, **D** y **G** implican que $w_n(x, y) < \infty$ para toda $x \geq 0$, $y \in Y$ y $n = 0, 1, \dots, N$, por lo tanto $\phi_g(\tau_{opt}) < \infty$, y esto también es valido para las funciones de aproximación $g_{n,p}^T(x)$.

Teorema 4.4.2 *Bajo las condiciones **A-E**, $\bar{\mathbf{F}}$ y **G**. El momento de Markov $\tau[p, T]$ es un momento de paro ε -óptimo para la optimización del problema (4.1.1) – (4.1.2) en el sentido de que*

$$\phi_g(\tau[p, T]) \geq \sup_{\tau \leq N} \phi_g(\tau) - \varepsilon_{p, T}.$$

Demostración. Sea $\tau \leq N$ un tiempo de paro, entonces

$$\begin{aligned} & E \left[e^{-R_{\tau}} g_{\tau}(S_{\tau}) \right] - E \left[e^{-R_{\tau}} g_{\tau, p}^T(S_{\tau}) \right] \leq \\ & \left\{ E \left[e^{-R_{\tau}} g_{\tau}(S_{\tau}) \right] - E \left[e^{-R_{\tau}} g_{\tau}^T(S_{\tau}) \right] \right\} + \left\{ E \left[e^{-R_{\tau}} g_{\tau}^T(S_{\tau}) \right] - E \left[e^{-R_{\tau}} g_{\tau, p}^T(S_{\tau}) \right] \right\} \end{aligned}$$

por el Teorema 4.4.1 se sabe que $0 \leq E \left[e^{-R_{\tau}} g_{\tau}^T(S_{\tau}) \right] - E \left[e^{-R_{\tau}} g_{\tau, p}^T(S_{\tau}) \right]$ y existe su supremo sobre los tiempos de paro, para acotar la primera parte simplemente se usará la suma sobre todos los tiempos de paro, entonces

$$\begin{aligned} & \sup_{\tau \leq N} E \left[e^{-R_{\tau}} g_{\tau}(S_{\tau}) \right] - \sup_{\tau \leq N} E \left[e^{-R_{\tau}} g_{\tau, p}^T(S_{\tau}) \right] \leq \\ & \sum_{n=0}^N E \left[e^{-R_n} g_n(S_n) \right] - E \left[e^{-R_n} g_n^T(S_n) \right] + \sup_{\tau \leq N} \left\{ E \left[e^{-R_{\tau}} g_{\tau}^T(S_{\tau}) \right] - E \left[e^{-R_{\tau}} g_{\tau, p}^T(S_{\tau}) \right] \right\} = \\ & \varepsilon_{p, T} \end{aligned}$$

por lo tanto

$$\sup_{\tau \leq N} \phi_g(\tau) - \phi_g(\tau[p, T]) \leq \varepsilon_{p, T}.$$

■

Al demostrar este último teorema se concluye con el análisis de los dominios de paro óptimo cuando se tiene una función de pago convexa.

En el siguiente capítulo se darán ejemplos de la aplicación de la teoría expuesta.

Capítulo 5

Ejemplos

En los capítulos anteriores se ha desarrollado la teoría en la cual se sustenta la existencia de los dominios de paro óptimo para instrumentos financieros con función de pago convexa. En el Teorema 4.2.3 se observa que para encontrar los dominios de paro óptimo basta con resolver la ecuación

$$g_n(x) = e^{-r_n} T_n w_{N-n-1}(x, I_n) \quad (5.0.1)$$

para cada $n \in \{0, 1, \dots, N-1\}$, y los dominios estarán dados por los subconjunto $[d_n, \infty)$ de \mathbb{R}^+ , donde d_n son las raíces de dicha ecuación.

Resolver la ecuación 5.0.1 no es trivial, debido a que la definición del proceso de rendimientos $w_n(x, y)$ está dado como:

$$\begin{aligned} w_0(x, y) &= g_N(x), y \\ w_k(x, y) &= \max \{g_{N-k}(x), e^{-r_{N-k}} T_{N-k} w_{k-1}(x, y)\}, \text{ para } k = 1, 2, \dots, N. \end{aligned} \quad (5.0.2)$$

Note que cuando n empieza a crecer el cálculo del proceso de rendimientos $w_n(x, y)$ se complica puesto que se deben calcular $n-1$ esperanzas condicionales para su cálculo, la idea de este capítulo es dar una forma alternativa para solucionar este problema. La idea está basada en el método de simulación Monte Carlo, el cual consiste en simular un número grande de veces el proceso que se estudia para obtener una muestra grande de datos que permitan hacer inferencia sobre el valor esperado del proceso.

En la siguiente sección se dará una descripción del algoritmo para la aproximación de los dominios de paro óptimo, el cual fue programado para ilustrar la teoría descrita por medio de ejemplos numéricos en la sección 5.2.

5.1 Elaboración del algoritmo para aproximar los dominios de paro óptimo

A continuación se describe el algoritmo con el cual se crearon los ejemplos de la sección 5. El algoritmo está basado en los artículos *Algorithms and Programs for optimal Monte Carlo pricing of American options* [9] y *Monte Carlo studies of american type call options with discrete time* [10].

La aproximación de los dominios de paro óptimo se hará en el caso descrito en el capítulo 4.2, es decir, cuando se tiene una opción americana de compra estándar, es decir, que se tiene una función de pago del tipo $g_n(x) = [x - K]^+$, para toda $n = 0, 1, \dots, N$. Para poder hacer el análisis se debe conocer el proceso que describe el precio del activo subyacente, $\{S_n\}_{n \in \mathbb{N}}$, el se supondrá como una caminata aleatoria geométrica, entonces el proceso $\{S_n\}_{n \in \mathbb{N}}$ puede ser escrito como:

$$S_n = S_{n-1}U_n, \quad n \in \mathbb{N},$$

donde $S_0 = S$ es constante y $\{U_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas que toman valores en los reales positivos.

Como se está trabajando con una opción americana estándar de compra la función de pagos es $g(x) = [x - K]^+$, donde K es el precio pactado y constante. En lugar de manejar las tasas libres de riesgo diferentes para cada periodo se utiliza una constante de descuento $v = e^{-rn}$.

Por el Teorema 4.2.3 para cada tiempo $n \in \{1, \dots, N - 1\}$ existe un valor $d_n \in \mathbb{R}^+$ tal que el dominio de paro óptimo toma la forma de un intervalo

$$\Gamma_n = [d_n, \infty),$$

para encontrar el valor d_n es necesario comparar los valores de las funciones $g(x_n)$ y $vT_n\omega_{N-n-1}(x)$, ya que d_n es la raíz de la ecuación $g(d_n) = vT_n\omega_{N-n-1}(d_n)$, debe de recordar que todo esto es cierto siempre y cuando se cumplan las propiedades **A-D** descritas en el capítulo 4.2.

Para poder elaborar la aproximación se debe fijar un intervalo donde se buscarán las d_n , denótese por $[S_L, S_U]$ a dicho intervalo. En el intervalo se debe asegurar que se encuentra el valor de d_n , para toda $n \in \{1, \dots, N - 1\}$, para poder asegurar que d_n se encontrará en dicho intervalo es conveniente tomar como valor inicial $S_L = K$, ya que antes de K el valor de ganancias es cero, para proponer un valor superior S_U no existe información, por lo que este debe considerarse de acuerdo a la experiencia o pruebas previas del algoritmo. El intervalo $[S_L, S_U]$ se particiona para obtener una sucesión de

precios, la partición se hace de acuerdo a la siguiente forma recursiva,

$$\begin{aligned} s_0 &= S_L \\ s_i &= s_{i-1} + i \cdot \Delta, \text{ para } i = 1, 2, \dots, J - 1 \end{aligned}$$

donde Δ es el incremento entre la sucesión de precios $\{s_i\}_{i \in \{0, 1, \dots, J-1\}}$ y $S_J = S_U$. La aproximación al verdadero valor de d_n depende en gran medida del tamaño de Δ , aunque esta no es la única variable involucrada en la precisión del algoritmo, mientras más pequeña sea Δ más precisa será la aproximación de los valores d_n , el problema de tomar una Δ muy pequeña es que obtiene un programa que tarda demasiado tiempo en ser procesado, por eso debe elegirse una Δ de tamaño adecuado que dependerá de los datos iniciales del problema, el comportamiento que se espera tener en el proceso y la precisión deseada.

