



**CENTRO DE INVESTIGACIÓN Y DE
ESTUDIOS AVANZADOS DEL
INSTITUTO POLITÉCNICO NACIONAL**

UNIDAD ZACATENCO

DEPARTAMENTO DE CONTROL AUTOMÁTICO

**“Aplicaciones de la Teoría de la
Probabilidad”**

TESIS

Que presenta

Maleni Maricarmen Pérez Sarmiento

Para Obtener el Grado de
MAESTRO EN CIENCIAS

En la especialidad de
CONTROL AUTOMÁTICO

Director de Tesis:

Dr. Jorge Alberto León Vázquez

Ciudad de México

Agosto, 2020

Agradecimientos

Al Consejo de Ciencia y Tecnología, CONACYT, por el apoyo económico brindado durante la maestría con la beca al No. de CVU 937198.

Al Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, CINVESTAV, por darme la oportunidad de formar parte de su comunidad y brindarme todo lo necesario para crecer profesionalmente durante dos años.

A mis padres, por su amor y apoyo incondicional e inmensurable. Gracias por guiarme y cuidarme siempre.

A mi hermano, por romper los momentos de estrés y tensión con sus divertidas ocurrencias cuando más lo necesitaba. Gracias por no dejarme sola.

A Lenín Álvarez, Ilse Pérez y Alfredo Figueroa, por escuchar mis quejas y mis miedos y regalarme motivación para seguir adelante. Gracias por creer en mí incluso cuando yo misma no lo hacía.

A mis profesores durante la maestría les agradezco por el conocimiento y la exigencia que aportaron a mi formación durante estos años.

Al Dr. Jorge Alberto León Vázquez, director de esta tesis, por los conocimientos, las asesorías y su tiempo. Gracias por darme la oportunidad de realizar este trabajo.

A Iván Rodríguez. No tengo cómo agradecerte por toda la ayuda y paciencia que me brindaste. Gracias por ser un amigo memorable.

A Juan Carlos Bocanegra, Ellis Frías, Gerson Portilla, Manolo de Urquijo, Héctor Vargas, Antonio Hurtado y Gladys Serra, les agradezco por todos los momentos de estrés y alegría compartidos. Gracias por ser tan buenos compañeros de clase, sobre todo, buenos amigos.

A José Luis Luna y Jessica Maldonado, por los consejos, las asesorías y los ánimos.

A Dios, por poner a todas estas increíbles personas en mi vida.

Índice general

Introducción	5
1. Definiciones básicas	8
2. Funciones de densidad e Histogramas de frecuencia.	24
3. Filtro de Kalman	35
3.1. Filtro de Kalman en tiempo continuo	45
4. El problema de Monty Hall	50
5. Juegos de Azar	53
5.1. Teoría de Riesgo	60
5.2. Cadenas de Markov	64
6. Opciones financieras	73
6.1. El modelo de Cox-Ross-Rubinstein	74
6.2. El modelo de Black-Scholes	89
Conclusiones	101
A. Ecuaciones en diferencias.	103
A.1. Solución de una ecuación en diferencias	104
B. Volatilidad histórica	107
Bibliografía	110

Introducción

Para la mayoría de las personas pensar en *probabilidad* es pensar inmediatamente en los juegos de azar: los naipes, las quinielas, incluso lanzar una moneda. Podría decirse que nuestra mente ha asociado este concepto a cualquier situación que presente las posibilidades de *ganar o perder* y aunque fueron dichos juegos los que inspiraron el inicio del estudio de la teoría de la probabilidad, la realidad actual es otra.

La curiosidad del ser humano por querer predecir eventos futuros hizo que se sintiera atraído a estudiar el comportamiento de fenómenos aleatorios, es decir, aquellos fenómenos con diferentes resultados posibles. De hecho, se cree que pudiesen haber sido los egipcios quienes crearon el antecesor del juego de dados, ellos tallaban el hueso del talón de ciertos animales de forma que éste pudiera caer en cuatro posiciones diferentes y registraban los resultados obtenidos.

En la historia de la probabilidad se considera que los matemáticos Pierre de Fermat (1601-1665) y Blaise Pascal (1623-1662) fueron quienes comenzaron a formalizar las teorías de la probabilidad analizando el juego de los dados [29], sin embargo nunca publicaron formalmente sus resultados. Los primeros escritos que intentaban interpretar el comportamiento de este juego fueron publicados en la edad media [12] y fue Christiaan Huygens (1629-1695) quien lo hizo basándose en la correspondencia que Pascal y Fermat mantenían [29].

Fue Jacob Bernoulli (1654-1705) el primero en publicar una definición del concepto de probabilidad [29], a la par, diversos matemáticos alrededor del mundo comenzaron a intentar darle formalidad al estudio de la misma, entre los cuales se encontraban Joseph Lagrange (1736-1819) y Thomas Bayes (1702-1761) quienes se dedicaron a unificar las ideas que se tenían sobre estos fenómenos, obteniendo la primera teoría general de las probabilidades [12].

Con el paso de los años y los aportes en la materia, la probabilidad comenzó a ser usada como respuesta al comportamiento no solo de los juegos de

azar, sino de fenómenos que pertenecían a otras áreas de estudio, ejemplo de ello fue Nikolas Bernoulli (1687-1759) quien dio solución a problemas en derecho o el comerciante John Graunt, quien la utilizó para estudiar fenómenos tanto biológicos como estadísticos [12] tales como la predicción de natalidad y mortalidad.

Otro ejemplo más es Gorrfried Achenwell (1719-1772) quien proponía el uso de la probabilidad y la estadística para estudiar fenómenos sociales [12], argumentando que el número de crímenes se relacionaba con la cantidad de personas que no habían tenido acceso a la educación.

Los avances más recientes de la teoría de la probabilidad se han basado en poder vincularla con la *Teoría de la Medida*, otra rama más de las matemáticas. De hecho, fue el matemático Andréi Kolmogórov (1903-1987) quien lo hizo [30]. Entre algunos de los matemáticos importantes que han contribuido en el desarrollo más reciente de esta materia se encuentran Nirbert Wiener (1894-1964) y Kiyoshi Itô (1915-2008).

Conforme la historia avanza, se ha relacionado a la teoría de la probabilidad con más ramas de diversas ciencias dándole así un sin fin de aplicaciones prácticas en ingeniería, física, ciencias sociales, medicina, economía, entre otros más, lo que la ha convertido en una parte importante de las matemáticas. La predicción del clima, el costo de las primas de seguros, las posibilidades de ganar de un partido político, la estimación del tiempo de vida útil de un artefacto, el porcentaje de éxito de un tratamiento médico e incluso la esperanza de vida de una población son sólo algunas de las aplicaciones actuales.

Esta tesis tiene como objetivo mostrar con más detalle el uso de la teoría de la probabilidad en nuestros días, de forma que, el concepto que tenemos pueda ir ampliándose. Para ello, se expondrán problemáticas específicas de cinco temas distintos y cómo se han resuelto con las herramientas de esta materia.

La estructura de este trabajo se compone de seis capítulos: el primero se ha utilizado para explicar los conceptos básicos, además de algunos teoremas esenciales, con la finalidad de que el lector pueda comprender el desarrollo de los siguientes capítulos.

En el Capítulo 2 se presenta el uso de la *Ley de los grandes números* para aproximar funciones de densidad por medio de histogramas de frecuencia.

En el Capítulo 3 se expondrá el uso del *Filtro de Kalman* para la estimación de variables. A lo largo del Capítulo 4 se explora un problema de elección basado en un programa televisivo norteamericano, conocido como *El*

problema de Monty Hall (o su equivalente mexicano *La catafixia*), abordado desde la perspectiva de la probabilidad condicional.

En el desarrollo del Capítulo 5 se tratan los juegos de azar y la ruina del jugador. Además, se aborda la *Teoría de Riesgo*, explicando uno de los problemas más comunes de probabilidad: el momento de ruina o quiebra de una empresa aseguradora. En la última sección de dicho capítulo, se exponen las *cadena de Markov* y cómo fueron implementadas para formar las bases de uno de los buscadores de información en línea más importante de nuestra época: *Google*.

Por último, en el Capítulo 6 se explica cómo funcionan las opciones financieras, el problema que existe para valuarlas y dos de los métodos que se han propuesto para hacerlo: *El modelo de Cox-Ross-Rubinstein* y *El modelo de Black-Scholes*. Al finalizar dicho capítulo se exponen las conclusiones generales de este trabajo.

Capítulo 1

Definiciones básicas

A lo largo de esta sección se introducen los conceptos claves para entender el lenguaje utilizado en la teoría de la probabilidad. Las nociones más significativas, utilizadas en este trabajo, se presentan a continuación.

Se denominan *fenómenos aleatorios* a aquellos fenómenos cuyo resultado no puede predecirse anticipadamente dentro de todas las diferentes posibilidades a obtener. A cada posible resultado obtenido del fenómeno aleatorio se le conoce como *punto muestral*. El conjunto de todos los puntos muestrales es llamado *espacio muestral* y en general, es denotado por Ω .

Definición 1.1 (Sigma-álgebra) Sea Ω un conjunto no vacío, una familia \mathcal{A} , de subconjuntos de Ω , es llamada **σ -álgebra** si satisface los puntos siguientes:

- $\Omega \in \mathcal{A}$
- Para todo $A \in \mathcal{A}$, $A^c \in \mathcal{A}$.
- Si $A_n \in \mathcal{A}$ para todo n , entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

Definición 1.2 Sea C una familia de subconjuntos de Ω . $\sigma(C) := \bigcap_{C \subset \mathcal{A}} \mathcal{A}$ es la σ -álgebra de Ω más pequeña que contiene a C , donde \mathcal{A} es una σ -álgebra. A $\sigma(C)$ se le conoce como la **σ -álgebra generada por C** .

Al par formado por un conjunto no vacío, Ω , y una σ -álgebra, \mathcal{A} , se le denomina *espacio medible* y es denotado, en general, por (Ω, \mathcal{A}) . A los elementos de \mathcal{A} se les denomina *eventos*.

Definición 1.3 (Medida) Sea \mathcal{A} una σ -álgebra de un conjunto Ω . Se conoce como **medida** a una función de conjuntos no negativa, $\mu : \mathcal{A} \rightarrow \mathbb{R}^+ \cup \{0\}$, la cual cumple:

- $\mu(\emptyset) = 0$, donde \emptyset es el conjunto vacío.
- Dada una sucesión $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$ de conjuntos disjuntos a pares, se tiene que:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Es decir, μ es σ -aditiva.

Dados dos espacios medibles, (Ω, \mathcal{A}) , (φ, \mathcal{B}) , una función $f : \Omega \rightarrow \varphi$ se dice *medible* si $f^{-1}(B) \in \mathcal{A}$, para todo $B \in \mathcal{B}$. En este caso, se denota por $f : (\Omega, \mathcal{A}) \rightarrow (\varphi, \mathcal{B})$.

Cotidianamente se suele emplear el término *probabilidad* como la idea del nivel de certeza de que un evento pueda ocurrir dentro de un panorama de incertidumbre. Formalmente, la probabilidad se define como un método por el cual es posible obtener una medida asociada a la frecuencia de un evento en un fenómeno aleatorio. Es decir, qué tan posible es que un evento específico ocurra antes de realizar un experimento: La probabilidad de obtener un número específico al lanzar un dado, la probabilidad de ganar la lotería al comprar un boleto, dos o diez o la probabilidad de lluvia al transcurrir la tarde.

Definición 1.4 Sea (Ω, \mathcal{A}) un espacio de medible. Una función $P : \mathcal{A} \rightarrow [0, 1]$ es llamada **medida de probabilidad** si satisface:

- $P(\Omega) = 1$
- P es una medida

Entre más probable es que un evento ocurra, mayor será el valor de su probabilidad en comparación con la de los demás posibles resultados. Note que, el hecho de que un evento tenga probabilidad cero no significa que éste sea imposible de ocurrir, sino que será un resultado menos probable de obtener

en relación con los otros.

La tripleta formada por $(\Omega, \mathcal{A}, \mu)$, es llamada *espacio de medida*, donde μ es una medida definida sobre el espacio medible (Ω, \mathcal{A}) . Cuando la medida definida sobre el espacio medible es una medida de probabilidad, entonces la tripleta (Ω, \mathcal{A}, P) se denomina *espacio de probabilidad*

Las siguientes propiedades se cumplen para cualquier espacio de probabilidad, las demostraciones pueden consultarse en el libro de Rocha Martínez [47].

- Sean $A_1, A_2 \in \mathcal{A}$. Si $A_1 \subset A_2$, entonces:

$$P(A_1) \leq P(A_2) \quad y \quad P(A_2 \setminus A_1) = P(A_2) - P(A_1)$$

- Sea A^c el complemento de A , entonces $P(A^c) = 1 - P(A)$
- Sea $A_1, A_2, \dots, A_n \in \mathcal{A}$ entonces,

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) \\ &\quad + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) \\ &\quad - \dots (-1)^{n+1} P(A_1 \cap A_2 \dots \cap A_n) \end{aligned}$$

Se dice que dos eventos, A y B , los cuales están definidos en algún espacio de probabilidad, con $P(B) > 0$, son *independientes* entre sí, si se cumple que:

$$P(A|B) = P(A)$$

Donde $P(A|B)$ es la *probabilidad condicional* de A dado B , definida como:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1.1)$$

Además, se tiene el siguiente teorema, cuya demostración puede consultarse en el libro de Protter (ver [32]).

Teorema 1.1 (Probabilidad total) *Sea $(B_n)_{n \geq 1}$ una partición finita de Ω , A un evento. Si $A \in \mathcal{A}$ entonces la probabilidad de A está dada por,*

$$P(A) = \sum_n P(A|B_n)P(B_n) \quad (1.2)$$

En un espacio medible (Ω, \mathcal{A}) , una *variable aleatoria* X es una función medible $X : (\Omega, \mathcal{A}) \rightarrow \mathbb{R}$, la cual establece un valor a los resultados obtenidos del fenómeno aleatorio.

Las variables aleatorias pueden clasificarse de acuerdo al tipo de función de distribución de probabilidad con la que se encuentren asociadas, tal es el caso de las variables aleatorias continuas, las absolutamente continuas y las discretas.

Se trabaja con variables aleatorias continuas cuando el contexto en el que se está desarrollando el fenómeno implica situaciones donde los valores que la variable aleatoria puede tomar se encuentran variando dentro de un intervalo. Por ejemplo, supongáse que se pide cronometrar el tiempo en que un atleta completa la carrera de los 100 metros planos. En cada repetición, el tiempo obtenido podría variar, es decir, la primera carrera posiblemente se logre terminar en 10.2s, mientras que la siguiente tal vez se complete en 9.63s o incluso 10.5s. Por lo que, hablar de la probabilidad de que la distancia se recorra antes de los 11 segundos tiene más sentido que pedir la probabilidad de terminarla en exactamente 10s. Con ello, se ilustra que la variable aleatoria $X := \text{“Tiempo en terminar la carrera”}$ toma valores dentro del intervalo $[9, 11]$.

Por el otro lado, se trabaja con variables aleatorias discretas cuando ésta solo puede tomar una cantidad a lo sumo numerable de valores distintos, tal es el caso de los lanzamientos de monedas.

Existe una función conocida como *función de distribución*, o en algunos casos *función de probabilidad acumulada*, la cual proporciona la probabilidad de aparición de los posibles valores asignados a su respectiva variable aleatoria.

Definición 1.5 (Función de distribución) Una función $F : \mathbb{R} \rightarrow [0, 1]$ se denomina *función de distribución* si cumple con lo siguiente:

- F es no decreciente y continua por la derecha con límites por la izquierda.
- $\lim_{t \rightarrow -\infty} F(t) = 0$
- $\lim_{t \rightarrow +\infty} F(t) = 1$

Un ejemplo de este tipo de funciones es:

$$F(x) = P(X \leq x), \quad x \in \mathbb{R} \tag{1.3}$$

y se conoce como la función de distribución de X .

Existen diversos tipos de distribuciones, a continuación se exponen las definiciones de algunas de las más usuales en la teoría de probabilidad, en especial las que serán esenciales en el desarrollo de esta tesis.

Definición 1.6 (Distribución normal) *Sea una variable aleatoria X , ésta tendrá una distribución normal, o también llamada **distribución gaussiana**, si su función de distribución esta dada por*

$$F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(v-\mu)^2}{2\sigma^2}} dv \quad (1.4)$$

donde μ representa la media o esperanza y σ la desviación típica o varianza, estos conceptos se presentan más adelante en esta sección. Usualmente se escribe $X \sim N(\mu, \sigma^2)$ para indicar que una variable X sigue este tipo de distribución.

Definición 1.7 (Distribución Log-normal) *Sea X una variable no negativa y Z una variable aleatoria definida como $Z := \ln X$. Se dice que X tiene una distribución **normal-logarítmica** con media μ y varianza σ si Z tiene una distribución normal con media μ y varianza σ . La función de probabilidad de esta distribución está definida como:*

$$f(x) = \begin{cases} \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[-\left(\frac{\ln x - \mu}{\sigma \sqrt{2}} \right)^2 \right] & \text{si } x > 0 \\ 0 & \text{de otra forma} \end{cases}$$

Definición 1.8 (Distribución binomial) *Supóngase un experimento aleatorio con dos únicos posibles resultados, cuyos eventos son independientes entre sí y para los cuales se tienen asociadas dos probabilidades: p la probabilidad de que el evento ocurra en una sola prueba o probabilidad de éxito y $q = 1 - p$, la probabilidad de que el suceso no ocurra o la probabilidad de fracaso. La función de probabilidad de este tipo de distribución está dada por:*

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

donde $x \in \{1, \dots, n\}$ representa el número de éxitos en las n veces en las que se repite el experimento.