Como el proceso es una cadena de Markov se puede utilizar recurrencias hacia atrás para resolver el problema, entonces el primer paso es encontrar el último valor d_n , el cual ya es conocido debido a que en la fecha de expiración $d_N = 0$. En el día $N - 1$ se debe calcular el valor d_{N-1} , para ello se aproxima el valor de $T_{N-1}[w_0(x)]$ para cada $x \in \{s_0, s_1, \dots, s_J\}$, para poder hacer esta aproximación se requiere de M valores de la variable aleatoria del incremento del día $N - 1$ al N , es decir, M realizaciones de la variable aleatoria U_N . A los M valores de realizados de la variable U_N se les denotará por $U_N^1, U_N^2, \dots, U_N^M$. Entonces para cada valor s_j la aproximación de $T_{N-1}[w_0(s_j)]$ está dada por

$$\hat{T}_{N-1}[w_0(s_j)] = \frac{1}{M} \sum_{i=1}^M vg(s_j * U_N^i),$$

de acuerdo al Teorema 4.2.3 existe un valor $d_{N-1} \in \mathbb{R}^+$ tal que $g(d_{N-1}) = T_{N-1}[w_0(d_{N-1})]$, $g(x) < T_{N-1}[w_0(x)]$ si $x \leq d_{N-1}$ y $g(x) > T_{N-1}[w_0(x)]$ si $x \geq d_{N-1}$. Para encontrar el valor d_n se comparan los valores $g(s_j)$ y $\hat{T}_{N-1}[w_0(s_j)]$, y la aproximación de d_{N-1} será $\hat{d}_{N-1} = \max \{s_j : g(s_j) < \hat{T}_{N-1}[w_0(s_j)]\}$.

La aproximación del valor d_{N-2} tiene una pequeña variación, esto se debe a que no se sabe si se ejercerá la opción en el tiempo $N - 1$ o N . De nuevo se debe realizar M veces a la variable U_{N-1} , para tener los datos $U_{N-1}^1, U_{N-1}^2, \dots, U_{N-1}^M$.

Al tiempo $N - 2$ el precio toma el valor x , si xU_{N-1} es mayor al valor \hat{d}_{N-1} la opción debe ser ejercida en el día $N - 1$ puesto que el precio a entrado en el dominio de paro óptimo, entonces la opción termina. En cambio, si xU_{N-1} es menor a \hat{d}_{N-1} la opción no debe ser ejercida puesto que no es un precio óptimo, entonces el precio tiene que ser simulado al tiempo N por $xU_{N-1}U_N$ donde será ejercida a este precio, puesto que en la fecha de vencimiento todo precio es óptimo. Entonces la aproximación de

$T_{N-2}[w_1(s_j)]$ está dada por

$$\begin{aligned} \widehat{T}_{N-2}[w_1(s_j)] &= \frac{1}{M} \sum_{i=1}^M \{vg(s_j * U_{N-1}^j) * \mathbb{I}(s_j * U_{N-1}^j)_{(s_j * U_{N-1}^j \geq d_{N-1})} + \\ &\quad v^2 g(s_j * U_{N-1}^j * U_N^j) * \mathbb{I}(s_j * U_{N-1}^j)_{(s_j * U_M^j < d_{N-1})}\}, \end{aligned}$$

donde $\mathbb{I}(s_j * U_{N-1}^j)_{(s_j * U_{N-1}^j \geq d_{N-1})}$ es la función indicadora de que el precio al tiempo $N - 1$ se encuentra en el dominio del paro óptimo y $\mathbb{I}(s_j * U_{N-1}^j)_{(s_j * U_M^j < d_{N-1})}$ la función indicadora de que el precio al tiempo $N - 1$ no se encuentre en el dominio de paro óptimo. Una vez que se ha aproximado a $T_{N-2}[w_1(s_j)]$ por $\widehat{T}_{N-2}[w_1(s_j)]$, el valor d_{N-2} es aproximado por $\widehat{d}_{N-2} = \max \{s_j : g(s_j) < \widehat{T}_{N-2}[w_1(s_j)]\}$.

De manera similar se procede para las aproximaciones $\widehat{d}_{N-3}, \widehat{d}_{N-4}, \dots, \widehat{d}_1$, teniendo presente que al calcular \widehat{d}_{N-n} se debe utilizar los dominios de paro óptimo para los tiempo $N - n + 1, N - n + 2, \dots, N$ generados por las aproximaciones $\widehat{d}_{N-1}, \widehat{d}_{N-2}, \dots, \widehat{d}_{N-n+1}$, ya que estos dominios son necesarios en el cálculo de $\widehat{T}_{N-n}[w_{N-n+1}(s_j)]$ para poder elegir el tiempo de paro $\widehat{d}_{N-n} = \max \{s_j : g(s_j) < \widehat{T}_{N-n}[w_{N-n+1}(s_j)]\}$.

5.2 Ejemplos numéricos

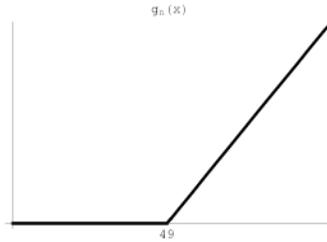
Con el algoritmo descrito han sido programado dos ejemplos en **Mathematica 5.0**, el programa usado se encuentra en el Apéndice A.

Ambos ejemplos suponen que el activo subyacente sigue una caminata aleatoria geométrica $\{X_n\}_{n \in \mathbb{N}}$, donde

$$X_n = X_{n-1}U_n$$

con la diferencia de que en el primer ejemplo 5.2.1 las variables aleatorias U_n se distribuyen *Bernoulli*(p, U, L), es decir, con probabilidad p la variable U_n toma el valor U y con probabilidad $(1 - p)$ toma el valor L ; en el segundo ejemplo 5.2.2 las variables aleatorias siguen una distribución *lognormal*(μ, σ).

En ambos ejemplos se supone que la opción americana es de compra con una función de pago $g_n(x) = [x - K]$, con precio pactado $K = 49$ y una duración de 10 días, $N = 10$; para efectos de valor presente se contempla una tasa de descuento anual del 7%.



Ejemplo 5.2.1 Bernoulli(0.5,1.1,0.9). El activo subyacente al cual se encuentra ligada la opción está descrito por una caminata aleatoria geométrica $\{X_n\}_{n \in \mathbb{N}}$, donde $X_n = X_{n-1}U_n$ y U_n se distribuye Bernoulli(0.5,1.1,0.9).

El intervalo donde se buscaron los dominios de paro óptimo fue $[49, 120]$ con incremento entre precios $\Delta = 1$, se propuso este intervalo porque el precio pactado es $K = 49$ y por simulaciones previas se observó que el dominio de paro óptimo se no salía de él.

Después de realizar 70 veces el algoritmo programado para la aproximación de los dominios de paro óptimo con muestra de tamaño $M=5000$, 10000 y 20000 se obtuvieron los siguientes estimadores.

Tabla 5.1:

M=5000	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
mín.	79	75	76	73	67	69	65	60	55	0
máx.	95	95	89	87	83	78	78	68	60	0
rango	16	20	13	14	16	9	13	8	5	0
media	85.87	83.16	82.1	78.07	75.66	72.96	67.86	61.76	55.83	0
varianza	10.95	18.66	7.95	7.78	13.85	3.64	3.95	3.29	1.25	0

Tabla 5.2:

M=10000	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
mín.	80	77	74	74	69	69	66	60	55	0
máx.	100	95	92	89	89	80	73	68	60	0
rango	20	18	18	15	20	11	7	8	5	0
media	88.24	86.37	83.16	80.5	78.37	74.04	68.11	61.93	56.11	0
varianza	15.26	14.24	11.61	12.28	10.06	3.81	2.16	2.15	1.75	0