Además de las funciones de distribución, también se encuentran las funciones de densidad. Las definiciones presentadas a continuación pueden encontrarse con mayor detalle en el libro de Feller [24] y en [47].

Definición 1.9 (Función de densidad discreta) Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es llamada **función de densidad discreta**, si cumple con lo siguiente:

- Para todo $x \in \mathbb{R}$, $0 \leq f(x) \leq 1$
- $\{x \in \mathbb{R} | f(x) \neq 0\}$ es un conjunto a lo sumo numerable.
- $\sum_{n=1}^{\infty} f(x_n) = 1$, $x_n \in \{x : f(x) \neq 0\}$

Definición 1.10 (Función de densidad continua) Se denomina **función de densidad continua** a una función, $f : \mathbb{R} \rightarrow \mathbb{R}^+ \cup \{0\}$, medible con respecto a la σ -álgebra de Borel, que satisface:

$$\int_{-\infty}^{+\infty} f(t)dt = 1$$

En las Definiciones 1.6 y 1.7 se pueden ver dos ejemplos de funciones de densidad.

Dada una variable aleatoria X , definida sobre (Ω, \mathcal{A}, P) , con función de distribución F . Se denomina, X o F como *absolutamente continuas* si y sólo si existe una función $f : \mathbb{R} \rightarrow \mathbb{R}$, de forma que:

$$F(x) = \int_{-\infty}^x f(t)dt$$

donde f es una función de densidad.

Note que, en este caso, por propiedades de la integral,

$$f(x) = F'(x) \tag{1.5}$$

para casi toda $x \in \mathbb{R}$. Es decir, existe $A \in \mathcal{B}(\mathbb{R})$, $\lambda(A^c) = 0$, y (1.5) se cumple para $x \in A$. Donde $\mathcal{B}(\mathbb{R})$ hace referencia a la σ -álgebra de Borel, la cual es la σ -álgebra generada por la colección de conjuntos abiertos en \mathbb{R} [47].

Ahora, se hablará sobre los conceptos de *esperanza* y los *momentos de orden superior*. La esperanza de una variable aleatoria indica un valor medio de los valores que ésta puede tomar cuando el fenómeno aleatorio sucede.

Definición 1.11 (Esperanza) Sea X una variable aleatoria con función de distribución de probabilidad, $f(x)$. La **esperanza o media** se define:

Si X es discreta:

$$\mu = E(X) = \sum_x xf(x) \quad (1.6)$$

Si X es absolutamente continua:

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx. \quad (1.7)$$

Además, se define la esperanza condicional como,

Definición 1.12 (Esperanza condicional) Sea (Ω, \mathcal{A}, P) un espacio de probabilidad, X una v.a. integrable. $\mathcal{G} \subset \mathcal{A}$ una sub- σ -álgebra. Se define la **esperanza condicional** de X dada \mathcal{G} como la única v.a $E(X|\mathcal{G})$, \mathcal{G} -medible e integrable, que satisface:

$$\int_A E(X|\mathcal{G})dP = \int_A XdP \quad \forall A \in \mathcal{G} \quad (1.8)$$

X y $E(X|\mathcal{G})$ integran lo mismo sobre conjuntos de \mathcal{G}

Cuando se hace mención a una familia de conjuntos \mathcal{C} se estará haciendo referencia a la σ -álgebra generada por éstos, es decir la σ -álgebra más pequeña que contiene a todos los conjuntos de \mathcal{C} . De igual forma, cuando se toma la esperanza condicional con respecto a una familia de variables aleatorias, se está haciendo referencia a la σ -álgebra más pequeña respecto a la cual todas las variables involucradas son medibles.

Teorema 1.2 Sea (Ω, \mathcal{A}, P) un espacio de probabilidad. Sean X y Y variables aleatorias integrables definidas en él y \mathcal{G} una sub- σ -álgebra de \mathcal{A} , entonces el valor esperado condicional satisface los siguientes puntos:

- $E(X|\mathcal{G}) \geq 0$ si $X \geq 0$ c.t.p.
- $E(c|\mathcal{G}) = c \quad c \in \mathbb{R}$
- $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G}) \quad a, b \in \mathbb{R}$
- $E(E[X|\mathcal{G}]) = E(X)$
- Si X es \mathcal{G} -medible, entonces $E(X|\mathcal{G}) = X$.

- Si X y \mathcal{G} son independientes, entonces $E(X|\mathcal{G}) = E(X)$
- Si Y es \mathcal{G} -medible y $E(|XY|) < \infty$ entonces $E(XY|\mathcal{G}) = YE(X|\mathcal{G})$
- Si $X \leq Y$ entonces $E(X|\mathcal{G}) \leq E(Y|\mathcal{G})$

La demostración del Teorema 1.2 puede consultarse en [13] y [36].

Definición 1.13 (Momento n-ésimo) Sea X una variable aleatoria. Se define el **momento n-ésimo** de X como:

$$\beta_n = E(X^n) \quad n \geq 1 \quad \text{si} \quad E(|X^n|) < \infty \quad (1.9)$$

Definición 1.14 (Momento central n-ésimo) Sea X una variable aleatoria. Se define el momento central n-ésimo de X como:

$$m_n = E([X - E(X)]^n) \quad n \geq 1 \quad (1.10)$$

El segundo momento central de X , en caso de existir, es conocido como *varianza* de X y es denotado como σ^2 .

Definición 1.15 (Covarianza) Sean X, Y variables aleatorias, se define la **covarianza** de X y Y como:

$$Cov(X, Y) = E(XY) - E(X)E(Y) \quad (1.11)$$

Para una $Cov(X, Y) = 0$, se dice que X y Y no están correlacionadas.

Teorema 1.3 (Ley de los grandes números) Sea X_n una sucesión de v.a independientes entre sí, con media μ y varianza σ^2 finitas. Entonces,

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{X_n}{n} - \mu \right| \geq \varepsilon \right) = 0 \quad (1.12)$$

Además,

$$\lim_{n \rightarrow \infty} \left(\frac{X_n}{n} \right) = \mu \quad (1.13)$$

A la ecuación (1.13) se le conoce como la **ley fuerte de los grandes números**.

Note que el valor esperado de una variable aleatoria es un promedio de los valores que ésta toma.

Lema 1.1 (Lema de Borel-Cantelli) *Sea $(A_n)_{n \in \mathbb{N}}$ una sucesión de eventos y $A = \limsup_{n \rightarrow \infty} A_n$, donde $\limsup_{n \rightarrow \infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{m \geq n} A_m$. Entonces,*

- Si $\sum_{n=1}^{\infty} P(A_n) < \infty$ entonces $P(A) = 0$
- Si los eventos que conforman $(A_n)_{n \in \mathbb{N}}$ son independientes, se cumple que si $\sum_{n=1}^{\infty} P(A_n) = \infty$ entonces $P(A) = 1$

Un fenómeno que evoluciona en el tiempo puede representarse por medio de un modelo, si la dinámica que rige este fenómeno está dada por el azar dicho modelo se conoce como un *proceso estocástico*.

Definición 1.16 (Proceso estocástico) *Sea (Ω, \mathcal{A}, P) un espacio de probabilidad. Se conoce como **proceso estocástico** a una función $X : \Omega \times [0, T] \rightarrow \mathbb{R}$, donde para cada $t \in [0, T]$, X_t es una variable aleatoria.*

Una definición equivalente de un proceso estocástico X es que se trata de una colección de funciones $\{X(\omega) : \mathbb{I} \rightarrow \mathbb{R}, \omega \in \Omega\}$ las cuales son conocidas como *trayectorias del proceso*. Donde $\mathbb{I} := [0, T]$.

En algunos casos es necesario estudiar secuencias que convergen a límites los cuales no son constantes sino variables aleatorias. De ello se hablará en el Capítulo 5. Las definiciones presentadas a continuación pueden consultarse para una mejor comprensión en el libro de Protter [32].

Definición 1.17 (Filtración) *Dado un espacio de probabilidad (Ω, \mathcal{A}, P) , se conoce como **filtración** $(\mathcal{F}_t)_{t \in [0, T]} = \{\mathcal{F}_t \subset \mathcal{F} : t \in [0, T]\}$ a una familia de σ -álgebras no decrecientes que cumplen:*

- $(\mathcal{F}_t)_{t \in [0, T]}$ es continua por la derecha.

$$\mathcal{F}_{t+} := \bigcap_{r > t} \mathcal{F}_r = \mathcal{F}_t$$

- $(\mathcal{F}_t)_{t \in [0, T]}$ es completa: (Ω, \mathcal{A}, P) es completo. Esto es, \mathcal{F}_t contiene a todos los conjuntos $A \in \mathcal{F}$ de medida cero.

A $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t \in [0, T]}, P)$ se le conoce como *espacio de probabilidad filtrado*.

Definición 1.18 (Martingala) Sea (Ω, \mathcal{A}, P) un espacio de probabilidad y sea $(\mathcal{A}_n)_{n \geq 0}$ una filtración en \mathcal{A} . Una secuencia de variables aleatorias $(X_n)_{n \geq 0}$ es llamada una **martingala** si:

- $E\{|X_n|\} < \infty$
- X_n es \mathcal{A}_n -medible para cada n .
- $E\{X_n | \mathcal{A}_m\} = X_m$ c.s. para toda $m \leq n$.

Definición 1.19 (Tiempo de paro) Sea $(\mathcal{A}_t)_{t \in \mathbb{I}}$ una filtración y $\tau : \Omega \rightarrow \mathbb{I} \cup \{+\infty\}$. τ es un **tiempo de paro** con respecto a la filtración (\mathcal{A}_t) , si el conjunto $\{\tau \leq t\}$ pertenece a la σ -álgebra \mathcal{A}_t , para cada $t \in \mathbb{I}$, donde \mathbb{I} es el conjunto de índices contenido, generalmente, en \mathbb{R}^+

El siguiente es un resultado importante para tiempos de paro, puede consultar la demostración y más detalles en el libro de Dudley (ver [15]) y el libro de Williams (ver [52]).

Teorema 1.4 (Teorema del paro Opcional) Sea $X := (X_n)_{n \geq 0}$ una martingala y sea τ un tiempo de paro. Si alguna de las siguientes situaciones se cumple:

- Para algún $N \in \mathbb{N}$, $\tau(\omega) \leq N \forall \omega$.
- Para algún $R \in \mathbb{R}^+$,

$$|X_n(\omega) - X_{n-1}(\omega)| \leq R \quad \text{para todo } (n, \omega)$$

y además $E(\tau) < \infty$.

entonces

$$E(X_\tau) = E(X_0)$$

Si X es una supermartingala entonces para los casos antes mencionados, se tiene que:

$$E(X_\tau) \leq E(X_0)$$

Definición 1.20 (Función de variación acotada) Sea una función $f : [a, b] \rightarrow \mathbb{R}$ definida en el intervalo $[a, b]$, se dice que f es de **variación acotada** si se cumple que:

$$\text{var}_1(f; [a, b]) := \sup \left\{ \sum_{i=1}^n |f(t_i) - f(t_{i-1})| : \pi = \{t_0 = a < t_1, \dots, < t_n = b\} \right.$$

es una partición de $[a, b]$ $\} < \infty$

Además, se define la variación cuadrática de f , cuando el límite existe, como:

$$\lim_{\|\pi\| \rightarrow 0} \sum_{i=1}^n (f(t_i) - f(t_{i-1}))^2$$

con

$$\|\pi\| = \max_{1 \leq i \leq n} (t_i - t_{i-1}) \rightarrow 0 \text{ cuando } n \rightarrow \infty$$

Definición 1.21 (Movimiento Browniano) Sea $B = \{B_t : t \in [0, T]\}$ un \mathcal{F}_t -proceso adaptado. Se dice que $(B, (\mathcal{F}_t)_{t \in [0, T]})$ es un **movimiento Browniano** si se cumplen los siguientes puntos:

- $B_0 = 0$ con probabilidad 1
- Para $t, s \in (0, T]$, $0 < s < t$, $B_t - B_s$ es independiente de \mathcal{F}_s
- Para todo $0 \leq s \leq t \leq T$, $B_t - B_s \sim N(0, t - s)$, es decir, $B_t - B_s$ tiene una distribución normal con media $\mu = 0$ y varianza $\sigma = t - s$ (ver Definición 1.8).
- B es un proceso continuo, i.e., $B(\omega)$ es una función continua para casi toda $\omega \in \Omega$.

Si se trata de un movimiento Browniano con valores en \mathbb{R}^n , es decir, multidimensional, entonces se hace referencia a un proceso estocástico de la forma:

$$B(t) = (B_1(t), B_2(t), \dots, B_n(t))$$

donde $B_1(t), B_2(t), \dots, B_m(t)$ son m movimientos Brownianos independientes unidimensionales.

El cálculo estocástico o también conocido como *cálculo de Itô* fue desarrollado por el matemático K. Itô en 1940 [14].

Un *proceso de Itô* es una expresión dada de la siguiente forma:

$$dX_t = f(t)dt + g(t)dB_t \quad (1.14)$$

la cual está definida para $t \in [0, T]$.

(1.14) puede describirse como el siguiente proceso estocástico:

$$X_t = X_a + \int_a^t f(s)ds + \int_a^t g(s)dB_s \quad (1.15)$$

donde X_a es una v.a \mathcal{F}_a -medible.

La primera integral en (1.15) es una integral de Riemann definida ω por ω , mientras la segunda se conoce como una *integral estocástica*.

De una forma muy general, la idea que Itô usó para definir la integral estocástica fue primero hacerlo para un proceso estocástico *simple adaptado* f_s , el cual está dado de la siguiente forma:

$$f_s = \sum_{n=0}^m f_{t_n} 1_{]t_n, t_{n+1}]}(s) \quad (1.16)$$

donde f_{t_n} es una variable aleatoria \mathcal{F}_{t_n} -medible. Además, $0 = t_0 < t_1 < \dots < t_{m+1} = T$ es una partición finita del intervalo $[0, T]$. El término $1_{]t_n, t_{n+1}]}$ se conoce como la función indicadora del intervalo $]t_n, t_{n+1}]$. Entonces, la integral estocástica para f_s con respecto a un movimiento Browniano $B = \{B_t, \mathcal{F}_t : t \in [0, T]\}$ se define como:

$$\int_0^T f_s dB_s := \sum_{n=0}^m f_{t_n} (B_{t_{n+1}} - B_{t_n}) \quad (1.17)$$

Lo siguiente en esta idea fue extender la integral estocástica a procesos un poco más generales, los cuales se piden que sean medibles y adaptados y que además cumplan con:

$$\int_0^T (f_s)^2 dt < \infty \quad \text{con probabilidad 1} \quad (1.18)$$

Si el proceso cumple lo anterior, entonces se puede encontrar una sucesión $\{f^{(m)} : m \geq 1\}$ de procesos simples adaptados tales que [38]:

$$\int_0^T (f_s^{(m)} - f_s)^2 ds \rightarrow 0 \quad \text{en probabilidad cuando } n \rightarrow \infty \quad (1.19)$$

Proposición 1.1 Sea f una función simple adaptada de la forma descrita en (1.16). Entonces para $N, c > 0$,

$$P \left[\left| \int_0^T f_s dB_s \right| > c \right] \leq \frac{N}{c^2} + P \left[\int_0^T (f_s)^2 ds > N \right] \quad (1.20)$$

La Proposición 1.1 implicaría que $\int_0^T (f_s^{(m)} - f_s)^2 ds$ es una sucesión convergente en probabilidad (ver [5]) y por lo tanto la sucesión de las integrales estocásticas de $(f_s^{(m)} - f_s)$ es una sucesión de Cauchy que converge en probabilidad, con lo que se puede enunciar la siguiente definición.

Definición 1.22 Sea f un proceso adaptado y medible, el cual cumple con (1.18) y dada una sucesión de procesos simples y adaptados $\{f^{(m)} : m \geq 1\}$ que cumplen con (1.19). La **integral estocástica** de f con respecto a un movimiento Browniano B se define como el límite en probabilidad de la sucesión $\{\int_0^T f_s^{(m)} dB_s : m \geq 1\}$.

Para los detalles del procedimiento de la construcción de la integral estocástica, además de las demostraciones necesarias, puede consultar [36], [38] y [46].

El siguiente teorema presentado es un resultado análogo al *Teorema fundamental del cálculo* y da una interpretación a la integral estocástica.

Teorema 1.5 Sea X_t un proceso de Itô dado por:

$$X_t = X_a + \int_a^t f(s) ds + \int_a^t g(s) dB_s \quad a \leq t \leq b$$

Sea $\phi : [a, b] \times \Omega \rightarrow \mathbb{R}$ una función continua la cual tiene derivadas parciales $\frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}, \frac{\partial^2 \phi}{\partial x^2}$. Entonces, $\phi(t, X_t)$ es un proceso de Itô y

$$\begin{aligned} \phi(t, X_t) &= \phi(a, X_a) + \int_a^t \frac{\partial \phi}{\partial x}(s, X_s) g(s) dB(s) \\ &+ \int_a^t \left[\frac{\partial \phi}{\partial t}(s, X_s) + \frac{\partial \phi}{\partial x}(s, X_s) f(s) + \frac{1}{2} \frac{\partial^2 \phi}{\partial x^2}(s, X_s) g(s)^2 \right] ds \end{aligned} \quad (1.21)$$

Usualmente la ecuación (1.21) se escribe de manera simbólica como:

$$d\phi(t, X_t) dt + \frac{\partial \phi}{\partial t}(t, X_t) dt + \frac{\partial \phi}{\partial x}(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2 \phi}{\partial x^2}(t, X_t) (dX_t)^2 \quad (1.22)$$

\times	dt	dB_t
dt	0	0
dB_t	0	dt

Tabla 1.1: Tabla de multiplicación de McKean

Note que en (1.21) aparece un término extra a los que se tendrían por el Teorema fundamental del cálculo, $\frac{1}{2} \frac{\partial^2 \phi}{\partial x^2}(t, X_t)(dX_t)^2$, esto se debe a que la variación cuadrática del movimiento Browniano $B(t)$ es distinta de cero [36].