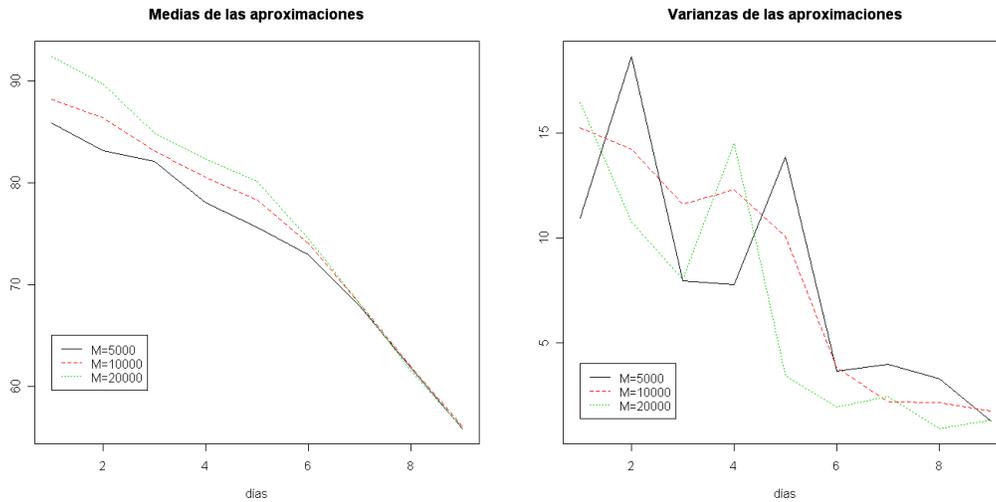
Tabla 5.3:

M=20000	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
mín.	84	79	77	75	75	71	66	60	55	0
máx.	101	98	92	91	84	78	76	65	60	0
rango	17	19	15	16	9	7	10	5	5	0
media	92.44	89.69	84.89	82.36	80.14	74.49	68.11	61.51	55.94	0
varianza	16.45	10.77	8.02	14.52	3.43	1.91	2.45	0.89	1.3	0

En las tablas anteriores se pueden observar los estimadores de las 70 realizaciones de las aproximaciones de los d_n para los diferentes valores de $M=5000$, 10000 y 20000. Como estimadores relevantes se encuentra el mínimo, máximo, rango cubierto, media

y varianza de las 70 realizaciones. Los datos obtenidos de las 70 realizaciones pueden ser consultados en el apéndice B.

A continuación se muestran las gráficas del comportamiento de la media y la varianza de los d_n 's aproximados.



Gráficamente se observa como los estimadores de la media tienen una convergencia en los días 7, 8 y 9, mientras el estimador se acerca al día del inicio de la opción la distancia entre los estimadores para los diferentes valores de M crece. Este comportamiento se debe a que se utilizó recursión para atrás en el algoritmo lo cual hace que exista una varianza mayor en las estimaciones de los primeros días de las d_n 's.

De acuerdo a la información obtenida de las simulaciones los mejores estimadores son los que se obtuvieron de las simulaciones realizadas con una tamaño de muestra $M=20000$, puesto que estos tienen una varianza menor en casi todo tiempo que los simulados con un valor de $M=5000$ o 10000 . Por estas razones los intervalos de paro óptimo que se proponen son:

Tabla 5.4:

día 1	día 2	día 3	día 4	día 5	día 6	día 7	día 8	día 9	día 10
[92.4, ∞)	[89.8, ∞)	[84.8, ∞)	[82.4, ∞)	[80.2, ∞)	[74.3, ∞)	[68, ∞)	[61.3, ∞)	[55.9, ∞)	[0, ∞)

Estos dominios son óptimos ya que al tiempo n permiten decidir si el precio del activo subyacente S_n ofrece la ganancia esperada máxima al poseedor de la opción si ejerce en este momento. Esta decisión es tomada al observar si el precio S_n se encuentra en el dominio de paro óptimo $[\hat{d}_n, \infty)$

Ejemplo 5.2.2 Lognormal(0,0.01). En el siguiente ejemplo se analiza la situación cuando el precio del activo subyacente al cual está vinculada la opción sigue una cami-

nata aleatoria $\{X_n\}_{n \in \mathbb{N}}$, donde $X_n = X_{n-1}U_n$ y las variables aleatorias U_n tienen una distribución lognormal(0,0.01).

Como ya se mencionó la función de pago de la opción es $g_n(x) = [x - K]^+$ para toda tiempo $n = 1, 2, \dots, N = 10$, donde $K = 49$ y la duración de la opción es de 10 días.

El intervalo que se analizó para estimar los valores de las d_n 's fue [49, 60], se utilizó este intervalo debido a que el precio pactado es 49 y por simulaciones de prueba se observó que las estimaciones se concentraban en el intervalo. La disminución que se presenta en la longitud del intervalo con respecto al ejemplo anterior se debe a que la variación de las variables lognormales(0,0.01) es menor a la que tienen las variables Bernoulli(0.5,1.1,0.9) usadas en el ejemplo anterior. A diferencia del ejemplo anterior esta vez se utilizó un incremento de precios de $\Delta = 0.1$ puesto que los valores que toman las variables U_n se concentran más al rededor del 1 que las variables Bernoulli(0.5,1.1,0.9), lo que produce cambios menores en los precios descritos por la caminata aleatoria geométrica lognormal(0,0,01).

A continuación se muestran los estimadores obtenidos después de haber realizado 70 simulaciones del programa de aproximación para las d_n 's.

Tabla 5.5:

M=5000	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
mín.	51.7	51.2	51.3	51.1	50.9	50.6	50.4	50.1	49.8	0
máx	53.4	53.2	52.5	52.9	52.5	52	51.9	51.4	51.9	0
rango	1.7	2	1.2	1.8	1.6	1.4	1.5	1.3	2.1	0
media	52.32	52.08	51.89	51.76	51.58	51.23	50.97	50.61	50.2	0
varianza	0.12	0.12	0.09	0.16	0.11	0.11	0.08	0.08	0.1	0

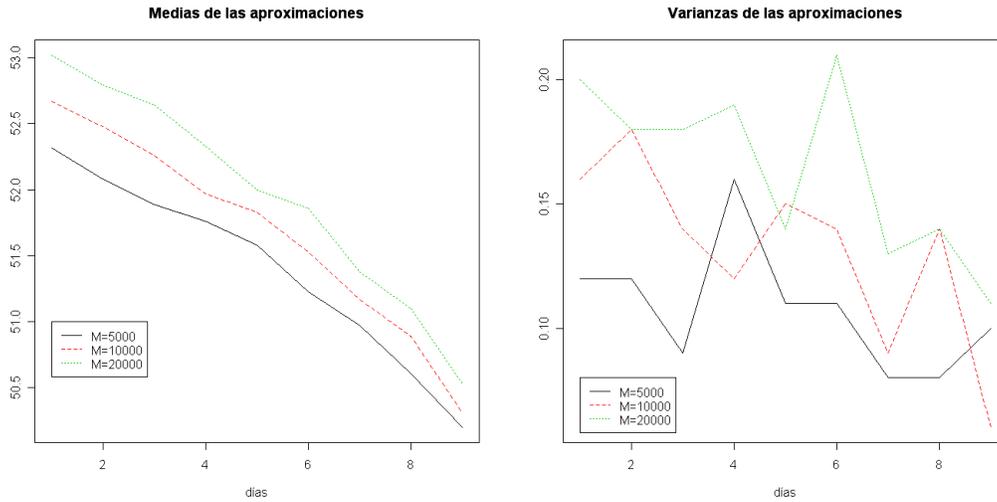
Tabla 5.6:

M=10000	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
mín.	51.7	51.8	51.6	51.4	51.1	50.7	50.5	50.4	49.9	0
máx	53.7	53.9	53.2	52.7	52.9	52.6	52	52.1	50.9	0
rango	2	2.1	1.6	1.3	1.8	1.9	1.5	1.7	1	0
media	52.67	52.48	52.26	51.97	51.83	51.53	51.17	50.89	50.31	0
varianza	0.16	0.18	0.14	0.12	0.15	0.14	0.09	0.14	0.06	0

Tabla 5.7:

M=20000	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
mín.	52.1	51.9	51.9	51.6	51.3	51.1	50.7	50.4	50.1	0
máx	53.9	54.2	53.7	54.1	52.9	52.9	52.3	52.1	51.6	0
rango	1.8	2.3	1.8	2.5	1.6	1.8	1.6	1.7	1.5	0
media	53.02	52.79	52.64	52.33	52	51.86	51.38	51.1	50.53	0
varianza	0.2	0.18	0.18	0.19	0.14	0.21	0.13	0.14	0.11	0

Como en el ejemplo anterior los estimadores utilizados para analizar el comportamiento de las aproximaciones de las d_n 's fueron el mínimo, máximo, rango, media y varianza de los aproximaciones simuladas.



De los datos obtenidos se observa que las medias de los estimadores de las d_n 's tienen un comportamiento ligeramente creciente conforme crece M , pero también se puede observar que la variación de las d_n 's disminuye con el crecimiento de M .

Esta característica que presentan las varianzas de disminuir con el crecimiento de M permite decir que al aumentar que el valor de las d_n 's se encuentra cercano a los estimadores encontrados.

Con las estimaciones obtenidas los dominios de paro óptimo que se proponen son los que se obtuvieron con el tamaño de muestra $M=20000$, debido a que presentan una varianza menor.

Tabla 5.8:

día 1	día 2	día 3	día 4	día 5	día 6	día 7	día 8	día 9	día 10
[53,∞)	[52.8,∞)	[52.6,∞)	[52.3,∞)	[52,∞)	[51.9,∞)	[51.4,∞)	[51.1,∞)	[50.5,∞)	[0,∞)

Capítulo 6

Conclusiones

La idea de porque desarrollar la teoría expuesta aquí surge por el deseo de optimizar opciones americanas de compra. Lo cual se hizo por medio de un teorema importante dentro de la teoría de optimización, teorema de paro óptimo de Shiryaev, el cual es sencillo en el sentido que pide muy pocas cosas para poder ser utilizado y sus resultados pueden ser simples de usar en el caso de tener un proceso de Markov. Estas características del teorema de paro óptimo son las que han permitido describir con más precisión los dominios de paro óptimo.

Aunque el problema de optimizar una opción americana es un problema continuo, se ha integrado una teoría que permite tener una aproximación al problema continuo por medio de discretizarlo, ya que como se observa si se desea resolver el problema en un intervalo $[0, T]$ continuo, la teoría aquí expuesta permite resolver el problema en una sucesión $x_0 = 0 < x_1 < \dots < x_{n-1} < x_n = T$ tan fina como se desee y se tengan los recursos tecnológicos necesarios para poder generar y aproximar los dominios de paro óptimo, con lo que se puede obtener una aproximación al problema continuo.

Como se muestra en el capítulo de ejemplos 5 la teoría expuesta tiene un problema al momento de aplicarse, ya que su solución exacta no es muy accesible debido a la forma en que se a construido el proceso de rendimientos $w_n(x, y)$. Debido a esto se ha recurrido a implementar un algoritmo para solucionar el problema numéricamente. En la actualidad los avances tecnológicos tienen un gran desarrollo lo que permite la creación y ejecución de los algoritmos aquí descritos en tiempos relativamente cortos, y de este modo tener una aproximación a la solución, la precisión de la aproximación dependerá de los recursos tecnológicos que se tengan y el tiempo disponible o que se desea invertir en calcular dichas aproximaciones. Recordando que la convergencia de las aproximaciones depende de diversos factores que son la distribución usada, la distancia en la partición del intervalo de análisis y el número de simulaciones utilizadas.

En la actualidad el desarrollo de la teoría de optimización a tenido un gran auge

en los métodos continuos que aunque teóricamente son precisos casi siempre tienen la condición de que al ser llevados a la realidad deben de ser aproximado por métodos numéricos que los discretizan para poder resolverlos, así que ante el desarrollo de los métodos continuos no se puede dejar de lado el análisis de problemas discretos que son una buena opción para aproximar problemas continuos.

Apéndice A

Algoritmo

En esta sección se describe el programa creado **Mathematica 5.0** usado en los ejemplos numéricos de la sección 5.2. Entre los dos ejemplos que se mostraron la única diferencia está dada por el tipo de distribución que siguen las variables aleatorias con las que se construyen las caminatas aleatorias que modelan el precio del activo subyacente.

En la sección de paquetes es suficiente con llamar al paquete que contiene la distribución a usar, en el caso de la distribución bernoulli el paquete a llamar es `<<Statistics'DiscreteDistributions'`, en el caso de la distribución lognormal el paquete es `<<Statistics'ContinuousDistributions'`.

■ PAQUETES USADOS

```
<<Statistics'ContinuousDistributions'  
<<Statistics'DiscreteDistributions'
```

A continuación se deben introducir los datos con los que ha de trabajar el programa, como: precio pactado (k), tasa anual de descuento (r), número de simulaciones a hacer (m), número de días de duración de la opción (n), el valor inicial y final del intervalo que se analizara (si y sf respectivamente), el nivel de distancia entre las particiones de los precios (h) y un número lo suficiente mente grande para distinguir cuando el programa no ha encontrado el dominio de paro óptimo debido a que este se encuentra fuera de $[si, sf]$ al que se le da el nombre de RANGOMAYOR.

■ DATOS INICIALES

```
{si,sf,h}={50,150,2};  
{k,r,RANGOMAYOR}={49,0.07,6666};  
{m,n}={100,10};  
rd=(1+r)1/360-1;
```

En esta sección se introduce la función de pago del instrumento que se analiza, el programa sólo a sido diseñado para una opción americana estándar, por lo que un

cambio esta función debe contemplar cambios en todo el programa.

■ FUNCIONES

$g[\theta_]:=Max[0,\theta-k];$

El siguiente paso es generar la partición del intervalo [si, sf], se introduce se introduce una corrección en la primera línea debido a que si $(sf-si)/h$ no es exacto, la partición no cubrirá a dicho intervalo y por eso se añaden un termino extra para asegurar esta cobertura.

■ PARTICIÓN DEL INTERVALO DE ANALISIS

$np=IntegerPart[\frac{sf-si}{h}]+2;$
 $x=Table[si,\{i,1,np\}];$
 $For[i=2,i\leq np,x[[i]]=x[[i-1]]+h;i++]$
 $gn=Table[g[x[[i]]],\{i,1,np\}];$

De las siguiente dos secciones del algoritmo sólo debe ser introducida la que corresponde a la distribución de las variables aleatorias U_n , la función berno en el caso de la distribución lognormal no da ninguna operación ya que se fija como la identidad, mientras que el caso de la distribución bernoulli es la que permite pasar de una Bernoulli(p) que toma valores 0 y 1 en la distribución Bernoulli(U,L,p) que toma valores U y L.

■ CASO BERNOULLI(U,L,p)

$\{U,L,p\}=\{1.1,0.9,0.5\};$
 $berno[\theta_]:=If[\theta==0,L,U];$
 $distribucionusada=BernoulliDistribution[p];$

■ CASO LOGNORMAL(U,L,p)

$\{\mu,\sigma\}=\{1.1,0.1\};$
 $berno[\theta_]=\theta;$
 $distribucionusada=LogNormalDistribution[\mu,\sigma];$

Para economizar espacio se define una función que genera los números aleatorios a usarse dentro de las iteraciones del programa en el calculo de los d_n 's, y para evitar una dependencia de los números aleatorios de la maquina se toma el tiempo de la maquina como valor inicial y se tiran 10000 muestras sin usar.

■ GENERADOR DE NÚMEROS ALEATORIOS

$ale[\theta_]=Table[berno[Random[distribucionusada]],\{i,1,\theta\}];$
 $ab[\theta_]=berno[Random[distribution]];$
 $SeedRandom[];$
 $semilla=Table[Random[],\{i,1,10000\}];$
 $ClearAll[semilla];$

La siguiente sección es la última parte del algoritmo, en esta se inicia con la definición de todos los dominios de paro óptimo en cero, para después hacer el calculo de

d_{n-1} por aparte de los d_n 's, esta diferencia se hace ya que el calculo de d_{n-1} no requiere de preguntas ni comparaciones, entonces al incluirlo en el calculo por grupo de los restantes d_n 's se estarían haciendo cálculos innecesarios que generan un costo de tiempo.