Además, para obtener $(dX_t)^2$ se utiliza la Tabla 1.1. Para más detalles sobre este procedimiento y la demostración del Teorema 1.5 puede consultar el libro de Kuo (ver [36]).

Un tipo de proceso estocástico que ha sido estudiado ampliamente es conocido como *cadena de Markov*. Una cadena de Markov se caracteriza por ser un proceso aleatorio para el cual la probabilidad de ocurrencia de un evento futuro únicamente depende del evento presente y no de su pasado. Este concepto fue introducido por el matemático ruso Andrei Andreevich Markov en el año de 1906 [9] y actualmente es utilizado para representar y resolver problemas en diversas ramas como lo son: medicina, ingeniería, economía, entre otros.

Definición 1.23 (Cadena de Markov.) Sea $\{X_0, X_1, \dots\}$ una secuencia de v.a. que toman valores en algún conjunto contable S denominado como espacio de estados. Un proceso estocástico $\{X_n, n = 0, 1, 2, \dots\}$ se conoce como **Cadena de Markov** si para todos los estados $i_0, i_1, \dots, i_{n-1}, i, j \in S$, $S = \{1, 2, \dots, s\}$, se cumple que:

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) \quad (1.23)$$

Para todo $n \geq 1$

A (1.23) se le conoce como la *propiedad Markoviana*.

Si dados dos estados i y j existe una probabilidad mayor a cero de que partiendo de i se pueda llegar a j después de un número finito de pasos entonces se dice que j es accesible para i . Se dice que una cadena de Markov es *irreducible* si todos sus estados son accesibles entre sí. Además, se dice que

un estado es *recurrente* si al comenzar en un estado i la probabilidad de que el proceso vuelva a entrar alguna vez en él es 1, si esto no sucede, entonces el estado se dice *transitorio*.

La evolución de una cadena está dada por las *probabilidades de transición*, p_{ij} , [31].

$$p_{ij} = P(X_{n+1} = j | X_n = i)$$

Si la probabilidad de transición es independiente de n entonces la cadena de Markov es una cadena con probabilidades estacionarias la cual es llamada *homogénea*.

Una matriz T cuyas entradas están formadas por probabilidades de transición p_{ij} se conoce como *matriz de transición*.

$$T = \begin{pmatrix} p_{11} & p_{12} & \cdot & \cdot & \cdot & p_{1s} \\ p_{21} & p_{22} & \cdot & \cdot & \cdot & p_{2s} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{s1} & p_{s2} & \cdot & \cdot & \cdot & p_{ss} \end{pmatrix}$$

Teorema 1.6 *La matriz de transición T es una matriz estocástica, es decir, cumple con los siguientes puntos:*

- $p_{ij} \geq 0 \quad \forall i, j$, es decir, T no tiene entradas negativas.
- $\sum_j p_{ij} = 1 \quad \forall i$.

La prueba del Teorema 1.6 es evidente dado que, para el primer punto, p_{ij} es una probabilidad condicional y por propiedades de las mismas p_{ij} no puede ser negativa. Para el segundo punto, se tiene que

$$\sum_{j=0}^{\infty} p_{ij} = \sum_{j=0}^{\infty} P(X_{n+1} = j | X_n = i) = 1 \quad \text{para } i = 0, 1, 2, \dots$$

Definición 1.24 (Período) *Sea i un estado perteneciente a una cadena de Markov. El **período** de i está dado por:*

$$\lambda(i) := MCD\{n \geq 1 : T_{ii}^n > 0\} \tag{1.24}$$

donde *MCD* es el máximo común divisor. Además, se dice que una cadena con período λ es *periódica* si para todo i , $\lambda(i) = \lambda \geq 2$. Si $\lambda(i) = 1$ entonces la cadena es llamada *aperiódica*.

Definición 1.25 (Distribución estacionaria) Sea $\rho(i), i \in S$, una distribución de probabilidad. Sea $X_n, n \geq 1$ una cadena de Markov con espacio de estados S y matriz de transición T . Si

$$\sum_i \rho(i)p_{ij} = \rho(j), \quad j \in S. \quad (1.25)$$

entonces, se dice que ρ es una **distribución estacionaria** para la cadena X_n . La ecuación (1.25) puede escribirse de forma matricial como $\rho T = \rho$, la cual si se itera n veces, se obtiene que:

$$\rho T^n = \rho \quad \forall n \geq 0 \quad (1.26)$$

Teorema 1.7 Sea $(X_n)_{n \geq 0}$ una cadena de Markov aperiódica, recurrente e irreducible, S el conjunto de estados de dicha cadena. Entonces se cumple que:

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \rho_j, \quad \text{para todo par de estados } i, j$$

Además, $\rho_j, j \in S$, es una solución única del sistema, tal que

$$\rho_j = \sum_{h=1}^s \rho_h p_{hj}, \quad j = 1, 2, \dots, s \quad \sum_{h=1}^s \rho_h = 1$$

$\rho = (\rho_h)_{h=1}^s$ es la distribución estacionaria de la cadena de Markov.

La demostración del Teorema 1.7 puede consultarse en [31].

Las deficiones que sean específicamente necesarias para entender un apartado en particular se presentarán en los mismos.

Capítulo 2

Funciones de densidad e Histogramas de frecuencia.

En el capítulo 1 se ha mencionado que las variables aleatorias se encuentran clasificadas por su tipo de distribución de probabilidad, en esta sección se abordará información relacionada a las variables aleatorias continuas, es decir, a aquellas cuya función de distribución es continua.

Es importante mencionar que uno de los métodos para obtener la función de distribución de una variable aleatoria absolutamente continua se basa en integrar la función de densidad que le corresponde (ver Definición 1.10).

Ahora bien, los histogramas son representaciones gráficas, específicamente diagramas de barras, utilizados en la descripción de datos numéricos, los cuales se encuentran clasificados por sus respectivas distribuciones de frecuencia, donde la cantidad de muestras obtenidas por clase se encuentra representada sobre el eje vertical, mientras que en el eje horizontal se posicionan los valores de la variable aleatoria [11].

La relación existente entre un histograma de frecuencias y la función de distribución es notoria, puesto que, entre más adecuados sean los rangos seleccionados para la representación más clara será la distribución de frecuencias en el histograma. Ahora bien, si se pide que el valor de la suma de áreas de los bloques en el histograma sea cercano o igual al valor de la unidad, entonces podrá observarse cómo la forma de éste se aproxima más a una curva que describe la densidad.

Uno de los principales motivos por los cuales es importante conocer este tipo de aplicación radica en que, por lo general, en la práctica las funciones de densidad o distribución no se encuentran disponibles explícitamente pero

sí los datos de donde ésta puede ser aproximada o ajustada.

El caso de la aproximación de una función de densidad por medio histogramas de frecuencia es posible debido a *La ley de los grandes números*, específicamente *La ley fuerte de los grandes números*, véase el Teorema 1.3. Este teorema propone que al realizar un promedio de las variables aleatorias, éste convergerá al valor de una constante, dependiendo de determinadas circunstancias. Básicamente, proporciona una justificación teórica para el proceso de promediar que se emplea en los experimentos para obtener la precisión de las mediciones [41].

En los siguientes párrafos se expone el contexto requerido para entender uno de los ejemplos de aplicación de estas técnicas.

El sol es un emisor de radiaciones del espectro electromagnético compuestas por luz, calor y radiación UV. Esta última se encuentra clasificada por las distintas formas en las que llega a la Tierra, es decir, radiaciones: UV-C, UV-B, UV-A, y aunque la más peligrosa se trata de la radiación UV-C, ésta es absorbida en gran medida por la atmósfera. Sin embargo, es la radiación UV-B la que tiene efectos importantes, por pequeños incrementos en las dosis, sobre la superficie de la piel y los ojos, siendo la primera causa de quemaduras, lesiones en la cornea, bronceado, entre otras enfermedades más fuertes, como es el cáncer en la piel.

Los tonos de piel más claros son los más susceptibles a este tipo de enfermedades y es necesario protegerlas para evitar daños severos, incluso aún cuando se encuentran bajo techo. Ahora bien, existe una unidad llamada Dosis Erimática Mínima (MED), que es el valor de la mínima irradiación necesaria para producir enrojecimiento en la piel, tras un determinado tiempo de exposición y que toma como unidad de referencia a los individuos con piel blanca ligeramente pigmentada.

Ejemplo 2.1 *Supóngase que una persona de tez morena debe realizar un trabajo de campo en Chalco, Edo. de México, en el cual se encuentra expuesto al sol durante el medio día. La empresa que lo contrata desea saber si es indispensable adquirir protección solar para esta persona o si supondría un gasto innecesario, por lo cual se requiere saber que tan probable es que los valores de la radiación superen los valores recomendados.*

Para el desarrollo de este ejemplo se tomará la información relacionada con la radiación UV-B registrada durante el año 2019 en la estación correspondiente a Chalco (véase [28]) y los valores equivalentes de riesgo para cada piel presentados en la Tabla 2.1.

Energía MED necesaria para generar eritema por tipo de piel	
Tipo de piel	MED
Muy clara	0.8
Clara	1
Morena	1.25
Morena Oscura	1.56
Oscura	1.95
Muy Oscura	2.44

Tabla 2.1: MED de valores promedios.

Tabla de frecuencias 12:00 horas		
Intervalo	Frecuencia absoluta	Frecuencia Relativa
(1,1.5]	7	0.02095
(1.5,2]	16	0.04790
(2,2.5]	30	0.08982
(2.5,3]	48	0.14371
(3,3.5]	82	0.24550
(3.5,4]	92	0.27544
(4,4.5]	54	0.16167
(4.5,5]	5	0.01497

Tabla 2.2: Tabla de frecuencias del Ejemplo 2.1

De acuerdo a los datos registrados durante 2019 para las 12:00 horas , se ha obtenido la tabla 2.2.

El problema pide el cálculo de una probabilidad, lo que implica que es necesario obtener la función de distribución o la función de densidad que describen los datos de la Tabla 2.2. Para ello se utilizará una aproximación basada en la ley de los grandes números, cuyos pasos generales se exponen a continuación.

Dado un intervalo $[a, b]$, donde $a < b$, y una sucesión de variables aleatorias independientes. El intervalo se divide en subintervalos I_k , y se define la siguiente sucesión:

$$Y_k^n = \frac{1}{n} \sum_{i=1}^n 1_{I_k}(X_i) \quad (2.1)$$

Aquí $\{X_i : i \in \mathbb{N}\}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas que representan los resultados del fenómeno aleatorio bajo estudio. La familia de variables aleatorias $I_k(X_i)$ con $i \in \{1, \dots, n\}$, posee una distribución de Bernoulli que toma valor 1 para la probabilidad de éxito y 0 para la probabilidad de fracaso, con parámetro $p_k = P(X_1 \in I_k)$, con lo que el Teorema 1.3 implica que $Y_k^n \rightarrow p_k$ c.s cuando $n \rightarrow \infty$.

Ahora, por la definición de función de densidad y la ley fuerte de los grandes números, se tiene que:

$$\lim_{n \rightarrow \infty} Y_k^n = \int_{I_k} f(x) dx \quad \text{c.s} \quad (2.2)$$

De forma que, realizando la suposición de que los intervalos I_k son lo suficientemente pequeños, además de que f es continua, se obtiene:

$$Y_k^n \approx f(a_k) \int_{I_k} dx \quad (2.3)$$

donde a_k es un punto perteneciente a I_k .

Con lo que finalmente se tiene:

$$f(x) \approx \frac{1}{\lambda(I_k)n} \sum_{i=1}^n 1_{I_k}(X_i), \quad x \in I_k \quad (2.4)$$

donde $\lambda(I_k)$ es la longitud del intervalo I_k . El procedimiento detallado y la demostración correspondiente puede consultarse en [50].

Aplicando lo anterior al Ejemplo 2.1, se se puede observar que:

- El intervalo $[a, b]$ representa las medidas MED/hora.
- La longitud del intervalo I_k seleccionado es de 0.5, por lo tanto $k \in \{1, 1.5, \dots, 5\}$
- El valor de n es igual a la cantidad total de mediciones registradas en el año 2019, en este caso 334.

Por lo que la aproximación (2.4) se reduce a

$$f(x) \approx \frac{1}{334(0,5)} \sum_{i=1}^{334} 1_{I_k}(X_i) \quad (2.5)$$

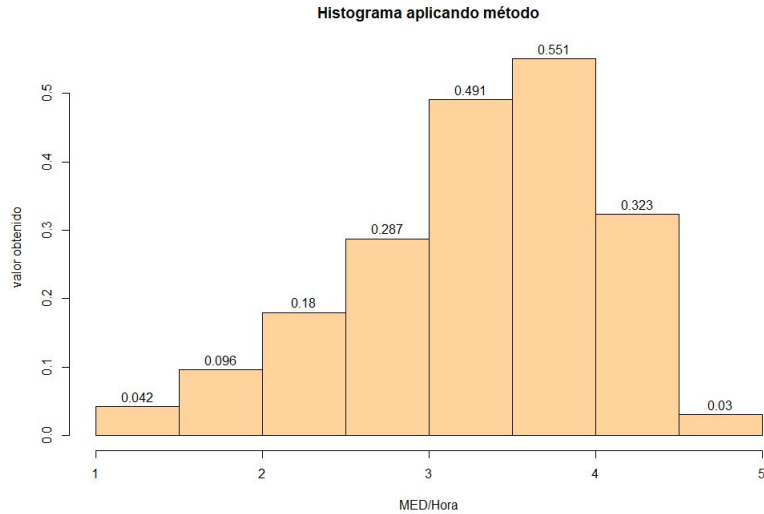


Figura 2.1: Aproximación de la función de densidad por histograma.

Note que el resultado obtenido al evaluar (2.5) es el valor que será utilizado al trazar la altura de cada barra del histograma en el correspondiente intervalo. La Figura 2.1 muestra lo obtenido para este caso.

El histograma obtenido refleja una aproximación de la forma de la función densidad. Ahora, con la ayuda visual que ha proporcionado esta herramienta, se ajustarán estos datos a una distribución específica.

Por la Figura 2.1 se puede suponer que los datos describen una distribución de tipo *Weibull*, cuya función de densidad está dada por:

$$f(x) = \begin{cases} abx^{b-1}e^{-ax^b}, & x > 0, \\ 0, & x \leq 0, \end{cases} \quad (2.6)$$

donde $b > 0$, $a > 0$ representan los parámetros de forma y escala respectivamente.

Utilizando el programa estadístico *R* se han obtenido los valores aproximados para los parámetros que podrían describir esta distribución, tal que, $b = 5.3974$, $a = 3.5931$. El siguiente paso es evaluar si la suposición realizada con respecto a la distribución ha sido correcta.

Una *prueba de bondad de ajuste* se emplea para decidir cuándo un conjunto de datos se apega a una distribución de probabilidad dada [18]. Estas pruebas evalúan la diferencia existente entre los valores reales y los valores

esperados descritos por el modelo. Existen pruebas correspondientes para distribuciones continuas o discretas.

Para realizar estas pruebas se necesitan especificar dos hipótesis sobre el comportamiento de la distribución, generalmente denotadas por H_0 y H_1 . La primera suele llamarse *hipótesis nula* y es aquella que se somete a evaluación. En caso de ser rechazada, H_1 se toma como verdadera.

Una vez que se tienen las hipótesis, se especifica el *nivel de significancia*, el cual es un valor fijo y que por lo general es elegido como 0.05 o 0.1. El valor de significancia se utiliza para comparar el *valor-p*, el cual es la probabilidad calculada al asumir que la hipótesis a probar es cierta. La forma de obtener el valor-p dependerá de la prueba elegida. Para detalles concretos de pruebas específicas, se puede consultar en el libro de Canavos [18].

Para este caso se ha utilizado la prueba de bondad de ajuste conocida como *Anderson-Darling*. La cual es una estadística de prueba definida como:

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{\{F_n(x) - F_0(x)\}^2}{\{F_0(x)\}\{1 - F_0(x)\}} dF_0(x) \quad (2.7)$$

donde F_n es la función de distribución empírica, definida como:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{[X_i \leq x]}, \quad (2.8)$$

$I_{[X_i \leq x]}$ es la función indicadora del evento $[X_i \leq x]$ y F_0 la función de distribución hipotética.

Lo anterior es equivalente a:

$$A_n^2 = -n - \frac{1}{n} \sum_{i=1}^n [(2i - 1) \log z_i + (2n + 1 - 2i) \log(1 - z_i)] \quad (2.9)$$

donde $z_i = F_0(x_i)$

Generalmente se suele emplear (2.9) en las pruebas de bondad.