■ CALCULO DE LOS VALORES d_n

```

dn=Table[0,{i,1,np}]
tiempoi=Date[]
tw=Table[0,{i,1,np},
(*calculo de  $d_{n-1}$ *)]
For[i=1,i≤np,{
  rand=ale[m],
  For[j=1,j≤m,tw[[i]]=tw[[i]]+g[x[[i]]*rand[[j]];j++],
  tw[[i]]=v $\frac{tw[[i]]}{m}$ 
];i++]
cont=Table[0,{i,1,np}];
i=1;
While[i≤np,{
  If[gn[[i]]<tw[[i]],cont[[i]]=i,cont[[i]]=0],
  If[cont[[i]]==0,i=np,0],
  i++
}]
If[Max[cont]+1>np,dn[[n-1]]=RANGOMAYOR,dn[[n-1]]=x[[Max[cont]+1]];
(*Calculo de los valores  $d_{n-2},d_{n-3},\dots,d_2,d_1$ *)
For[fecha=n-2,fecha≤1,{
  tw=Table[0,{i,1,np}];
  For[i=1,i≤np,{
    rand=ale[m],
    For[j=1,j≤m,{
      xn=x[[i]]*rand[[j]],
      fn=0,
      dia=fecha+1,
      pot=0,
      While[fn==0,{
        If[xn≤dn[[dia]],{tw[[i]]=tw[[i]]+g[xn]*v≤pot,fn=1},xn=xn*ab[1]],
        dia++,
        pot++
      }
    }
  ];j++]

```

```

    tw[[i]]=v  $\frac{tw[[i]]}{m}$ 
  };i++]
  cont=Table[0,{i,1,np}],
  i=1,
  While[i<=np,{
    If[gn[[i]]<tw[[i]],cont[[i]]=i,cont[[i]]=0],
    If[cont[[i]]==0,i=np,0],
    i++
  }]
  If[Max[cont]+1>np,dn[[fecha]]=RANGOMAYOR,dn[[fecha]]=x[[Max[cont]+1]]]
};fecha=fecha-1];
fecha=fecha-1]
tiempof=Date[]
tiempof-tiempoi
dn

```

Los datos que arroja este programa son análogos a los siguientes:

```

{ año, mes, día, hora, minutos, segundos }
{  $d_0, d_1, \dots, d_{n-1}, d_n$  }
{ año, mes, día, hora, minutos, segundos }

```

donde la primera y última línea hacen referencia al tiempo de inicio y termino del programa respectivamente, y la segunda línea muestra los valores estimados de los d_n 's.

Apéndice B

Datos simulados

Este apéndice está creado para consulta de los datos simulados por medio de los programas que se encuentran en el apéndice A y que fueron usados para la exposición de los ejemplos 5.2.1 y 5.2.2.

En las tablas B.1, B.2 y B.3 se encuentran los datos simulados para los valores de $M=5000, 10000$ y 20000 respectivamente utilizados en el ejemplo $\text{bernoulli}(0.5, 1.1, 0.9)$ 5.2.1. En las tablas B.4, B.5 y B.6 se encuentran los datos simulados para $M=5000, 10000$ y 20000 respectivamente en el caso del ejemplo $\text{lognormal}(0, 0.01)$ 5.2.2.

Tabla B.1:

d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
79	75	76	73	67	69	65	60	55	0
80	77	76	73	68	70	65	60	55	0
81	77	76	74	68	70	65	60	55	0
81	77	77	74	69	70	65	60	55	0
81	78	77	75	69	70	66	60	55	0
81	78	78	75	70	70	66	60	55	0
82	78	79	75	71	70	66	60	55	0
82	78	79	75	71	71	66	60	55	0
82	78	79	75	71	71	66	60	55	0
83	78	80	75	72	71	66	60	55	0
83	78	80	75	72	71	66	60	55	0
83	79	80	75	72	71	66	60	55	0
83	79	80	76	73	71	66	61	55	0
83	79	80	76	73	71	67	61	55	0
83	79	80	76	73	71	67	61	55	0
83	79	80	76	73	72	67	61	55	0
84	79	80	76	73	72	67	61	55	0
84	79	80	76	73	72	67	61	55	0
84	79	80	76	73	72	67	61	55	0
84	79	80	76	73	72	67	61	55	0
84	80	80	76	74	72	67	61	55	0
84	80	81	76	74	72	67	61	55	0
84	80	81	76	74	72	67	61	55	0
84	81	81	77	74	72	67	61	55	0
84	81	81	77	74	72	67	61	55	0
84	81	81	77	74	72	67	61	55	0
85	81	81	77	74	72	67	61	55	0
85	81	81	77	75	72	67	61	55	0
85	81	81	77	75	72	67	61	55	0
85	82	81	77	75	72	67	61	55	0
85	82	82	77	75	73	67	61	55	0
85	83	82	78	75	73	67	61	55	0
85	83	82	78	75	73	67	61	55	0
85	83	82	78	75	73	67	61	55	0
85	83	82	78	75	73	67	61	55	0
86	83	82	78	75	73	67	61	55	0
86	83	82	78	75	73	67	61	55	0
86	83	82	78	76	73	68	61	55	0
86	84	82	78	76	73	68	61	56	0
86	84	83	78	76	73	68	61	56	0
86	84	83	78	76	73	68	61	56	0
86	84	83	79	77	73	68	61	56	0
86	84	83	79	77	73	68	61	56	0
86	84	83	79	77	73	68	61	56	0
87	84	83	79	77	74	68	62	56	0
87	84	83	79	77	74	68	62	56	0
87	85	83	79	77	74	68	62	56	0
87	85	83	79	77	74	68	62	56	0
87	85	83	79	77	74	68	62	56	0
87	86	83	79	77	74	68	62	56	0
87	86	84	79	78	74	68	62	56	0
87	86	84	79	79	74	68	62	56	0
88	86	84	79	79	74	68	62	56	0
88	86	84	79	79	74	69	62	56	0
88	87	84	79	79	74	69	62	56	0
88	87	84	80	80	74	69	63	57	0
88	87	84	80	80	74	69	63	57	0
89	87	84	80	80	74	69	63	57	0
89	88	85	80	80	75	69	63	57	0
89	88	85	81	80	75	70	63	57	0
89	88	85	81	80	75	70	63	57	0
89	88	85	81	80	75	70	63	57	0
90	89	85	81	80	75	70	63	57	0
90	89	85	81	80	75	70	63	57	0
91	89	86	81	80	76	70	63	57	0
91	89	86	82	80	76	70	64	58	0
91	90	86	82	80	76	70	66	58	0
93	90	87	84	81	76	71	67	58	0
93	90	88	85	81	77	72	67	58	0
93	92	88	85	83	77	72	68	59	0
95	95	89	87	83	78	78	68	60	0

Tabla B.2:

d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
80	77	74	74	69	69	66	60	55	0
82	79	75	75	70	71	66	60	55	0
82	79	77	75	72	71	66	60	55	0
82	79	78	75	73	72	66	61	55	0
82	80	79	76	74	72	66	61	55	0
82	81	79	76	74	72	67	61	55	0
83	82	79	76	74	72	67	61	55	0
83	82	79	76	74	72	67	61	55	0
83	83	80	76	75	72	67	61	55	0
84	83	80	76	76	72	67	61	55	0
84	83	80	76	76	72	67	61	55	0
84	83	80	77	76	72	67	61	55	0
84	83	80	77	76	72	67	61	55	0
85	83	80	77	76	73	67	61	55	0
85	83	81	77	76	73	67	61	55	0
85	83	81	78	76	73	67	61	55	0
85	84	81	78	76	73	67	61	55	0
86	84	81	78	77	73	67	61	55	0
86	84	82	78	77	73	67	61	55	0
86	84	82	78	77	73	67	61	55	0
86	84	82	78	77	73	67	61	55	0
86	84	82	78	77	73	67	61	55	0
86	85	82	78	77	73	67	61	55	0
87	85	82	78	77	73	67	61	55	0
87	85	82	78	77	73	67	61	55	0
87	85	82	78	77	73	67	61	55	0
87	85	82	78	77	73	67	61	55	0
87	85	82	79	77	73	67	61	55	0
87	85	82	79	78	74	67	61	55	0
87	85	82	79	78	74	68	61	55	0
88	85	82	79	78	74	68	61	55	0
88	86	82	80	78	74	68	61	55	0
88	86	83	80	78	74	68	61	56	0
88	86	83	80	78	74	68	61	56	0
88	86	83	80	78	74	68	61	56	0
88	86	83	80	79	74	68	61	56	0
88	86	83	80	79	74	68	61	56	0
88	86	83	80	79	74	68	61	56	0
88	86	83	80	79	74	68	62	56	0
88	86	83	81	79	74	68	62	56	0
88	86	83	81	79	74	68	62	56	0
89	87	83	81	79	74	68	62	56	0
89	87	83	81	79	74	68	62	56	0
89	87	83	81	80	74	68	62	56	0
89	87	83	82	80	74	68	62	56	0
89	87	83	82	80	74	68	62	56	0
89	87	83	82	80	74	68	62	56	0
90	88	83	82	80	74	68	62	56	0
90	88	84	82	80	74	68	62	56	0
90	88	84	82	80	74	68	62	56	0
90	89	84	82	80	74	68	62	56	0
90	89	84	83	80	74	68	62	56	0
90	89	84	83	80	75	68	62	57	0
90	89	84	83	80	75	69	62	57	0
90	89	85	83	80	75	69	62	57	0
91	89	86	83	80	75	69	62	57	0
91	89	86	84	80	75	69	63	57	0
91	89	86	84	80	75	69	63	57	0
91	90	86	84	80	75	69	63	57	0
92	90	87	84	80	75	69	63	57	0
92	90	87	84	81	75	69	63	57	0
92	90	87	84	81	75	69	63	57	0
92	90	87	84	81	75	69	63	57	0
93	91	87	85	81	76	70	63	57	0
93	91	88	85	82	76	70	63	58	0
93	91	88	85	82	76	70	64	58	0
93	91	88	85	82	76	71	64	58	0
93	92	88	85	82	77	71	64	59	0
93	92	89	86	82	77	71	64	59	0
94	92	89	86	83	79	71	65	59	0
96	93	89	87	83	79	71	65	59	0
97	94	89	87	83	80	72	67	60	0
100	95	92	89	89	80	73	68	60	0

Tabla B.3:

d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
84	79	77	75	75	71	66	60	55	0
85	83	79	75	76	72	66	60	55	0
85	84	79	76	77	72	66	61	55	0
86	85	80	76	78	73	67	61	55	0
86	85	81	77	78	73	67	61	55	0
86	85	81	78	78	73	67	61	55	0
86	85	81	78	78	73	67	61	55	0
86	86	82	78	78	73	67	61	55	0
87	86	82	78	78	73	67	61	55	0
88	86	82	78	78	73	67	61	55	0
88	86	83	78	78	73	67	61	55	0
89	86	83	79	78	73	67	61	55	0
89	86	83	79	78	73	67	61	55	0
89	87	83	79	79	73	67	61	55	0
89	87	83	79	79	73	67	61	55	0
89	87	83	79	79	74	67	61	55	0
89	87	83	79	79	74	67	61	55	0
89	87	83	80	79	74	67	61	55	0
89	88	83	80	79	74	67	61	55	0
89	88	83	80	79	74	67	61	55	0
90	88	83	80	79	74	67	61	55	0
90	89	84	80	79	74	67	61	55	0
90	89	84	80	79	74	67	61	55	0
90	89	84	80	79	74	67	61	55	0
91	89	84	80	79	74	67	61	55	0
91	89	84	81	80	74	67	61	55	0
91	89	84	81	80	74	67	61	55	0
91	89	84	81	80	74	67	61	55	0
92	89	84	81	80	74	67	61	55	0
92	89	84	81	80	74	67	61	55	0
92	90	84	82	80	74	68	61	55	0
93	90	84	82	80	74	68	61	55	0
93	90	84	82	80	74	68	61	55	0
93	90	85	82	80	74	68	61	56	0
93	90	85	82	80	74	68	61	56	0
93	90	85	82	80	74	68	61	56	0
93	90	85	82	80	74	68	61	56	0
93	90	85	82	80	75	68	61	56	0
93	90	85	82	80	75	68	61	56	0
94	91	85	83	80	75	68	61	56	0
94	91	85	83	80	75	68	61	56	0
94	91	85	83	81	75	68	61	56	0
94	91	86	83	81	75	68	61	56	0
94	91	86	83	81	75	68	61	56	0
94	91	86	83	81	75	68	62	56	0
94	91	86	83	81	75	68	62	56	0
95	91	86	84	81	75	68	62	56	0
95	91	86	84	81	75	68	62	56	0
95	91	86	84	81	75	69	62	56	0
95	92	86	85	81	75	69	62	56	0
95	92	87	85	81	75	69	62	57	0
95	92	87	85	81	75	69	62	57	0
96	92	87	86	81	75	69	62	57	0
96	92	87	86	81	75	69	62	57	0
96	92	87	86	81	75	69	62	57	0
96	92	87	86	81	75	69	62	57	0
96	92	87	86	82	75	69	62	57	0
96	92	87	86	82	75	69	62	57	0
97	92	88	87	82	75	69	62	57	0
97	92	88	87	82	75	70	62	57	0
97	93	88	87	82	76	70	62	57	0
97	93	88	87	83	76	70	62	57	0
97	94	89	88	83	76	70	62	57	0
98	94	89	88	83	77	70	63	58	0
98	94	89	88	83	77	70	63	58	0
98	94	89	88	83	77	70	64	58	0
98	95	89	89	83	77	70	64	58	0
100	95	90	90	84	78	71	64	58	0
100	96	91	90	84	78	72	64	59	0
101	98	92	91	84	78	76	65	60	0

Tabla B.4:

d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
51.7	51.2	51.3	51.1	50.9	50.6	50.4	50.1	49.8	0
51.8	51.5	51.3	51.1	50.9	50.6	50.5	50.2	49.9	0
51.8	51.6	51.4	51.2	51	50.7	50.5	50.2	49.9	0
51.8	51.6	51.4	51.2	51.1	50.8	50.6	50.2	49.9	0
51.8	51.6	51.5	51.3	51.1	50.8	50.6	50.2	49.9	0
51.9	51.7	51.5	51.3	51.1	50.8	50.7	50.2	49.9	0
51.9	51.7	51.5	51.3	51.2	50.8	50.7	50.3	49.9	0
51.9	51.7	51.5	51.3	51.2	50.8	50.7	50.3	49.9	0
51.9	51.7	51.5	51.3	51.2	50.9	50.7	50.3	49.9	0
51.9	51.7	51.5	51.3	51.2	50.9	50.7	50.3	49.9	0
52	51.7	51.6	51.4	51.2	50.9	50.7	50.3	49.9	0
52	51.7	51.6	51.4	51.2	50.9	50.7	50.4	49.9	0
52	51.8	51.6	51.4	51.2	50.9	50.7	50.4	49.9	0
52	51.8	51.6	51.4	51.3	50.9	50.7	50.4	50	0
52	51.8	51.6	51.4	51.3	51	50.8	50.4	50	0
52	51.8	51.7	51.4	51.3	51	50.8	50.4	50	0
52	51.8	51.7	51.4	51.3	51	50.8	50.4	50	0
52.1	51.8	51.7	51.5	51.4	51	50.8	50.4	50	0
52.1	51.8	51.7	51.5	51.4	51	50.8	50.4	50	0
52.1	51.8	51.7	51.5	51.4	51	50.8	50.4	50	0
52.1	51.8	51.7	51.5	51.4	51	50.8	50.4	50	0
52.1	51.8	51.7	51.5	51.5	51	50.8	50.4	50.1	0
52.1	51.9	51.7	51.5	51.5	51	50.8	50.4	50.1	0
52.2	51.9	51.7	51.5	51.5	51.1	50.8	50.5	50.1	0
52.2	51.9	51.8	51.5	51.5	51.1	50.8	50.5	50.1	0
52.2	51.9	51.8	51.6	51.5	51.1	50.8	50.5	50.1	0
52.2	51.9	51.8	51.6	51.5	51.1	50.8	50.5	50.1	0
52.2	52	51.8	51.6	51.5	51.1	50.8	50.5	50.1	0
52.2	52	51.8	51.6	51.5	51.1	50.8	50.5	50.1	0
52.2	52	51.8	51.6	51.6	51.1	50.9	50.5	50.1	0
52.2	52	51.8	51.7	51.6	51.1	50.9	50.5	50.1	0
52.2	52	51.8	51.7	51.6	51.1	50.9	50.5	50.1	0
52.2	52	51.8	51.7	51.6	51.2	50.9	50.5	50.1	0
52.3	52	51.9	51.7	51.6	51.2	50.9	50.5	50.1	0
52.3	52	51.9	51.8	51.6	51.2	51	50.6	50.2	0
52.3	52.1	51.9	51.8	51.6	51.2	51	50.6	50.2	0
52.3	52.1	51.9	51.8	51.6	51.2	51	50.6	50.2	0
52.3	52.1	51.9	51.8	51.6	51.2	51	50.6	50.2	0
52.3	52.1	51.9	51.8	51.6	51.2	51	50.6	50.2	0
52.3	52.1	51.9	51.8	51.6	51.3	51	50.6	50.2	0
52.3	52.2	51.9	51.8	51.6	51.3	51	50.7	50.2	0
52.4	52.2	51.9	51.8	51.6	51.3	51	50.7	50.2	0
52.4	52.2	52	51.8	51.6	51.3	51	50.7	50.2	0
52.4	52.2	52	51.9	51.6	51.3	51	50.7	50.2	0
52.4	52.2	52	51.9	51.7	51.4	51	50.7	50.2	0
52.4	52.3	52	51.9	51.7	51.4	51	50.7	50.2	0
52.4	52.3	52	51.9	51.7	51.4	51.1	50.7	50.3	0
52.5	52.3	52	51.9	51.7	51.4	51.1	50.8	50.3	0
52.5	52.3	52	51.9	51.7	51.4	51.1	50.8	50.3	0
52.5	52.3	52.1	52	51.7	51.5	51.1	50.8	50.3	0
52.5	52.3	52.1	52	51.8	51.5	51.1	50.8	50.3	0
52.5	52.4	52.1	52	51.8	51.5	51.1	50.8	50.3	0
52.5	52.4	52.1	52	51.8	51.5	51.1	50.9	50.3	0
52.6	52.4	52.1	52	51.8	51.5	51.1	50.9	50.3	0
52.6	52.4	52.2	52	51.9	51.6	51.2	50.9	50.3	0
52.6	52.4	52.2	52	51.9	51.6	51.2	50.9	50.4	0
52.6	52.4	52.2	52.1	51.9	51.6	51.2	50.9	50.4	0
52.6	52.4	52.2	52.2	51.9	51.6	51.2	50.9	50.4	0
52.6	52.5	52.2	52.2	51.9	51.6	51.3	50.9	50.4	0
52.8	52.5	52.2	52.2	52	51.6	51.3	50.9	50.4	0
52.8	52.5	52.3	52.3	52	51.6	51.3	51	50.5	0
52.9	52.5	52.3	52.3	52	51.7	51.3	51	50.5	0
52.9	52.6	52.3	52.4	52	51.7	51.4	51	50.5	0
52.9	52.6	52.4	52.5	52.1	51.7	51.5	51.1	50.6	0
52.9	52.6	52.5	52.6	52.1	51.9	51.5	51.1	50.6	0
53	52.6	52.5	52.7	52.2	51.9	51.6	51.2	50.9	0
53.3	52.7	52.5	52.8	52.4	52	51.9	51.3	51.1	0
53.4	53.2	52.5	52.9	52.5	52	51.9	51.4	51.9	0