Prosiguiendo con la prueba de bondad para este caso, se plantean las siguientes dos hipótesis:

- H_0 : los valores de radiación, MED/Hora, registrados para las 12:00 hrs en el 2019 siguen una distribución Weibull con parámetros de forma y escala, $b = 5.3974$, $a = 3.5931$, respectivamente.

```

Anderson-Darling test of goodness-of-fit
Null hypothesis: weibull distribution
with parameters shape = 5.39741044430762, scale = 3.59319134288359

data: med_hora
An = 2.3565, p-value = 0.05901

```

Figura 2.2: Resultados obtenidos para la prueba de bondad Anderson-Darling

- H_1 : los valores de radiación, MED/Hora, registrados para las 12:00 hrs en el 2019 *no* siguen una distribución Weibull con parámetros de forma y escala, $b = 5.3974$, $a = 3.5931$, respectivamente.

Nivel de significancia: 0.05

Posteriormente, con ayuda del programa de estadística R , se ha calculado el valor de la prueba Anderson, y el valor-p asociado a ésta, obteniendo como resultado, *valor-p* = 0.05901. La Figura 2.2 presenta la información entregada por el programa.

Lo anterior se resume en que el valor-p es mayor al nivel de significancia y por lo tanto la hipótesis nula no se rechaza, por lo tanto se concluye que los datos siguen una distribución Weibull, con los parámetros antes mencionados. En la Figura 2.3 puede observarse la comparación del área que describe la distribución contra la del método anterior.

Una vez encontrada la función que describe la densidad del conjunto de datos, se puede realizar el cálculo de cualquier probabilidad.

De acuerdo a la Tabla 2.1, el sujeto que realizará el trabajo puede exponerse a no más de 1.25 MED antes de que esta exposición le genere algún problema de salud. Entonces, sería prudente calcular cuál es la probabilidad de que a las 12:00 horas la radiación sea mayor a 1.25 MED/hora.

$$P(X > 1.25) = \int_{1.25}^{\infty} abx^{b-1}e^{-ax^b} dx \approx 0.9966 \quad (2.10)$$

Por lo tanto existe una probabilidad de 99 % de que la radiación sea mayor a 1.25 MED/hora. Por último, supongamos que aún así la empresa solicita que se realice el cálculo para un valor de radiación que evidentemente causará daño, es decir, se necesitaría calcular la probabilidad de que la radiación sea mayor a 3.4 MED. Con lo que se obtiene:

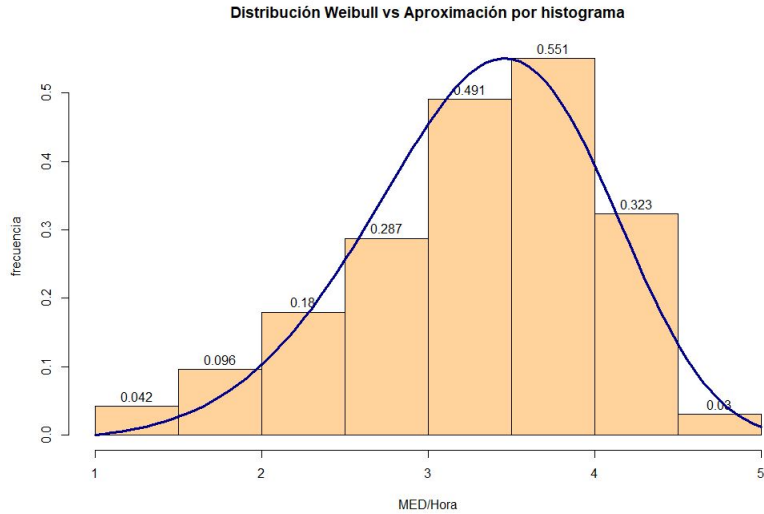


Figura 2.3: Comparación de la distribución Weibull con parámetros $b = 5.3974$, $a = 3.5931$ y el histograma obtenido aplicando la aproximación de densidad por la ley fuerte de los grandes números.

$$P(X > 3.4) \approx 0.476063 \quad (2.11)$$

Es decir, la probabilidad de que la radiación sea mayor a 3.4 MED es de 47% lo cual sigue representando un valor de ocurrencia alto.

Es importante mencionar que el método de aproximación por histograma no se restringe a una distribución en específico. Por lo que para concluir este apartado se desea mostrar una aplicación donde los datos sigan otro tipo de distribución. El siguiente ejemplo ha sido tomado del libro de Sanz i Solé [50].

Ejemplo 2.2 *Los tiempos presentados en la Tabla 2.3 se han obtenido de la observación de la duración (en horas) de un componente electrónico de un sistema de alarma. Se desea aproximar el ajuste de tiempo a una función de distribución de probabilidad.*

De la Tabla 2.3 se obtiene la tabla de frecuencias 2.4, donde los intervalos elegidos tienen una longitud de 20. Con lo que el histograma quedaría como el mostrado en la Figura 2.5. Es evidente que el tipo de distribución que

Tiempo de vida en horas									
262.8	1	36.4	4	59.4	35.3	70.5	22.6	3.7	5.8
32.1	0.5	17.4	77.6	46.7	182.4	76.7	3.5	13.4	29.7
6.1	15.1	110.5	45.9	31.7	22.4	27.8	10	33	26.7
8	6.8	63	70.9	30	12.2	29.6	3.3	32.2	12.3
128.2	24.6	7	39.8	71.1	19.4	5.4	4.4	54.4	24.8

Tabla 2.3: Tiempo de vida en hora de 50 componentes electrónicos.

muestran estos datos se basa en una distribución exponencial. La cual tiene una función de densidad como la siguiente:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0 & \text{cualquier otro caso} \end{cases} \quad (2.12)$$

De igual forma que en el Ejemplo 2.1, se calculan los parámetros para la distribución supuesta y se realiza la prueba de bondad de ajuste.

El valor esperado de una variable aleatoria, X , con una distribución exponencial está dada por:

$$E(X) = \frac{1}{\lambda} \quad (2.13)$$

Como consecuencia del Teorema 1.3, realizando el cálculo de la esperanza, se puede obtener el parámetro λ de la ecuación (2.13).

Para este caso, la esperanza de los datos tiene un valor de $E(X) = 39.222$, con lo que $\lambda = 0.02549$.

Se proponen las hipótesis y el nivel de significancia como:

- H_0 : los valores registrados en la tabla 2.3 siguen una distribución esponencial con parámetros $\lambda = 0.02549$.
- H_1 : los valores registrados en la tabla 2.3 no siguen una distribución esponencial con parámetros $\lambda = 0.02549$.

Nivel de significancia: 0.05

Utilizando el programa R , se ha realizado la prueba de Anderson-Darling, obteniendo un $valor-p=0.7948$. La Figura 2.4 muestra los datos obtenidos.

El valor-p es mayor al valor de significancia pedido, por lo tanto, se puede concluir que los datos siguen una distribución exponencial con parámetro $\lambda = 0.02549$. En la Figura 2.6 se puede observar la comparación entre la distribución propuesta y la aproximación por histograma.

Tabla de frecuencias de componentes		
Intervalo	Frecuencia absoluta	Frecuencia Relativa
(0,20]	20	0.40
(20,40]	16	0.32
(40,60]	4	0.08
(60,80]	6	0.12
(80,100]	0	0
(100,120]	1	0.02
(120,140]	1	0.02
(140,160]	0	0
(160,180]	0	0
(180,200]	1	0.02
(200,220]	0	0
(220,240]	0	0
(240,260]	0	0
(260,280]	1	0.02

Tabla 2.4: Tabla de frecuencias para el Ejemplo 2.2

```

Anderson-Darling test of goodness-of-fit
Null hypothesis: exponential distribution
with parameter rate = 0.0254958951608791

data: horas
An = 0.4525, p-value = 0.7948

```

Figura 2.4: Datos obtenidos de la prueba Anderson-Darling en el programa R.

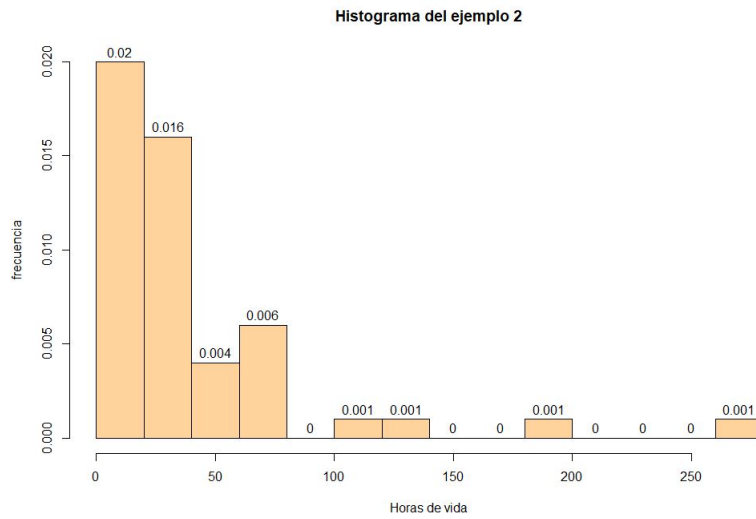


Figura 2.5: Histograma Ejemplo 2

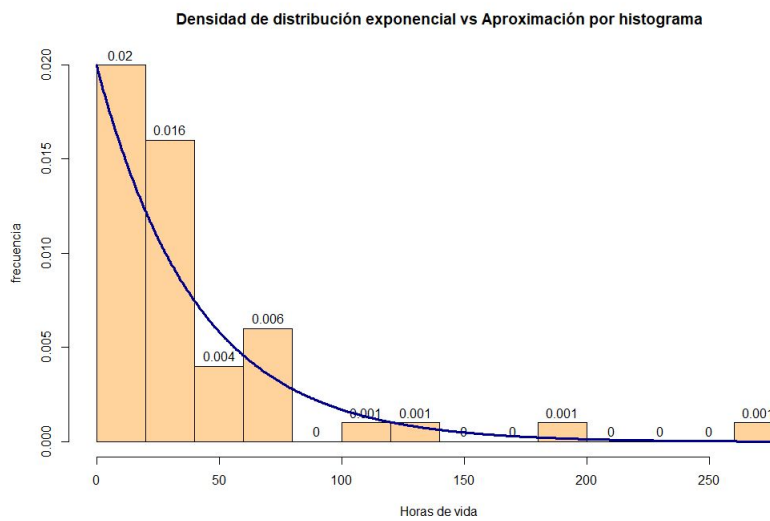


Figura 2.6: Comparación de densidad de distribución exponencial contra aproximación por histograma.

Capítulo 3

Filtro de Kalman

En el desarrollo de esta sección se retomará el concepto dado en el Capítulo 1 sobre esperanza condicional (ver Definición 1.12).

Se conoce como un *sistema dinámico* a un proceso que evoluciona en el tiempo. Analizar un sistema dinámico implica el estudio, a lo largo del tiempo, de las variables medibles y observables del sistema para saber cómo afectan el comportamiento de éste [44].

Al conjunto más pequeño de parámetros en $t = t_0$ que en conjunto con el conocimiento de la entrada para $t \geq t_0$, determina por completo el comportamiento del sistema para cualquier tiempo $t \geq t_0$, se le denomina *estado del sistema* [42]. A los parámetros que componen este dicho grupo se les conoce como *variables de estado*.

Las variables de estados forman un espacio n-dimensional al cual se le denomina *espacio de estados*. Un estado puede ser representado mediante un punto en el espacio de estados [42].

Un sistema en tiempo discreto es descrito por el modelo de estados formado por (3.1) y (3.2).

$$x_{k+1} = A_k x_k + b_k + \xi_k \quad (3.1)$$

$$y_k = H_k x_k + h_k + w_k, \quad k = 1, 2, \dots \quad (3.2)$$

donde

$x_k \in \mathbb{R}^n$ es el estado actual del proceso.

$y_k \in \mathbb{R}^m$ es la salida actual disponible.

$b_k \in \mathbb{R}^n$ y $h_k \in \mathbb{R}^m$ son vectores con información disponible del modelo y la salida respectivamente.

$A_k \in \mathbb{R}^{n \times n}$, $H_k \in \mathbb{R}^{m \times n}$ son matrices deterministas.

$\xi_k \in \mathbb{R}^n$ y $w_k \in \mathbb{R}^m$ son vectores aleatorios independientes entre sí, definidos en un espacio de probabilidad (Ω, \mathcal{A}, P) , los cuales tienen una distribución gaussiana con media cero y covarianza Θ_k y W_k respectivamente (ver Def. (1.6) y (1.15)). Es decir,

$$\begin{aligned} E\{\xi_k\} &= 0, & E\{w_k\} &= 0 \\ E\{\xi_k \xi_k^T\} &= \Theta_k, & E\{w_k w_k^T\} &= W_k \end{aligned}$$

A ξ_k se le conoce como el ruido del estado, mientras que w_k representa el ruido de la salida del sistema.

En la práctica no siempre es posible o incluso conveniente obtener mediciones de cada una de las variables que se desean controlar, por lo que, en la mayoría de los casos, se busca una manera de conseguir un valor lo más cercano posible al real.

El filtro de Kalman, desarrollado en los últimos años de la década de los cincuenta por Rudolf E. Kalman [4], se trata de una herramienta teórica aplicada a problemas de estimación en tiempo real, se utiliza dentro de diversos campos de la ingeniería y proporciona un marco de referencia para inferir las variables no medibles desde mediciones indirectas y ruidosas [4]. También es usado para la predicción del curso de sistemas dinámicos que no pueden ser controlados.

Dado un proceso inobservable $x = x(t, \omega)$ y un proceso observable $y = y(s, \omega)$, los cuales se encuentran distribuidos conjuntamente en un espacio de probabilidad (Ω, \mathcal{A}, P) . Este filtro, obtiene un estimado del proceso que no puede observarse al tiempo $(t + \tau)$, $\hat{x}(t + \tau, \omega)$, el cual es óptimo con respecto al criterio de mínimos cuadrados, y que es posible obtener por medio de mediciones de $y = y(s, \omega)$ en $[0, T]$. Las demostraciones e información complementaria pueden encontrarse en el libro de Poznyak [45]

Proposición 3.1 *Sea X una variable aleatoria cuadrado integrable y \mathcal{G} una σ -álgebra. Se tiene que,*

$$\arg \min_{Y \sim \mathcal{G}} E((X - Y)^2) = E(X|\mathcal{G}) \quad (3.3)$$

donde $Y \sim \mathcal{G}$ es la familia de todas las variables aleatorias cuadrado integrables y medibles con respecto a \mathcal{G} . Recuerde que los elementos de \mathcal{G} se interpretan como la información del sistema a la cual se puede tener acceso.

La igualdad mostrada en la Proposición 3.1 hace referencia a que el mínimo de $\{E((X - Y)^2) : Y \sim \mathcal{G}\}$ se alcanza en $E(X|\mathcal{G})$ [45] y por lo tanto, se puede concluir que es el mejor estimador de la variable X dada la información obtenida \mathcal{G} .

El algoritmo del Filtro de Kalman se compone de dos secciones importantes: las ecuaciones de predicción y las de corrección, las cuales se presentan a continuación. La demostración puede verse en [45].

Teorema 3.1 (Filtro de Kalman) Sean A_k, H_k matrices deterministas, $\{\xi_i, W_j : i, j \in \mathbb{N}\}$ una familia gaussiana de variables aleatorias independientes con media cero y b_k, h_k medibles con respecto a \mathcal{G}_k . Se tiene que:

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(y_k - \tilde{y}_{k|k-1}) \\ \hat{x}_{k|k-1} &= A_{k-1}\hat{x}_{k-1|k-1} + b_{k-1} \\ \tilde{y}_{k|k-1} &= H_k\hat{x}_{k|k-1} + h_k\end{aligned}$$

donde

$$\begin{aligned}K_k &= S_k^T D_k^{-1} \\ S_k &= H_k(\Theta_{k-1} + A_{k-1}P_{k-1}A_{k-1}^T) \\ D_k &= W_k + S_k H_k^T \\ P_k &= A_{k-1}P_{k-1}A_{k-1}^T + \Theta_k - S_k^T D_k^{-1} S_k\end{aligned}$$

$$\bar{x}_{0|0} = E\{x_0\}, \quad P_0 = cov\{x_0\}$$

$$\hat{x}_{k|k} := E(x_k|\mathcal{G}_k); \quad \hat{x}_{k|k-1} := E(x_k|\mathcal{G}_{k-1})$$

$\hat{x}_{k|k-1}$ se interpreta como el valor del estado actual con respecto a las observaciones del estado anterior. De forma similar, $\tilde{y}_{k|k-1}$, se interpreta como la salida actual con respecto a la salida del estado anterior.

Las aplicaciones más populares de este algoritmo se basan en el monitoreo y control de sistemas dinámicos tales como procesos de manufactura continua, aeronaves o navíos [4], además de procesamiento de señales biológicas, la predicción de variables económicas, entre otras. El Ejemplo 3.1 es una aplicación unidimensional del filtro de Kalman.

Ejemplo 3.1 *Se requiere medir la temperatura de un líquido dentro de un contenedor industrial. Se asume que la temperatura de este líquido es constante sin embargo, no se descarta el hecho de que existan pequeños cambios. Las mediciones se toman cada segundo y se considera que el sensor con el que se han realizado tiene un error cuya desviación estándar es de 0.1 °C.*

El sistema del Ejemplo 3.1 puede expresarse como

$$x_k = T + w_k \quad (3.4)$$

donde

T es la temperatura constante.

w_k representa el ruido del proceso, de forma que, $w_k \sim N(0, \sigma_p)$.