Tabla B.5:

d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
51.7	51.8	51.6	51.4	51.1	50.7	50.5	50.4	49.9	0
51.9	51.8	51.7	51.4	51.1	50.9	50.6	50.4	49.9	0
52	51.8	51.7	51.4	51.2	50.9	50.7	50.5	50	0
52	51.9	51.8	51.5	51.2	51	50.7	50.5	50	0
52	51.9	51.8	51.5	51.2	51.1	50.7	50.5	50	0
52.1	51.9	51.8	51.5	51.3	51.1	50.7	50.5	50.1	0
52.2	52	51.8	51.5	51.4	51.2	50.8	50.5	50.1	0
52.2	52	51.8	51.5	51.4	51.2	50.8	50.5	50.1	0
52.3	52	51.8	51.5	51.4	51.2	50.8	50.6	50.1	0
52.4	52	51.9	51.6	51.4	51.2	50.9	50.6	50.1	0
52.4	52	51.9	51.6	51.4	51.2	50.9	50.6	50.1	0
52.4	52.1	51.9	51.6	51.5	51.2	50.9	50.6	50.1	0
52.4	52.1	51.9	51.6	51.5	51.2	50.9	50.6	50.1	0
52.4	52.1	51.9	51.6	51.5	51.2	50.9	50.6	50.1	0
52.4	52.1	51.9	51.6	51.5	51.2	50.9	50.6	50.1	0
52.4	52.1	52	51.6	51.5	51.3	50.9	50.6	50.1	0
52.4	52.2	52	51.6	51.5	51.3	50.9	50.6	50.1	0
52.4	52.2	52	51.7	51.5	51.3	51	50.6	50.1	0
52.5	52.2	52	51.7	51.5	51.3	51	50.6	50.1	0
52.5	52.2	52	51.7	51.6	51.3	51	50.6	50.1	0
52.5	52.2	52	51.7	51.6	51.3	51	50.7	50.1	0
52.5	52.2	52.1	51.8	51.6	51.3	51	50.7	50.2	0
52.5	52.2	52.1	51.8	51.6	51.3	51	50.7	50.2	0
52.5	52.2	52.1	51.8	51.7	51.3	51	50.7	50.2	0
52.5	52.2	52.1	51.8	51.7	51.3	51.1	50.7	50.2	0
52.5	52.3	52.1	51.8	51.7	51.3	51.1	50.7	50.2	0
52.5	52.3	52.1	51.8	51.7	51.3	51.1	50.7	50.2	0
52.6	52.3	52.1	51.8	51.8	51.4	51.1	50.7	50.2	0
52.6	52.4	52.1	51.9	51.8	51.4	51.1	50.7	50.2	0
52.6	52.4	52.1	51.9	51.8	51.4	51.1	50.7	50.2	0
52.6	52.4	52.2	51.9	51.8	51.4	51.1	50.7	50.2	0
52.6	52.4	52.2	51.9	51.8	51.4	51.1	50.8	50.2	0
52.6	52.4	52.2	51.9	51.8	51.4	51.1	50.8	50.2	0
52.6	52.4	52.2	51.9	51.8	51.5	51.1	50.8	50.2	0
52.6	52.5	52.2	52	51.8	51.5	51.2	50.8	50.2	0
52.6	52.5	52.2	52	51.9	51.5	51.2	50.8	50.3	0
52.7	52.5	52.2	52	51.9	51.6	51.2	50.8	50.3	0
52.7	52.5	52.3	52	51.9	51.6	51.2	50.8	50.3	0
52.7	52.5	52.3	52.1	51.9	51.6	51.2	50.9	50.3	0
52.7	52.6	52.3	52.1	51.9	51.6	51.2	50.9	50.3	0
52.7	52.6	52.3	52.1	51.9	51.6	51.2	50.9	50.3	0
52.7	52.6	52.3	52.1	51.9	51.6	51.2	50.9	50.3	0
52.8	52.6	52.3	52.1	52	51.6	51.2	50.9	50.3	0
52.8	52.6	52.4	52.1	52	51.6	51.3	50.9	50.3	0
52.8	52.6	52.4	52.2	52	51.6	51.3	50.9	50.4	0
52.8	52.6	52.4	52.2	52	51.6	51.3	51	50.4	0
52.8	52.6	52.4	52.2	52	51.7	51.4	51	50.5	0
52.8	52.7	52.4	52.2	52	51.7	51.4	51	50.5	0
52.9	52.7	52.4	52.2	52	51.7	51.4	51	50.5	0
52.9	52.7	52.5	52.2	52	51.7	51.4	51	50.5	0
52.9	52.7	52.5	52.3	52	51.7	51.4	51	50.5	0
52.9	52.7	52.5	52.3	52.1	51.7	51.4	51.1	50.5	0
52.9	52.7	52.5	52.3	52.1	51.8	51.4	51.1	50.5	0
53	52.8	52.5	52.3	52.1	51.8	51.4	51.1	50.6	0
53	52.8	52.6	52.3	52.1	51.9	51.4	51.2	50.6	0
53.1	52.8	52.6	52.3	52.1	51.9	51.4	51.2	50.6	0
53.1	52.8	52.6	52.3	52.2	51.9	51.5	51.2	50.6	0
53.1	52.9	52.6	52.4	52.2	51.9	51.5	51.2	50.6	0
53.1	53	52.7	52.4	52.2	52	51.5	51.2	50.6	0
53.1	53.1	52.7	52.4	52.2	52	51.5	51.3	50.6	0
53.1	53.1	52.8	52.4	52.3	52	51.5	51.3	50.6	0
53.2	53.1	52.8	52.4	52.3	52.1	51.6	51.4	50.7	0
53.3	53.1	52.9	52.5	52.4	52.1	51.6	51.4	50.7	0
53.3	53.2	53	52.5	52.5	52.1	51.7	51.5	50.7	0
53.3	53.2	53.1	52.5	52.5	52.2	51.7	51.9	50.8	0
53.5	53.3	53.1	52.6	52.6	52.2	51.7	51.9	50.8	0
53.7	53.4	53.1	52.6	52.7	52.5	51.8	52.1	50.8	0
53.7	53.9	53.2	52.7	52.9	52.6	52	52.1	50.9	0