σ_p es la desviación estándar del proceso, que se supondrá como 0.0001. Además,

$$y_k = T + v_k \quad (3.5)$$

donde

v_k representa el ruido de medición, $v_k \sim N(0, \sigma_m)$.

σ_m es la desviación estándar de las mediciones, que se supondrá como 0.1

Suponga que la temperatura real del líquido ronda cerca de los 250 °C y que durante 20 iteraciones los valores del comportamiento real del sistema son los encontrados en la Tabla 3.1

Ahora, considere la Tabla 3.2 como los valores medidos por el sensor en dichas iteraciones. Los valores de las Tablas 3.1 y 3.2 fueron simulados con ayuda del programa Matlab.

Para poder realizar el filtro de Kalman se considera lo siguiente: El error de medición se supone con una desviación típica $\sigma_m=0.1$, con lo que la varianza del error de medición es $\sigma_m^2=0.01$

Suponga que al no conocer la temperatura real, se hace la suposición de que el líquido se encuentra inicialmente a 50°C, con lo que $X_0=50$ y se considera una $P_0=10$. Con lo que, al aplicar el algoritmo del filtro de Kalman, se obtienen los valores mostrados en la Tabla 3.3.

En la Figura 3.1 puede observarse la comparación gráfica de la simulación de los valores reales, los valores obtenidos por mediciones y las aproximaciones realizadas.

Valores del sistema			
Iteración	Temperatura	Iteración	Temperatura
1	250.0066	11	250.0145
2	249.9930	12	250.0138
3	250.0050	13	250.0010
4	250.0099	14	249.9957
5	249.9931	15	250.0051
6	250.0099	16	249.9934
7	249.9931	17	249.9987
8	249.9914	18	249.9947
9	250.0005	19	250.0011
10	249.9934	20	250.0113

Tabla 3.1: Temperaturas reales del sistema para el Ejemplo 3.1

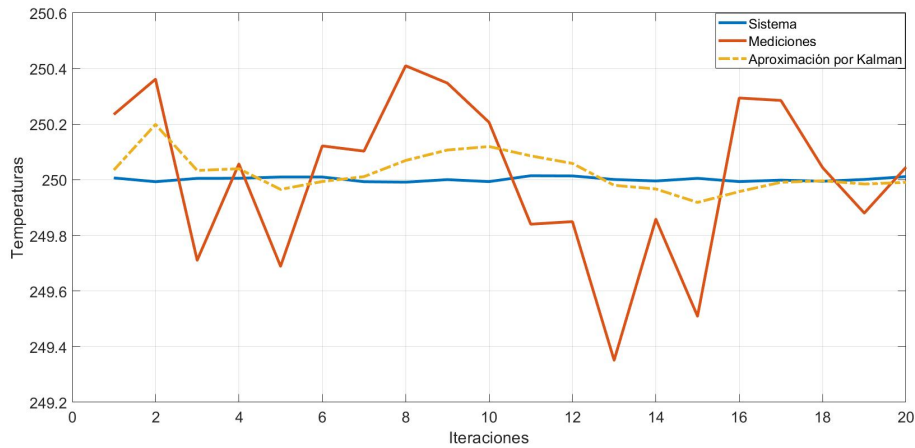


Figura 3.1: Comparación de los valores reales, medidos y aproximados del Ejemplo 3.1

Ejemplo 3.2 *Se desea conocer la posición y la velocidad de un vehículo cuyas mediciones son imprecisas. Considere que se trata de un vehículo sobre una vía sin fricción y que inicialmente éste se encuentra en reposo. Ahora, suponga que una aceleración aleatoria no controlable es ejercida sobre él y que se continua de esta manera. La medición de la posición es tomada cada*

Valores medidos del sistema			
Iteración	Temperatura	Iteración	Temperatura
1	250.2348	11	249.8403
2	250.3616	12	249.8495
3	249.7107	13	249.3512
4	250.0569	14	249.8582
5	249.6890	15	249.5095
6	250.1217	16	250.2940
7	250.1030	17	250.2852
8	250.4099	18	250.0437
9	250.3476	19	249.8803
10	250.2066	20	250.0452

Tabla 3.2: Temperaturas medidas del sistema para el Ejemplo 3.1

Δt segundos.

Las componentes del vector X representan la posición x y la velocidad \dot{x} del vehículo y dada la información anterior, se desea obtener los valores de dichas componentes. Así,

$$X = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$$

Suponga que, entre los lapsos de tiempo $(k - 1)$ y k la aceleración a_k es una variable aleatoria normalmente distribuida, de forma que $a_k \sim N(0, \sigma_p)$.

El sistema puede describirse de la siguiente manera

$$X_k = AX_{k-1} + Ga_k \quad (3.6)$$

donde

$$A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad G = \begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \end{bmatrix} \quad (3.7)$$

Las matrices A , G son constantes. Además, note que no existen entradas de control como tal sino que la aceleración a_k puede considerarse como una entrada que agrega ruido al sistema y que por consecuente afecta al vector X . Con lo que el sistema puede representarse de la siguiente forma

Aproximación por Kalman			
Iteración	Temperatura	Iteración	Temperatura
1	250.0248	11	250.0760
2	250.1889	12	250.0488
3	250.0236	13	249.9701
4	250.0296	14	249.9569
5	249.9554	15	249.9086
6	249.9837	16	249.9475
7	250.0013	17	249.9809
8	250.0594	18	249.9863
9	250.0969	19	249.9747
10	250.1095	20	249.9807

Tabla 3.3: Aproximaciones de la temperatura por filtro de Kalman para el Ejemplo 3.1

$$X_k = AX_{k-1} + \xi_k \quad (3.8)$$

donde

$$\xi_k \sim N(0, Q)$$

Q es de la forma:

$$Q = GG^T \sigma_p^2 = \begin{bmatrix} \frac{\Delta T^2}{4} & \frac{\Delta T^3}{2} \\ \frac{\Delta T^3}{2} & \Delta T^2 \end{bmatrix} \sigma_p^2 \quad (3.9)$$

Recuerde que, las mediciones que se realizan de la posición del vehículo también se encuentran perturbadas por ruido w_k , de igual forma, suponga que se tiene una distribución normal no precisamente centrada, $w_k \sim N(0, \sigma_m)$. Con lo que,

$$Y_k = HX_k + w_k \quad (3.10)$$

donde

$$H = [1 \quad 0]$$

A partir de lo anterior se pueden realizar las aproximaciones por el método de Kalman. Suponga que las mediciones de posición y velocidad se realizan cada 0.1 segundos, esto es $\Delta t = 0.1$. Suponga además que $\sigma_p = 2$ y que $\sigma_m = 5$

Considere la Tabla 3.4 como los valores reales del sistema y la Tabla 3.5 como las mediciones realizadas. Por cuestiones de espacio, en ambas tablas, se presentan únicamente las primeras veinte iteraciones de las cien realizadas.

Valores reales del sistema					
Iteración	Posición	Velocidad	Iteración	Posición	Velocidad
1	0	0	11	-2.7677	0.6721
2	0.5606	0.2232	12	10.1436	0.9779
3	4.1862	0.1788	13	-3.2812	1.1218
4	-3.4838	0.2178	14	0.5514	1.0432
5	3.1490	0.2761	15	-0.4285	0.7875
6	-1.6356	0.0290	16	6.2489	0.8478
7	3.8069	0.3532	17	3.2350	1.0746
8	4.9267	0.2773	18	-11.5889	1.0476
9	9.7454	0.3705	19	7.4804	0.7561
10	-4.2009	0.3009	20	-6.0884	1.1600

Tabla 3.4: Valores de la posición y velocidad del sistema para el Ejemplo 3.2

Mediciones del sistema			
Iteración	Posición	Iteración	Posición
1	-0.5477	11	-2.7677
2	0.5606	12	10.1436
3	4.1862	13	-3.2812
4	-3.4838	14	0.5514
5	3.1490	15	-0.4285
6	-1.6356	16	6.2489
7	3.8069	17	3.2350
8	4.9267	18	-11.5889
9	9.7454	19	7.4804
10	-4.2009	20	-6.0884

Tabla 3.5: Mediciones de la posición y velocidad del sistema para el Ejemplo 3.2

Los valores iniciales que se establecen para el filtro en este ejemplo son los siguientes:

$$x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$P_0 = \begin{bmatrix} \sigma_p^2 & 0 \\ 0 & \sigma_m^2 \end{bmatrix}$$

Aplicando el algoritmo, se obtienen los valores mostrados en la Tabla 3.6

En la Figura 3.2 y 3.3, puede observarse la comparación entre los valores reales, las medidas tomadas y las aproximaciones por el filtro de Kalman de la posición y la velocidad respectivamente.

Aproximaciones por Kalman					
Iteración	Posición	Velocidad	Iteración	Posición	Velocidad
1	-0.0700	0.1754	11	1.6030	1.7441
2	0.0357	0.2280	12	3.4470	3.4579
3	0.7140	1.1502	13	2.4254	2.1646
4	0.0819	0.1195	14	2.2414	1.5822
5	0.6604	0.7090	15	1.8923	1.2341
6	0.2698	0.3710	16	2.7619	2.0206
7	1.0317	1.2473	17	3.0084	2.0269
8	1.9568	2.3143	18	0.7966	0.0644
9	3.7819	4.0720	19	1.8675	1.1663
10	2.4534	2.5651	20	0.7488	0.1294

Tabla 3.6: Aproximaciones de la posición y velocidad del sistema por filtro de Kalman para el Ejemplo 3.2

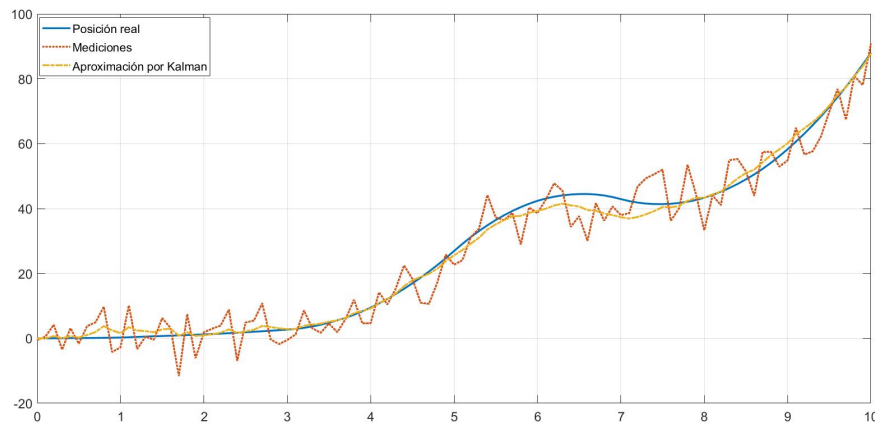


Figura 3.2: Comparación de los valores de posición reales, medidos y aproximados del Ejemplo 3.2

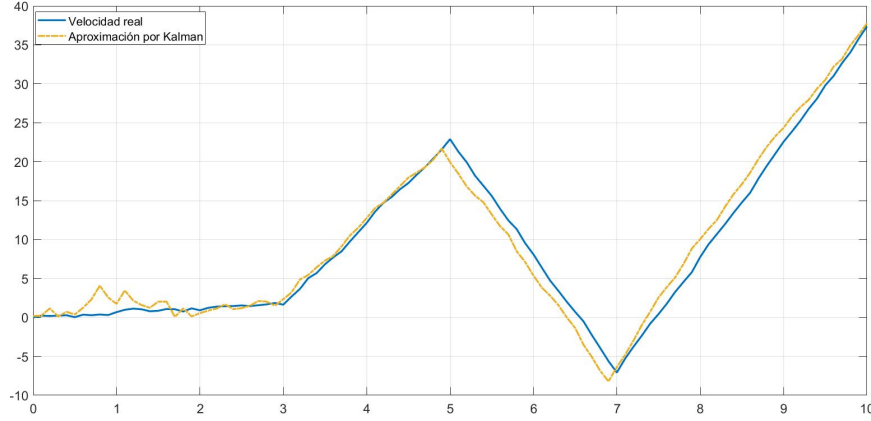


Figura 3.3: Comparación de los valores de velocidad reales y aproximados del Ejemplo 3.2

3.1. Filtro de Kalman en tiempo continuo

Comunmente llamado *Filtro Kalman-Bucy*, por las aportaciones del matemático Richard Snowden Bucy, este algoritmo es una versión en tiempo continuo del filtro de Kalman visto en la sección anterior. En este apartado se expondrán las ecuaciones que lo componen, para información más detallada puede consultar el libro de Oksendal [43].

Sea un sistema representado por la siguiente ecuación diferencial estocástica

$$dX_t = A(t)X_t dt + B(t)dU_t \quad (3.11)$$

Para el cual, las observaciones tienen la forma

$$dZ_t = G(t)X_t dt + D(t)dV_t \quad (3.12)$$

donde

$A(t) \in \mathbb{R}^{n \times n}$, $B(t) \in \mathbb{R}^{n \times p}$, $G(t) \in \mathbb{R}^{m \times n}$, $D(t) \in \mathbb{R}^{m \times r}$ son matrices deterministas. Se asume X_0 con una distribución normal y $Z_0 = 0$.

U es un movimiento Browniano p -dimensional independiente de X_0 .

V es un movimiento Browniano r -dimensional independiente de U y X_0 .

Al igual que en el filtro de tiempo discreto, el problema de filtración al que se busca dar solución es en encontrar el *mejor estimado* \hat{X}_t del estado X_t basado en las observaciones dadas por Z_s donde $0 \leq s \leq t$ [43]. Es decir,

$$\int_{\Omega} \left| X_t - \hat{X}_t \right|^2 dP = \inf \{ E[X_t - Y]^2; Y \in \mathcal{B} \}$$

donde

$$\mathcal{B} := \{ Y : \Omega \rightarrow \mathbb{R}^n; Y \in L^2(\Omega, P) \text{ y } Y \text{ es } \mathcal{G}_t\text{-medible} \}$$

\mathcal{G}_t es la σ -álgebra generada por $\{Z_s(\cdot), s \leq t\}$, es decir, la σ -álgebra más pequeña para la cual Z_t es adaptado para $s \leq t$.

Teorema 3.2 (Filtro de Kalman continuo multidimensional) *Sea $\hat{X}_t = E[X_t | \mathcal{G}_t]$, donde \hat{X}_t es \mathcal{G}_t -medible. La solución para el sistema formado por (3.12) y (3.11), satisface la ecuación estocástica diferencial*

$$d\hat{X}_t = (A_t - S_t G_t^T (D_t D_t^T)^{-1} G_t) \hat{X}_t dt + S_t G_t^T (D_t D_t^T)^{-1} dZ_t \quad (3.13)$$

donde

$\hat{X}_0 = E[X_0]$, $S(t) := E[(X_t - \hat{X}_t)(X_t - \hat{X}_t)^T] \in \mathbb{R}^{n \times n}$ satisface la ecuación de Riccati:

$$\frac{dS}{dt} = AS + SA^T - SG^T(DD^T)^{-1}GS + BB^T \quad (3.14)$$

$$S(0) = E[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T]$$

con

$$\hat{X}_t = \phi(t) \left(\hat{X}_0 + \int_0^t \phi(s)^{-1} S_s G_s^T (D_s D_s^T)^{-1} dZ_s \right) \quad (3.15)$$

donde ϕ es la solución fundamental asociada a (3.13).

Para más detalle sobre soluciones fundamentales y ecuaciones diferenciales estocásticas se aconseja consultar [6] y [36].

Cuando se trata del caso de una sola dimensión, entonces el Teorema 3.2 queda de la siguiente manera (ver [43]).

Teorema 3.3 *Dado el sistema lineal (3.16) y las observaciones lineales (3.17)*

$$dX_t = F(t)X_t dt + C(t)dU_t \quad (3.16)$$

$$dZ_t = G(t)X_t dt + D(t)dV_t \quad (3.17)$$

donde $F(t), D(t), G(t), C(t) \in \mathbb{R}$

En el caso donde el problema de filtrado es lineal unidimensional, la solución $\hat{X}_t = E[X_t | \mathcal{G}_t]$, satisface la ecuación estocástica diferencial dada por:

$$d\hat{X}_t = \left(F(t) - \frac{G^2(t)S(t)}{D^2(t)} \right) \hat{X}_t dt + \frac{G(t)S(t)}{D^2(t)} dZ_t \quad (3.18)$$

donde

$$\begin{aligned} \hat{X}_0 &= E[X_0] \\ S(t) &= E[(X_t - \hat{X}_t)^2] \text{ satisface la ecuación de Riccati} \end{aligned}$$

$$\frac{dS}{dt} = 2F(t)S(t) - \frac{G^2(t)}{D^2(t)} S^2(t) + C^2(t), \quad S(0) = E[(X_0 - E[X_0])^2] \quad (3.19)$$

Para complementar la información dada, se expondrá un ejemplo unidimensional, éste ha sido tomado de [43].

Ejemplo 3.3 *Suponga un proceso constante (3.20) cuyas observaciones están perturbadas por ruido (3.21),*

$$dX_t = 0, \quad \text{i.e., } X_t = X_0 \quad (3.20)$$

$$E[X_0] = 0, \quad E[X_0^2] = a^2$$

$$dZ_t = X_t dt + m dV_t \quad (3.21)$$

$$Z_0 = 0$$

Se desea estimar los valores de X .