Tabla B.6:

d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
52.1	51.9	51.9	51.6	51.3	51.1	50.7	50.4	50.1	0
52.1	52.1	51.9	51.6	51.3	51.1	50.8	50.5	50.1	0
52.2	52.1	51.9	51.7	51.4	51.2	50.8	50.6	50.1	0
52.3	52.2	52	51.7	51.4	51.2	50.9	50.6	50.1	0
52.4	52.2	52	51.7	51.5	51.2	50.9	50.6	50.1	0
52.4	52.3	52.1	51.7	51.5	51.2	50.9	50.7	50.1	0
52.5	52.3	52.1	51.9	51.5	51.3	50.9	50.7	50.2	0
52.5	52.3	52.1	51.9	51.5	51.3	50.9	50.7	50.2	0
52.5	52.4	52.2	51.9	51.6	51.4	51	50.7	50.2	0
52.5	52.4	52.2	51.9	51.6	51.4	51	50.7	50.2	0
52.5	52.4	52.2	51.9	51.6	51.4	51	50.7	50.2	0
52.5	52.4	52.2	52	51.6	51.4	51	50.7	50.2	0
52.5	52.4	52.2	52	51.7	51.5	51	50.7	50.3	0
52.6	52.4	52.2	52	51.7	51.5	51	50.8	50.3	0
52.6	52.4	52.3	52	51.7	51.5	51	50.8	50.3	0
52.6	52.4	52.3	52	51.7	51.5	51	50.8	50.3	0
52.7	52.5	52.3	52	51.7	51.6	51.1	50.8	50.3	0
52.7	52.5	52.3	52.1	51.7	51.6	51.1	50.8	50.3	0
52.8	52.5	52.4	52.1	51.7	51.6	51.1	50.8	50.3	0
52.8	52.5	52.4	52.1	51.8	51.6	51.1	50.9	50.3	0
52.8	52.5	52.4	52.1	51.8	51.6	51.2	50.9	50.4	0
52.9	52.6	52.4	52.1	51.8	51.6	51.2	50.9	50.4	0
52.9	52.6	52.4	52.1	51.8	51.6	51.2	50.9	50.4	0
52.9	52.6	52.5	52.1	51.8	51.6	51.2	50.9	50.4	0
52.9	52.6	52.5	52.2	51.8	51.6	51.2	50.9	50.4	0
52.9	52.6	52.5	52.2	51.9	51.6	51.2	50.9	50.4	0
52.9	52.6	52.5	52.2	51.9	51.7	51.2	50.9	50.4	0
52.9	52.7	52.5	52.2	51.9	51.7	51.3	50.9	50.4	0
52.9	52.7	52.5	52.2	51.9	51.7	51.3	51	50.4	0
53	52.8	52.5	52.2	51.9	51.7	51.3	51	50.4	0
53	52.8	52.5	52.2	51.9	51.7	51.3	51	50.5	0
53	52.8	52.6	52.2	51.9	51.8	51.3	51	50.5	0
53	52.8	52.6	52.3	51.9	51.8	51.3	51	50.5	0
53	52.8	52.6	52.3	52	51.8	51.4	51.1	50.5	0
53	52.8	52.6	52.3	52	51.8	51.4	51.1	50.5	0
53.1	52.8	52.7	52.3	52	51.9	51.4	51.1	50.5	0
53.1	52.8	52.7	52.3	52	51.9	51.4	51.1	50.5	0
53.1	52.8	52.7	52.3	52.1	51.9	51.4	51.1	50.5	0
53.1	52.8	52.7	52.4	52.1	51.9	51.4	51.2	50.5	0
53.1	52.9	52.7	52.4	52.1	51.9	51.4	51.2	50.6	0
53.1	52.9	52.8	52.4	52.1	51.9	51.4	51.2	50.6	0
53.2	52.9	52.8	52.4	52.1	51.9	51.5	51.2	50.6	0
53.2	52.9	52.8	52.4	52.2	52	51.5	51.2	50.6	0
53.2	52.9	52.8	52.4	52.2	52	51.5	51.2	50.6	0
53.3	53	52.9	52.4	52.2	52	51.5	51.3	50.6	0
53.3	53	52.9	52.5	52.2	52	51.6	51.3	50.6	0
53.3	53	52.9	52.5	52.2	52	51.6	51.3	50.6	0
53.3	53	52.9	52.5	52.3	52	51.6	51.3	50.6	0
53.3	53	52.9	52.5	52.3	52	51.6	51.3	50.7	0
53.3	53	52.9	52.5	52.3	52.1	51.6	51.3	50.7	0
53.3	53	52.9	52.6	52.3	52.1	51.6	51.3	50.7	0
53.4	53.1	53	52.6	52.3	52.1	51.7	51.4	50.7	0
53.4	53.1	53	52.6	52.3	52.3	51.7	51.4	50.7	0
53.4	53.1	53	52.6	52.4	52.3	51.8	51.4	50.7	0
53.4	53.1	53	52.7	52.4	52.4	51.8	51.4	50.7	0
53.4	53.1	53.1	52.7	52.4	52.4	51.8	51.4	50.7	0
53.4	53.2	53.1	52.7	52.4	52.5	51.8	51.4	50.8	0
53.5	53.2	53.1	52.8	52.4	52.5	51.8	51.5	50.8	0
53.5	53.3	53.1	52.8	52.4	52.5	51.8	51.6	50.9	0
53.6	53.3	53.2	52.8	52.4	52.6	51.9	51.7	50.9	0
53.6	53.3	53.2	52.9	52.5	52.6	51.9	51.7	51	0
53.7	53.4	53.4	52.9	52.6	52.6	51.9	51.8	51	0
53.7	53.6	53.4	53.1	52.6	52.6	51.9	51.8	51.2	0
53.9	53.6	53.4	53.2	52.6	52.7	52	51.9	51.2	0
53.9	53.6	53.4	53.3	52.6	52.8	52	51.9	51.4	0
53.9	53.8	53.6	53.3	52.8	52.8	52.1	51.9	51.5	0
53.9	54.2	53.7	54.1	52.9	52.9	52.3	52.1	51.6	0

Bibliografía

- [1] Javier Martín Pliego, Luis Ruiz Pérez. *Estadística I: probabilidad*. Editorial AC; Madrid, 1995.
- [2] John C. Hull. *Introduction to futures and options markets*. Editorial Prentice Hall; New Jersey, 1998.
- [3] Miguel A. García Álvarez. *Introducción a la teoría de la probabilidad*, primer y segundo curso. Editorial Fondo de Cultura Económica, 2005.
- [4] Onesimo Hernandez-Lerma, Jean B. Lasserre. *Discrete-time Markov control processes : basic optimality criteria*. Editorial Springer; New York, 1995.
- [5] Kukush A. G., Silvestrov D. S. "Optimal pricing for American type options with discrete time". In: *Proceedings of the Eighth International School on Mathematical and Statistical Methods in Economics, Finance and Insurance*, 2004.
- [6] Shiryaev Albert N. *Essentials of stochastic finance: facts, models, theory*. Editorial World scientific, Singapore 1999.
- [7] Shiryaev Albert N. *Optimal stopping rules*. Editorial Springer-Verlag; New York, 1978.
- [8] Shiryaev Albert N. *Probability*. Editorial Springer; New York, 1995.
- [9] Silvestrov D.S., Galochkin V.G., Sibirtsev V.G., "Algorithms and Programs for optimal Monte Carlo pricing of American options", In: *Proceedings of the Second International School on Actuarial and Financial Mathematics*. Kiev, 1999.
- [10] Jönsson H., "Monte Carlo studies of american type call options with discrete time", *Theory of stochastic processes*, Vol. 7(23), 2001.
- [11] Tapas K. Chandra, Dipak Chatterjee. *A first course in probability*. Editorial Narosa Publishing House; New Delhi, 2001.

- [12] Feller, W. *An introduction to probability theory and Its applications, vol. I.* Editorial John Wiley; New York, 1971.
- [13] Feller, W. *An introduction to probability theory and Its applications, vol. II.* Editorial John Wiley; New York, 1971.