De acuerdo al Teorema 3.3 los datos que se tienen del Ejemplo 3.3 pueden resumirse como:

$$F(t) = 0; \quad C(t) = 0; \quad G(t) = 1; \quad D(t) = m$$

Con lo que, la ecuación de Ricatti, quedaría como:

$$S(t) = E[(X_t - \hat{X}_t)^2] :$$

$$\frac{dS}{dt} = -\frac{1}{m^2}S^2, \quad S(0) = a^2 \quad (3.22)$$

Resolviendo (3.22) se obtiene,

$$S(t) = \frac{a^2 m^2}{a^2 t + m^2} \quad (3.23)$$

Por (3.18) se tiene que,

$$d\hat{X}_t = -\frac{a^2}{a^2 t + m^2} \hat{X}_t dt + \frac{a^2}{a^2 t + m^2} dZ_t; \quad \hat{X} = E[X_0] \quad (3.24)$$

Para resolver (3.24), considere la siguiente ecuación

$$\phi(t, \hat{X}_t) = \hat{X}_t \exp\left(\int_0^t \frac{a^2}{m^2 + a^2 s} ds\right)$$

La cual tiene las siguientes derivadas parciales,

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \hat{X}_t \exp\left(\int_0^t \frac{a^2}{m^2 + a^2 s} ds\right) \left(\frac{a^2}{m^2 + a^2 t}\right) \\ \frac{\partial \phi}{\partial \hat{x}} &= \exp\left(\int_0^t \frac{a^2}{m^2 + a^2 s} ds\right) \\ \frac{\partial^2 \phi}{\partial \hat{x}^2} &= 0 \end{aligned}$$

Aplicando la fórmula de Itô (ver ecuación (1.22)) a ϕ se tiene que

$$\begin{aligned} d\phi(t, \hat{X}_t) &= \frac{\partial \phi}{\partial t} dt + \frac{\partial \phi}{\partial \hat{x}}(t, \hat{X}_t) d\hat{X}_t + \frac{1}{2} \frac{\partial^2 \phi}{\partial \hat{x}^2}(t, \hat{X}_t) (d\hat{X}_t)^2 \\ &= \hat{X}_t \exp\left(\int_0^t \frac{a^2}{m^2 + a^2 s} ds\right) \left(\frac{a^2}{m^2 + a^2 t}\right) dt \\ &\quad + \exp\left(\int_0^t \frac{a^2}{m^2 + a^2 s} ds\right) d\hat{X}_t \end{aligned}$$

Para simplificar la escritura de los cálculos, llamemos

$$\beta(t) = \exp\left(\int_0^t \frac{a^2}{m^2 + a^2s} ds\right)$$

Sustituyendo el valor de $d\hat{X}_t$ en la ecuación anterior, se obtiene

$$\begin{aligned} d\phi(t, \hat{X}_t) &= \hat{X}_t\beta(t) \left(\frac{a^2}{m^2 + a^2t}\right) dt + \beta(t) \left(\frac{-a^2}{m^2 + a^2t} \hat{X}_t dt + \frac{a^2}{m^2 + a^2t} dZ_t\right) \\ &= \hat{X}_t\beta(t) \left(\frac{a^2}{m^2 + a^2t}\right) dt - \beta(t) \left(\frac{a^2}{m^2 + a^2t}\right) \hat{X}_t dt + \beta(t) \left(\frac{a^2}{m^2 + a^2t}\right) dZ_t \\ &= \exp\left(\int_0^t \frac{a^2}{m^2 + a^2s} ds\right) \frac{a^2}{m^2 + a^2t} dZ_t \end{aligned}$$

Entonces,

$$d\left(\hat{X}_t \exp\left(\int_0^t \frac{a^2}{m^2 + a^2s} ds\right)\right) = \exp\left(\int_0^t \frac{a^2}{m^2 + a^2s} ds\right) \frac{a^2}{m^2 + a^2t} dZ_t \quad (3.25)$$

Además,

$$\exp\left(\int_0^t \frac{a^2}{m^2 + a^2s} ds\right) = \exp\left(\ln\left|\frac{m^2 + a^2t}{m^2}\right|\right) = \frac{m^2 + a^2t}{m^2} \quad (3.26)$$

Sustituyendo (3.26) en (3.25)

$$d\left(\hat{X}_t \frac{m^2 + a^2t}{m^2}\right) = \left(\frac{m^2 + a^2t}{m^2}\right) \frac{a^2}{m^2 + a^2t} dZ_t = \frac{a^2}{m^2} dZ_t \quad (3.27)$$

De (3.27) se obtiene que \hat{X} está dada por

$$\hat{X}_t = \frac{a^2}{a^2t + m^2} Z_t$$

Capítulo 4

El problema de Monty Hall

El *problema de Monty Hall* se trata de un problema de probabilidad condicional que se encuentra basado originalmente en un programa televisivo estadounidense, cuya versión mexicana es presentada como *la Catafixia*.

La solución a dicho problema fue realizada formalmente por Marilyn von Savant, quien trabajaba como escritora para una columna de preguntas en una revista. Su respuesta era correcta, sin embargo, recibió cientos de comentarios y burlas provenientes de matemáticos reconocidos argumentando que estaba equivocada [22]. Ella mantuvo su postura, pidió a su público que jugara cuantas veces fuera necesario para darse cuenta de ello y al final, ha sido su respuesta la que prevaleció.

En esta sección, analizaremos la solución a este juego de elección, el cual se presenta en el Ejemplo 4.1.

Ejemplo 4.1 *Considere que se encuentra en un juego de elección, usted está frente a tres puertas marcadas como A , B y C . Una de las puertas oculta un premio mayor, supóngase un auto, mientras que en cada una de las restantes se encuentra una cabra. Usted se llevará lo que haya detrás de la puerta seleccionada, considere que el único que conoce como se han acomodado los premios detrás de cada puerta es el presentador del juego y ahora le piden a usted que escoja entre las puertas. Suponga que ha elegido A , con lo que el presentador abre una de las restantes, específicamente la que contenga una cabra, ya sea B o C y después de esta revelación, se le da la opción de mantener su elección original A , o bien, cambiarla por la otra puerta restante ¿Qué es más conveniente, cambiar su elección o quedarse con la puerta original?*

Marilyn von Savant afirmaba que para aumentar las probabilidades de ganar el premio mayor, el jugador debería cambiar la elección original por la puerta que queda sin elegir pero ¿Por qué?

Al inicio de la sección se mencionó que esto puede verse como un problema de probabilidad condicional (ver la igualdad (1.1)), lo cual se explicará con más detalle a continuación.

La probabilidad condicional de un evento se interpreta como la probabilidad de que el evento suceda dado que uno o más eventos relacionados a éste han ocurrido. Es decir, $P(B|A)$ es la probabilidad de que B se lleve a cabo una vez que A ha sucedido.

Por ejemplo, suponga los eventos

$$A = \{\text{Realizar actividad física}\} \quad B = \{\text{tener dolor de cuerpo}\}$$

entonces $P(B|A)$ sería la probabilidad de tener dolor de cuerpo dado que se realizó alguna actividad física o, en otro caso, $P(A|B)$ que se interpretaría como la probabilidad de haber realizado alguna actividad física dado que se tiene dolor de cuerpo.

Regresando al Ejemplo 4.1. Considere dos tipos de jugadores: los que siempre cambian de puerta J_1 , y los que siempre se quedan con su elección original J_2 .

Suponga los siguientes tres eventos:

$$S_a := \{\text{El jugador inicialmente selecciona la puerta que esconde el auto}\}$$

$$S_c := \{\text{El jugador inicialmente selecciona una puerta que esconde una cabra}\}$$

$$G := \{\text{El jugador gana el auto}\}$$

Se desea calcular la probabilidad de ganar $P(G)$ para cada tipo de jugador. Se tiene que, para calcular $P(G)$, G puede escribirse como $G = (G \cap S_a) \cup (G \cap S_c)$, es decir la unión de los dos eventos donde es posible ganar. Con lo que,

$$\begin{aligned} P(G) &= P((G \cap S_a) \cup (G \cap S_c)) \\ &= P(G \cap S_a) + P(G \cap S_c) \end{aligned}$$

De la igualdad (1.1) se sabe que,

$$P(A \cap B) = P(A|B)P(B)$$

Entonces,

$$P(G) = P(G|S_a)P(S_a) + P(G|S_c)P(S_c)$$

Dado que se tienen tres posibles elecciones, de las cuales una de ellas es un auto y las dos restantes son cabras, entonces se puede suponer que la $P(S_a) = \frac{1}{3}$, mientras que la $P(S_c) = \frac{2}{3}$. Considerado todo lo mencionado anteriormente, se analizarán los dos tipos de jugadores.

Para J_1 , el jugador que siempre cambia su elección original, se tiene que,

$$P(G|S_a) = 0, \quad \& \quad P(G|S_c) = 1$$

Con lo que,

$$P(G)_{J_1} = 0 + 1 \left(\frac{2}{3} \right) = \frac{2}{3} \quad (4.1)$$

Por el otro lado, para J_2 , el jugador que siempre mantiene su elección original se tiene,

$$P(G|S_a) = 1, \quad \& \quad P(G|S_c) = 0$$

Entonces,

$$P(G)_{J_2} = 1 \left(\frac{1}{3} \right) + 0 = \frac{1}{3} \quad (4.2)$$

De las ecuaciones (4.1) y (4.2) se puede concluir que el jugador que cambia su elección original tiene mayores probabilidades de ganar, lo cual afirmaba Marilyn von Savant.

Capítulo 5

Juegos de Azar

Para el desarrollo de esta sección, se retomarán los conceptos sobre martingala y filtración dados en el Capítulo 1.

Suponga el escenario descrito a continuación: Una gimnasta se encuentra parada en medio de una viga de equilibrio, cada minuto lanza una moneda, si el resultado es águila dará un paso hacia adelante, de lo contrario retrocederá un paso, si suponemos que no se cuenta con una moneda justa, entonces la gimnasta avanzará con probabilidad p y retrocederá con probabilidad $1 - p$. A este fenómeno se le conoce como una caminata aleatoria simple, la cual es un tipo de proceso estocástico.

Definición 5.1 (Caminata aleatoria) *Se conoce como caminata aleatoria a una sucesión de variables aleatorias $\{S_n\}$, para la cual $S_0 = 0$, definida como:*

$$S_n = \sum_{k=1}^n A_k \quad (5.1)$$

donde las variables aleatorias $\{A_k\}$ son independientes e idénticamente distribuidas, i.e., que todas las v.a. tienen la misma función de distribución (ver igualdad (1.3)).

Las caminatas aleatorias se utilizan para estudiar la evolución de sucesos que pueden describirse como una serie de pasos aleatorios. En algunas ocasiones, el estudio del instante en que un fenómeno particular ocurre por primera vez dentro de un proceso estocástico, es de gran importancia y se conoce como *tiempo de paro*.

Si se utiliza como contexto el escenario de la gimnasta, se podría estudiar el primer momento en el que ella se cae de la viga y la probabilidad que tenga de hacerlo.

Los tiempos de paro (ver Definición 1.19) comunmente son utilizados para el estudio de capital de empresas, tal es el caso de las compañías de seguros, donde el fenómeno particular es el momento en el que ésta llega a la quiebra. A continuación se expondrá un ejemplo basado en juegos de azar.

Ejemplo 5.1 *Suponga que un jugador desea apostar en un casino, para lo cual dispone de j pesos, mientras que el casino está dispuesto a perder hasta c pesos. Éste se trata de un juego justo, donde en cada ronda se puede ganar o perder un peso. La pregunta principal se basa en saber si se jugará un tiempo finito antes de que ocurra la ruina del jugador o del casino.*

En este ejemplo $S_n = \sum_{i=1}^n \xi_i$ donde la familia $\{\xi_i : i \in \mathbb{N}\}$ son variables aleatorias independientes e indenticamente distribuidas. Para resolver el Ejemplo 5.1 consideramos el tiempo de paro:

$$T = \inf\{n : S_n = -j \text{ o } c\}, \quad \inf\emptyset = +\infty \quad (5.2)$$

Note que T indica el primer momento en el que el jugador pierde j pesos o en dado caso, cuando él le ha ganado al casino los c pesos.

Se definen los siguientes conjuntos

$$A_k = \{\xi_{k+1} = 1, \dots, \xi_{k+d} = 1\}, \text{ para todo } k \geq 0 \quad (5.3)$$

donde $d = j + c$ es el dinero total de la apuesta y A_k representa los juegos al tiempo k para los cuales el jugador ha ganado d -partidas continuas. De forma que los conjuntos $\{A_{kd}\}_{k=0}^{\infty}$ son independientes,

$$\begin{aligned} k = 0 & \quad A_0 = \{\xi_1 = 1, \dots, \xi_d = 1\} \\ k = 1 & \quad A_d = \{\xi_{d+1} = 1, \dots, \xi_{2d} = 1\} \\ k = 2 & \quad A_{2d} = \{\xi_{2d+1} = 1, \dots, \xi_{3d} = 1\} \\ & \quad \vdots \end{aligned} \quad (5.4)$$

En cada ronda del juego el jugador tiene la mitad de probabilidad de ganar, esto es,

$$P(\xi_k = 1) = \frac{1}{2} \quad P(\xi_k = -1) = \frac{1}{2} \quad (5.5)$$

Retomando los conjuntos A_{kd} de las rachas ganadoras del jugador, note que la probabilidad de estos conjuntos, por ser independientes es,

$$P(A_{kd}) = \frac{1}{2} \times \frac{1}{2} \times \dots \times \frac{1}{2} = \frac{1}{2^d} \quad (5.6)$$

De la ecuación (5.6) se tiene que,

$$\sum_{k=1}^{\infty} P(A_{kd}) = \sum_{k=1}^{\infty} \frac{1}{2^d} = +\infty \quad (5.7)$$

Entonces, por el Lema de Borel-Cantelli, ver Capítulo 1, Lema 1.1, la ecuación (5.7) implica que,

$$P(\limsup_{k \rightarrow \infty} A_{kd}) = 1 \quad (5.8)$$

Lo anterior indica que los eventos $\{A_{kd}\}_{k=0}^{\infty}$, además ocurren una infinidad de veces. Con esto, se sabe que existe con probabilidad uno un número aleatorio τ tal que A_{τ} ocurre.

Con lo que se tienen dos posibles casos, $S_{\tau} \in (-j, c)$ o $S_{\tau} \notin (-j, c)$. Note que τ se interpreta como la primera vez que el jugador tiene una racha ganadora. Por lo que el primer caso es el de interés.

Entonces, para $S_{\tau} \in (-j, c)$,

$$\begin{aligned} S_{\tau+d} &= \xi_1 + \dots + \xi_{\tau} + \dots + \xi_{\tau+d} \\ &= S_{\tau} + \xi_{\tau+1} + \dots + \xi_{\tau+d} \\ &= S_{\tau} + d \\ &> -j + d \\ &= c \end{aligned} \quad (5.9)$$

Note que la siguiente propiedad de continuidad se cumple para S_n ,

Si $S_n > s \in \mathbb{N}$, entonces, existe $i < n$ tal que $S_i = s$

Con lo que, para $S_{\tau+d} > c$, se obtiene que $T < \tau + d$

Con lo anterior se ha demostrado que con probabilidad uno $T < \infty$, es decir, el juego tendrá un número finito de rondas.

Un evento que resulta de gran interés cuando se trata de juegos de azar, es calcular la probabilidad que tiene el jugador de perder su fortuna, es decir,

la probabilidad de ruina. Esto puede calcularse por medio de probabilidad condicional (ver igualdad (1.1)). Para responder a esta cuestión suponga los siguientes eventos,

$$R = \{\text{El jugador pierde}\} \quad \& \quad R^c = \{\text{El jugador gana}\}$$

Se define H_j como la probabilidad que tiene el jugador de perder habiendo iniciado el juego con un capital de j pesos. Esto es,

$$H_j := P(R|S_0 = j)$$

Entonces, después de la primer partida,

$$H_{j+1} = P(\text{Perder}|S_0 = j \wedge \text{ganar la 1ra ronda})$$

$$H_{j-1} = P(\text{Perder}|S_0 = j \wedge \text{perder la 1ra ronda})$$

Para simplificar la escritura, se tomará $P(h_j) = P(R|S_0 = j)$

Por el Teorema 1.1, se tiene que

$$\begin{aligned} P(h_j) &= P(\text{Perder}|S_0 = j \wedge \text{perder la 1ra ronda})P(\xi_1 = -1) \\ &\quad + P(\text{Perder}|S_0 = j \wedge \text{ganar la 1ra ronda})P(\xi_1 = 1) \\ &= P(R|S_1 = j - 1)P(\xi_1 = -1) + P(R|S_1 = 1 + j)P(\xi_1 = 1) \\ &= P(R|S_1 = j - 1)\frac{1}{2} + P(R|S_1 = 1 + j)\frac{1}{2} \end{aligned}$$

Lo cual puede escribirse como:

$$H_j = \frac{1}{2}H_{j+1} + \frac{1}{2}H_{j-1} \tag{5.10}$$

Además, se sabe que, si se inicia el juego con un capital igual a cero, entonces la ruina es segura, i.e, $H_0 = 1$. Por el otro lado, si se inicia con un capital $d = c + j$, es decir, habiendo ganado al casino, entonces la ruina es practicamente imposible, i.e, $H_d=0$.

El procedimiento detallado para solucionar una ecuación en diferencias, como lo es (5.10), puede consultarse en [37] o de forma más general en el Apéndice A.

Ahora, dichos pasos para resolver (5.10), implican que se obtiene una ecuación cracterística, la cual en este caso es $\frac{1}{2}y^2 - y + \frac{1}{2} = 0$, o bien, para fines prácticos $y^2 - 2y + 1 = 0$

Llame r a las raíces de la ecuación característica, las cuales en este caso son $r_1 = r_2 = 1$. Lo que implica que, se propone $H_j = j$.

Se propone además una solución de la forma

$$H_j = C_1 r^j + C_2 j r^j \quad (5.11)$$

El cual quedaría, para este caso $r = 1$ como

$$H_j = C_1 + C_2 j \quad (5.12)$$

Lo siguiente es calcular las constantes a partir de las condiciones iniciales dadas, $H_0 = 1$ y $H_d = 0$, con ello las constantes quedarían como

$$H_0 = C_1 = 1$$

$$H_d = 1 + C_2(c + j) \Rightarrow C_2 = \frac{-1}{c + j}$$

Por lo tanto, sustituyendo C_1 y C_2 en (5.11), la probabilidad que tiene el jugador de perder su fortuna partiendo de un capital j es,

$$H_j = \frac{c}{c + j} \quad (5.13)$$

Hasta este punto del problema se ha dado respuesta a dos cuestiones, conocer la probabilidad de ruina del jugador y saber si el juego será finito, pero ¿Cúántas partidas podrían esperarse jugar antes de que el juego termine? Para dar respuesta a esto, suponga lo siguiente.

Sea N el número restante de partidas y $E_j := E(N|S_0 = j)$ el tiempo esperado para ganar o perder del jugador iniciando con un capital de j pesos.

$$\begin{aligned} E_j &= E(N|S_0 = j) \\ &= E(N|S_0 = j \wedge \text{ganar la 1ra ronda})P(\text{ganar la 1ra ronda}) \\ &\quad + E(N|S_0 = j \wedge \text{perder la 1ra ronda})P(\text{perder la 1ra ronda}) \\ &= E(N|S_1 = j + 1)P(\text{ganar la 1ra ronda}) \\ &\quad + E(N|S_n = n - 1)P(\text{perder la 1ra ronda}) \end{aligned}$$

Recordando que se trata de un juego justo se tiene que,
 $P(\text{ganar la 1ra ronda}) = P(\text{perder la 1ra ronda}) = \frac{1}{2}$

Si el primer juego es exitoso, es decir el jugador gana, entonces el juego continua ahora comenzando desde $S_0 + 1$, por lo que la duración del juego con éxito en la primera ronda se podría expresar como $E_{j+1} + 1$. De la misma forma cuando la primera ronda resulta en pérdida la duración del juego se representaría como $E_{j-1} + 1$. Por lo que la igualdad de E_j quedaría como,

$$\begin{aligned} E_j &= \frac{1}{2}(E(N|S_0 = j + 1) + 1) + \frac{1}{2}(E(N|S_0 = j - 1) + 1) \\ &= 1 + \frac{1}{2}E_{j+1} + \frac{1}{2}E_{j-1} \end{aligned} \quad (5.14)$$

Note que la ecuación (5.14) es una ecuación en diferencias no homogénea de segundo grado que se resolverá de una forma similar al procedimiento realizado anteriormente cuando se calculó la ruina del jugador. Además, se sabe que el juego termina cuando el capital inicial llega a cero $j = 0$ o cuando se le ha ganado al casino, $j = d$. y por lo tanto la duración esperada del juego en ambos casos es 0, i.e, $E_{j=0}=0$ y $E_{j=d} = 0$

Para más información sobre el procedimiento para resolver ecuaciones en diferencias no homogéneas consulte el Apéndice A, para detalles específicos y demostraciones puede ver [26] y [37].

Se tiene que (5.14) es lo mismo que

$$E_{j+1} - 2E_j + E_{j-1} = -2 \quad (5.15)$$

Ahora, para este tipo de ecuaciones se propone como solución general la siguiente ecuación,

$$E_j = C_1 + jC_2 \quad (5.16)$$

Además, se propone una solución particular de la forma $E_j = Cj^2$, donde C es una constante que será calculada sustituyendo dicha solución particular en (5.15), con ello se tiene

$$C(j + 1)^2 - 2Cj^2 + C(j - 1)^2 + 2 = 0 \Rightarrow C = -1 \quad (5.17)$$

Entonces, la solución particular $E_j = Cj^2$ quedaría como $E_j = -j^2$.

Para obtener la solución completa de (5.14), se suman la solución particular con la solución general, lo que da como resultado la ecuación siguiente,

$$C_1 + C_2j - j^2 \quad (5.18)$$

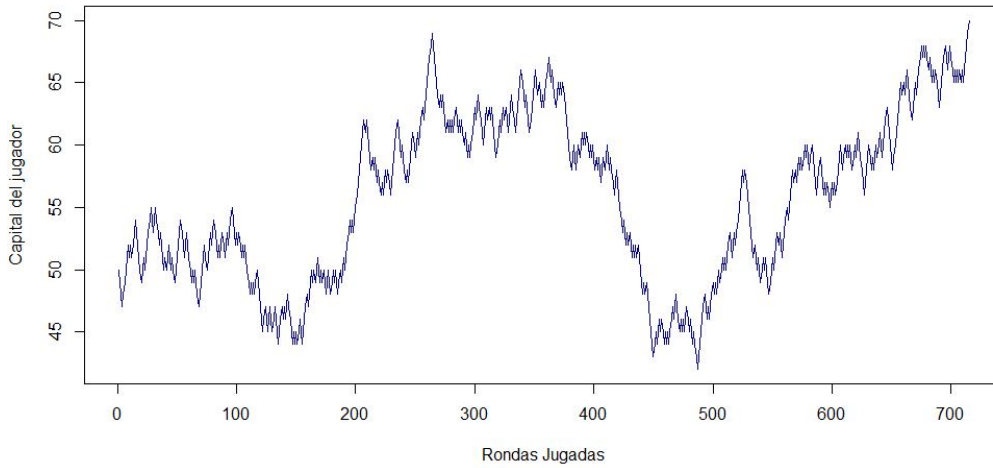


Figura 5.1: Caminata del Ejemplo 5.1, donde $c = 20$, $j = 50$ y el jugador gana

Se necesitan calcular los valores de las constantes involucradas en (5.18). Por las condiciones dadas, $E_{j=0} = E_{j=d} = 0$, se obtiene que $C_1 = 0$ y $C_2 = d$. Sustituyendo C_1 y C_2 en (5.18),

$$\begin{aligned}
 E_j &= (c + j)j - j^2 \\
 &= cj + j^2 - j^2 \\
 &= cj
 \end{aligned}
 \tag{5.19}$$

Por lo tanto, el tiempo esperado de la duración del juego es $E_j = cj$

En las Figuras 5.1 y 5.2 se pueden observar una caminata aleatoria obtenida de la simulación de una serie de juegos donde el jugador gana y pierde respectivamente con $c = 20$ y $j = 50$.

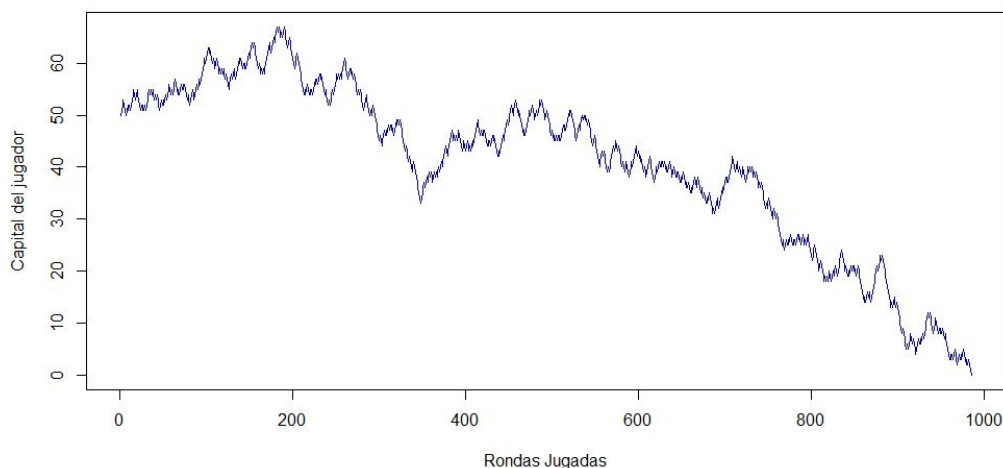


Figura 5.2: Caminata del Ejemplo 5.1, donde $c = 20$, $j = 50$ y el jugador pierde

5.1. Teoría de Riesgo

En la sección anterior se estudió la probabilidad de que un jugador llegue a la ruina realizando una serie de apuestas, es importante notar que la teoría de riesgo va más allá de saber si se ganará o perderá un juego.

Para toda empresa es necesario conocer las consecuencias de las decisiones que se realizan sobre el capital que ésta posee, en especial, saber sobre su probabilidad de ruina, de forma que puedan contar con estrategias para un manejo correcto de las reservas.

Un ejemplo claro son las compañías de seguros, para las cuales los ingresos totales que se generan están ligados a fenómenos aleatorios y por lo tanto su capital presenta fluctuaciones. En el Ejemplo 5.2 se estudiará una aproximación del cálculo de la probabilidad de ruina para estas empresas.

Ejemplo 5.2 *Suponga que desea abrir una compañía de seguros, la cual funciona de la siguiente manera: un conjunto de individuos, los cuales se encuentran expuestos a un tipo de riesgo específico le pagan a usted una cantidad fija cada determinado tiempo, a cambio, su empresa se compromete a pagar un monto predeterminado en el momento en que este incidente ocurra. Lo que a usted le interesa saber es cual es la probabilidad de que su compañía quiebre.*

Es importante notar que en este tipo de casos no se puede saber con certeza cuantas personas contratarán el servicio, tampoco la cantidad de compensaciones que se deberán pagar en el año, aun así la empresa debe ser capaz de cubrir los gastos.

Sea $X_0 > 0$ el capital inicial de la compañía, X_n las ganancias en el año n , $p \geq 0$ las primas constantes pagadas por los asegurados en el año y C_n las compensaciones cobradas por los usuarios, donde C_1, C_2, \dots , son variables aleatorias independientes idénticamente distribuidas con función de distribución $N(\mu, \sigma^2)$ (ver Definición 1.6). Suponga que $\mu < p$, lo cual significa que las indemnizaciones son menores que los ingresos en el año.

El modelo que describe la manera en la que opera la aseguradora está dado por:

$$X_{n+1} = X_n + p - C_{n+1} \quad (5.20)$$

La ecuación (5.20) es conocida como *modelo clásico de riesgo* o *modelo de Cramer-Lundberg*.

Una compañía aseguradora quedará en la ruina si para cuando el año haya terminado las ganancias son menores o iguales a cero. Con lo que el momento de ruina es,

$$\tau = \inf\{n : X_n \leq 0\}$$

Se tiene que la probabilidad de ruina está definida por

$$P(X_n \leq 0, \text{ para algún } n)$$

La función generadora de momento de una variable aleatoria con distribución normal con media μ y varianza σ^2 está dada por

$$E(e^{tX}) = e^{\mu t + \frac{\sigma^2 t^2}{2}}, \quad t \in \mathbb{R}$$

Entonces, tomando $p - C_1$, la función generadora es

$$\begin{aligned} M_t &= E(e^{t(p-C_1)}) \\ &= E(e^{tp} e^{-tC_1}) \\ &= e^{tp} E(e^{-tC_1}) \\ &= e^{(p-\mu)t + \frac{\sigma^2 t^2}{2}} \end{aligned} \quad (5.21)$$

Para simplificación de los cálculos, se pide que $M(t_0) = 1$, lo que implicaría que

$$(p - \mu)t_0 + \frac{\sigma^2 t_0^2}{2} = 0$$

con lo que el valor de t_0 , para cumplir que $M(t_0) = 1$, quedaría como

$$\begin{aligned} t_0 \left(p - \mu + \frac{\sigma^2 t_0}{2} \right) &= 0 \\ \frac{\sigma^2 t_0}{2} &= -p + \mu \\ t_0 &= -\frac{2(p - \mu)}{\sigma^2} < 0 \end{aligned}$$

donde t_0 se refiere al primer instante de tiempo a ser evaluado. Ahora, se definen

$$Z_n = \min\{e^{t_0 X_n}, 1\}, \quad \mathcal{F}_n = \sigma\{C_1, C_2, \dots, C_n\}, \forall n \in \mathbb{N}.$$

Z_n está acotada por 1 y es integrable, i.e., $E(|Z_n|) < \infty$. Además, se tiene que,

$$Z_{n+1} = \min\{e^{t_0 X_{n+1}}, 1\}$$

Por lo tanto $Z_{n+1} \leq e^{t_0 X_{n+1}}$.

De acuerdo al Teorema 1.2 y recordando el hecho de que C_{n+1} es independiente de C_1, \dots, C_n . Se tiene que,

$$\begin{aligned} E(Z_{n+1} | \mathcal{F}_n) &\leq E(e^{t_0 X_{n+1}} | \mathcal{F}_n) \\ &= E(e^{t_0(X_n + p - C_{n+1})} | \mathcal{F}_n) \\ &= E(e^{t_0 X_n} e^{t_0(p - C_{n+1})} | \mathcal{F}_n) \end{aligned}$$

Note que $e^{t_0 X_n}$ es \mathcal{F}_n -medible y por la hipótesis i.i.d, i.e., independientes e idénticamente distribuidas, sobre la familia $\{C_i\}$, es decir, que son independientes e idénticamente distribuidas, entonces

$$\begin{aligned} E(e^{t_0 X_n} e^{t_0(p - C_{n+1})} | \mathcal{F}_n) &= e^{t_0 X_n} E(e^{t_0(p - C_{n+1})} | \mathcal{F}_n) \\ &= e^{t_0 X_n} E(e^{t_0(p - C_{n+1})}) \\ &= e^{t_0 X_n} M_{t_0} \\ &= e^{t_0 X_n} \end{aligned}$$

Ahora, como $E(Z_{n+1}|\mathcal{F}_n)$ está acotada por $e^{t_0 X_n}$ y además también está acotada por 1, entonces,

$$\begin{aligned} E(Z_{n+1}|\mathcal{F}_n) &\leq \min\{1, e^{t_0 X_n}\} \\ &= Z_n \end{aligned}$$

Observe que $\{Z_n, \mathcal{F}_n\}$ es una supermartingala.

Se define el tiempo de paro τ relacionado al primer instante en el que la compañía queda en ruina como

$$\tau = \inf\{n : X_n \leq 0\} = \inf\{n : Z_n = 1\} \quad (5.22)$$

Para los siguientes pasos, es necesario recordar el *Teorema de paro Opcional* (ver Teorema 1.4).

Tomando $\tau \wedge m$, $m \in \mathbb{N}$, como el tiempo de paro, el cual es acotado y dado que $\{Z_n, \mathcal{F}_n\}$ es una supermartingala, entonces por el Teorema del paro Opcional se tiene que,

$$E(Z_{\tau \wedge m}|\mathcal{F}_0) \leq Z_0 \quad c.s \quad (5.23)$$

Calculando la esperanza para ambos términos se obtiene

$$E(E(Z_{\tau \wedge m}|\mathcal{F}_0)) \leq E(Z_0)$$

Por propiedades de la esperanza condicional (ver Teorema 1.2) se tiene que

$$E(E(Z_{\tau \wedge m}|\mathcal{F}_0)) = E(Z_{\tau \wedge m}) \quad (5.24)$$

De (5.23) y (5.24) se implica que

$$E(Z_0) \geq E(Z_{\tau \wedge m}) \quad (5.25)$$

Note que $Z_{\tau \wedge m}$ es positiva, por definición de la esperanza, se tiene que,

$$\begin{aligned} E(Z_{\tau \wedge m}) &= \int_{\Omega} Z_{\tau \wedge m} dP \\ &\geq \int_{[\tau \leq m]} Z_{\tau \wedge m} dP \\ &= \int_{[\tau \leq m]} Z_{\tau} dP \end{aligned}$$

Dado que $\tau(\omega) = \inf\{n : Z_n(\omega) = 1\}$ entonces $Z_\tau(\omega) = Z_{\tau(\omega)}(\omega)$. De forma que

$$E(Z_{\tau \wedge m}) \geq \int_{[\tau \leq m]} 1dP$$

Con

$$\int_{[\tau \leq m]} 1dP = P\{\tau \leq m\}$$

Por lo que

$$E(Z_{\tau \wedge m}) \geq P\{\tau \leq m\} \quad (5.26)$$

De las desigualdades en (5.25) y (5.26) se deduce que

$$E(Z_0) \geq P\{\tau \leq m\} \quad \forall m \in \mathbb{N} \quad (5.27)$$

Recuerde que $Z_0 = \min\{e^{t_0 X_0}, 1\}$ y se tiene que $t_0 < 0$, por lo que

$$Z_0 = e^{t_0 X_0}$$

Por consecuente $Z_0 < 1$, además,

$$\begin{aligned} E(Z_0) &= e^{t_0 X_0} \\ &= e^{\frac{-2(p-\mu)}{\sigma^2} X_0} \end{aligned}$$

Con lo que,

$$E(Z_0) \geq P\{\tau \leq m\}, \forall m \in \mathbb{N} \quad (5.28)$$

Donde $\tau \leq m$ se interpreta como que para alguna $n \leq m$, $X_n \leq 0$, lo que sería la ruina.

Si se hace tender $m \rightarrow \infty$ y se obtiene el límite, entonces de la desigualdad (5.28) se obtiene que la probabilidad de ruina de la compañía sería:

$$P\{X_n \leq 0, \text{ para algún } n\} \leq e^{\frac{-2(p-\mu)}{\sigma^2} X_0}.$$

5.2. Cadenas de Markov

A lo largo de este capítulo se ha hablado sobre temas que implican procesos estocásticos, tales son los juegos de azar y las probabilidades de ruina, un

tipo de proceso estocástico más que se expondrá son las *cadena de Markov* (ver Definición 1.23).

Entre los problemas que las cadenas de Markov han podido optimizar se encuentra una de las herramientas más importantes de nuestros días: los motores de búsqueda dentro de la *World Wide Web*, específicamente *Google*, el cual se describe en el Ejemplo 5.3 que se estudiará en los párrafos siguientes. Para poder desarrollar y entender dicho ejemplo se requiere de exponer previamente ciertos conceptos relacionados a la *Teoría de grafos*, los cuales se presentan a continuación.

Definición 5.2 (Grafo) *Se conoce como un grafo G a un par ordenado (N, V) formado por un conjunto no vacío de nodos (o vértices) N y un conjunto de aristas (o vínculos) llamado V (el cual es disjunto de N). Este par está asociado a una función de incidencia ϕ_G la cual asigna un par no ordenado $\{xy\}$ de vértices con una arista, es decir $\phi_G(e) = \{x, y\}$, con lo que se dice que e conecta a x con y [16].*

Si la arista asigna una relación de dirección entre los nodos x, y , entonces ésta se conoce como una *arista dirigida*, \vec{xy} . El nodo desde el cual parte el vínculo se conoce como *nodo de salida* y por consecuente, el nodo a donde llega se denomina *nodo de llegada* [8].

A las aristas de los grafos orientados se les puede asignar un valor el cual debe ser no negativo, si este es el caso, el grafo se dice *ponderado*. Generalmente, el valor de estas ponderaciones es menor o igual a 1, de forma que la suma de las ponderaciones de las aristas que comparten un mismo nodo de salida debe ser igual a uno [8].

Un ejemplo para entender lo anterior sería ver las ciudades de un estado como los nodos mientras que las carreteras que las conectan entre sí son las aristas. Además, si supone que usted vive en una ciudad A y trabaja en una ciudad B conectadas entre sí por tres carreteras entre las cuales una de ellas pasa por una ciudad C , usted diariamente tendría que elegir por cuál carretera regresar a su casa y esta decisión únicamente dependería de factores presentes en el momento que está en marcha, tal como es el tráfico, el tiempo, el clima, etc. Con lo que esto, además de verse como un grafo, también podría interpretarse como una cadena de Markov.

Una forma alternativa para describir un grafo es representarlo de forma matricial, esta representación se conoce como *matriz de adyacencia* [25].

Definición 5.3 (Matriz de adyacencia) Sea $G = (N, V)$ un grafo. Se conoce como **matriz de adyacencia** de G a la matriz $A = (a_{i,j})$ cuadrada de orden n , para la cual las entradas $(a_{i,j})$ representan el número de aristas que unen el nodo n_i con el nodo n_j , para $i, j = \{1, \dots, s\}$

Una vez entendidos los conceptos anteriores se puede explicar como una cadena de Markov se relaciona con un grafo. Básicamente, el grafo es una representación visual de la cadena, ésta tiene un conjunto de estados, cada uno de ellos representará un nodo para el grafo y cada paso es visto como un enlace entre estados, ponderado con la probabilidad de transición [13]. Este grafo es llamado *diagrama de transición de estados*. Además, note que si la matriz de adyacencia del grafo es estocástica entonces se trata de la matriz de probabilidades de transición.

Prosiguiente con el objetivo principal de esta sección, el estudio del algoritmo utilizado por *Google*, se presenta el Ejemplo 5.3.

Ejemplo 5.3 *En la actualidad, la cantidad de páginas web que conforman la World Wide Web es inmensa, de forma que, si no existiera una manera de ordenar los resultados de las búsquedas, la mayor parte de éstas obtendrían miles o incluso millones de resultados no útiles. Google es el buscador de archivos almacenados en servidores web más utilizado. Fue desarrollado en 1997 por Larry Page y Sergey Brin y su eficiencia se basa en el algoritmo PageRank, por lo que es natural cuestionarse ¿Cómo funciona PageRank para entregar los resultados más óptimos?*

El escenario del Ejemplo 5.3 puede plantearse como lo siguiente: una persona va explorando la web, yendo de una página a otra escogiendo de forma aleatoria los enlaces, si continua así, eventualmente llegará a una página que no cuente con un vínculo de salida o en dado caso su búsqueda podría volverse cíclica entre un grupo de páginas interconectadas, entonces el Ejemplo 5.3 se puede interpretar como un grafo orientado ponderado, donde los nodos son todas las páginas que existen en la *World Wide Web* mientras que las aristas son los enlaces que las interconectan entre sí, de forma que el usuario puede llegar a una página desde otras.

Además, este problema representado como un grafo es una cadena de Markov, donde cada nodo de la red es un estado de la cadena, pues se asume que la siguiente página que el usuario escogerá únicamente depende de lo relevante que es la página actual para su búsqueda.

Con este contexto, la búsqueda podría representarse definiendo un proceso $\{X_n; n \in \mathbb{N}\}$, para el cual X_n indica en qué página web está el usuario. Defina i como la página actual, j la página futura, E_j el número total de enlaces de salida que se encuentran en j y n los pasos de la búsqueda, con lo que la probabilidad de estar en la página j dado que el usuario se encontraba en la página i está definida como

$$P(X_{n+1} = j | X_n = i) = \frac{1}{E_i} \quad (5.29)$$

Se ha dicho que el grafo que representa a la web puede ser ponderado, esta ponderación básicamente funciona de la siguiente manera: la importancia de una página web puede asignarse por medio de las páginas que tienen un enlace hacia ella. Si i tiene un vínculo hacia j entonces el contenido de j es relevante para i . Si existen varias páginas con enlaces hacia j entonces es de común acuerdo que j es importante [8], es decir, cada vínculo con dirección a j se interpretan como un voto para j .

Ahora, si suponemos que j tiene un enlace asociado, el cual proviene de una página k que tiene reconocimiento en el tema de la búsqueda entonces este voto tiene más valor que el de otros enlaces de páginas menos relevantes. Además, si j sólo tiene un enlace asociado pero éste es de la página k entonces se considera que k transfiere su autoridad a j indicando que éste es importante [8].

PageRank calcula los valores de importancia de manera recursiva de la siguiente forma:

$$r_{n+1}(i) = \sum_{j \in C_i} \frac{r_n(j)}{E_j}$$

donde:

C_i es el conjunto de páginas que tienen un enlace que llega a la página i .

E_j es el número de enlaces de salida que tiene la página j .

$r_n(j)$ es el cálculo de importancia de la página j a la etapa n

El valor inicial de importancia asignado a la página será $\frac{1}{k}$, donde k es el número total de estados.

El algoritmo PageRank define la importancia de cada página basándose en las distribuciones estacionarias. De forma matricial, las cadenas de Markov están asociadas a un vector de probabilidades iniciales r , esto junto con la matriz de transición T determinan las probabilidades de cada estado en el

siguiente paso de la cadena por medio del vector dado por rT . Entre mayor sea la probabilidad de una página, mayor será el valor de importancia de ella para la red [39].

Sea r_n el vector formado por los valores PageRank calculados en la iteración n de las páginas web y sea T la matriz de transición asociada a la cadena de Markov dada por

$$T := (p_{i,j}) = \begin{cases} \frac{1}{E_j} & \text{si existe un enlace de } j \text{ a } i \\ 0 & \text{en otro caso} \end{cases} \quad (5.30)$$

Las probabilidades de llegar a una página específica se modificarán con cada iteración que se haga con la matriz de transición, para que esto suceda, el algoritmo PageRank establece que:

$$\text{PageRank} = r_{n+1} = r_n T^n \quad (5.31)$$

donde r es un vector fila que representa los valores de importancia de cada página. El vector de probabilidades iniciales tiene un valor equitativo para todas las entradas, es decir,

$$r_0(i) = \frac{1}{s} \quad \text{para } i=1,\dots,s$$

Entonces, el algoritmo iterará n veces de forma que para un n lo suficientemente grande el vector r presentará los valores de la distribución estacionaria de los estados de la cadena de Markov, en este caso de la localización del usuario en una página i en la etapa n , esto con base en el Teorema 1.7.

Para entender esto mejor, suponga una red como la que se presenta en la Figura 5.3, de acuerdo a la ecuación (5.30), se puede obtener la matriz de transición de la cadena de Markov asociada a ésta considerando las filas como el punto de partida del usuario y las columnas como el punto de llegada es decir, para obtener la entrada $T_{3,4}$ se tiene:

$$T_{3,4} = p_{D,C} = \frac{1}{E_C} = \frac{1}{3}$$

De forma que,

$$T = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad (5.32)$$

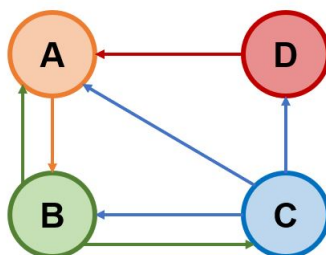


Figura 5.3: Representación de una red con cuatro páginas web en un grafo para el Ejemplo 5.3

Note que T cumple con las condiciones para ser una matriz estocástica. Para iniciar las iteraciones del algoritmo, se propone el vector r como:

$$r = \left(\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \right) \quad (5.33)$$

Es decir, inicialmente todas las entradas presentan el mismo valor de importancia para cada página. Ahora, sustituyendo (5.32) y (5.33) en la ecuación (5.31) e iterando para $n \geq 15$ se obtiene:

$$\begin{aligned} r &= (0.25 \quad 0.25 \quad 0.25 \quad 0.25) \\ rT &= (0.4583 \quad 0.3333 \quad 0.1250 \quad 0.0833) \\ rT^2 &= (0.3472 \quad 0.3472 \quad 0.2500 \quad 0.0556) \\ rT^3 &= (0.3148 \quad 0.4282 \quad 0.1852 \quad 0.0718) \\ &\vdots \\ rT^{15} &= (0.3333 \quad 0.4000 \quad 0.2000 \quad 0.0667) \end{aligned}$$

Por lo tanto, los valores de importancia de las páginas de esta red son:

$$PageRank = \begin{matrix} & A & B & C & D \\ (0.3333 & 0.4000 & 0.2000 & 0.0667) \end{matrix} \quad (5.34)$$

De (5.34) se interpreta que la página más relevante para la búsqueda del usuario es B .

En el caso anterior no se encontró problema alguno para realizar el cálculo PageRank, pero en ciertas ocasiones existen inconvenientes que este algoritmo básico puede presentar, uno de los cuales es la convergencia para llegar a la distribución estacionaria, pues ésta puede depender del vector de probabilidades inicial [35]. El Teorema 1.7 afirma que esto no sucederá si se

trabaja con una matriz de transición que cumpla ser aperiódica, irreducible y recurrente.

Algunos ejemplos de los problemas que pueden encontrarse en la matriz de transición, haciendo que no cumpla con lo deseado son, por ejemplo, cuando el usuario después de cierta etapa en su búsqueda, se encuentra en un *callejón sin salida* que es el momento en que la persona llega a un nodo que no cuenta con un enlace de salida. O en otro caso, cuando la búsqueda se estanca en una *trampa araña*, lo cual se refiere a que ésta se torna cíclica haciendo que no se pueda navegar por toda la red y por lo tanto se crea un problema en la distribución de los votos haciendo que el valor de importancia de las páginas no se vea modificado. Esto último puede verse claramente en la matriz de transición cuando todas las componentes de una misma fila son iguales a cero.

Estas cuestiones han sido consideradas y resueltas en el algoritmo PageRank, presentando algunas variaciones en la matriz de transición, la cual ahora estará dada por:

$$\overline{\overline{T}} = (1 - \beta)E + \beta\overline{T} \quad (5.35)$$

donde:

- El término $(1 - \beta)$ representa la probabilidad del usuario de reiniciar la búsqueda. β se conoce como el *factor de amortiguamiento o desecho* y normalmente se propone como un valor entre 0.8 y 0.9.
- $E = \frac{e^T e}{s}$, es una matriz de perturbaciones, e es un vector fila para el cual todas sus entradas son iguales a 1 y s es el orden de la matriz de transición original.
- \overline{T} es una modificación de la matriz de transición realizada previamente que se usa para hacer la matriz irreducible. \overline{T} se obtiene reemplazando la fila de ceros en la matriz de transición original T por una fila la forma $\frac{e}{s}$.

Una vez que se haya calculado la nueva matriz de transición $\overline{\overline{T}}$ entonces se procede a realizar el algoritmo básico visto previamente ecuación (5.31).

Para ejemplificar el procedimiento descrito anteriormente, suponga una red de páginas como las mostradas en el grafo de la Figura 5.4

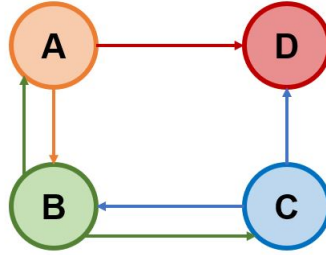


Figura 5.4: Grafo que representa una red compuesta de cuatro páginas web con un nodo sin enlaces de salida

La matriz de transición asociada a esta cadena de Markov está dada por:

$$T = \begin{matrix} & \begin{matrix} A & B & C & D \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

Si realizamos el algoritmo *PageRank* con esta matriz T , sin ninguna modificación, para un $n = 20$ se obtiene el siguiente vector:

$$PageRank = \begin{matrix} A & B & C & D \\ (0.2220 & 0.2220 & 0.2220 & 0.2220) \times 10^{-15} \end{matrix} \quad (5.36)$$

Note que (5.36) presenta un vector fila con todas las entradas iguales y que además pueden considerarse prácticamente iguales a cero, por lo tanto este vector no sirve para asignar un valor de importancia a cualquiera de las páginas.

Dado que la última fila de T es un renglón con entradas iguales a cero, esto implica que se está trabajando con una cadena reducible, por lo que, se necesita modificar la matriz de transición para poder aplicar el algoritmo.

Para poder aplicar el algoritmo, primero se debe obtener \bar{T} . Para ello se reemplaza la fila de entradas cero por el vector $\frac{e}{s}$, en este caso $s = 4$ entonces

$$\frac{e}{s} = \left(\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \right)$$

Por lo que \bar{T} quedaría como

$$\bar{T} = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \quad (5.37)$$

Para calcular $\bar{\bar{T}}$ se propone $\beta = 0.85$ y se calcula la matriz E

$$E = \frac{e^T e}{s} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

Entonces,

$$\begin{aligned} \bar{\bar{T}} &= 0.15 \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} + 0.85 \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \\ &= \begin{pmatrix} 0.0375 & 0.4625 & 0.0375 & 0.4625 \\ 0.4625 & 0.0375 & 0.4625 & 0.0375 \\ 0.0375 & 0.4625 & 0.0375 & 0.4625 \\ 0.2500 & 0.2500 & 0.2500 & 0.2500 \end{pmatrix} \end{aligned}$$

Ahora que se ha calculado $\bar{\bar{T}}$ se puede aplicar la ecuación (5.31), proponiendo el vector de probabilidades iniciales como:

$$r = \left(\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \right)$$

Iterando para $n \geq 20$ se obtiene que el vector de distribución estacionaria para este ejemplo es:

$$PageRank = \begin{matrix} & A & B & C & D \\ (0.2176 & 0.2824 & 0.2176 & 0.2824) \end{matrix} \quad (5.38)$$

De acuerdo a (5.38) las páginas B y D pueden considerarse como las más relevantes dentro del grafo, ambas teniendo el mismo valor de importancia.

Capítulo 6

Opciones financieras

Un *mercado financiero* puede ser definido como un espacio, ya sea físico o virtual, donde se realizan operaciones de compra-venta de activos financieros y en el cual se determina su precio [10], además se consideran en él los mecanismos y procedimientos, que son el medio para llevar a cabo dichas operaciones [40].

Se conoce como *activo financiero* a un título, generalmente emitido por alguna entidad económica, que otorga al comprador el derecho de recibir un ingreso futuro proveniente del vendedor [34].

Dentro de los mercados financieros se encuentran diferentes tipos de activos: aquellos que presentan riesgo y los que son libres de él. Cuando se conoce el rendimiento de un activo con certeza, es decir, las ganancias generadas por éste, de forma que se sabe exactamente cuál será su valor al finalizar un período establecido, entonces se le conoce como *activo libre de riesgo*. Los bonos y las cuentas de ahorro, son ejemplos de activos libres de riesgos. Por el contrario, las acciones y las tasas de interés son activos con riesgo.

En los mercados financieros se utilizan las estrategia de inversión o *portafolios*, que son un conjunto de distintos activos financieros en los que un individuo o institución ha invertido parte de su capital.

Las *opciones* son tipo de contrato que da el derecho, más no la obligación, de realizar una transacción sobre un activo financiero S llamado *subyacente*, para el cual se fija un *precio de ejercicio* K por un determinado periodo de tiempo T al cual se le conoce como *tiempo de ejecución*.

Si la opción da el derecho de adquirir el activo se conoce como *opción de compra* o *call*, por el contrario si da el derecho de venderlo entonces es llamado *opción de venta* o *put*.

Cuando el poseedor puede reclamar la compra o venta en cualquier momento del período se trata de una opción *americana*, pero si sólo puede hacerlo hasta que el período establecido termine, entonces se conoce como una opción *europea*.

Las opciones pueden verse como un seguro, puesto que permiten tener una cobertura ante fluctuaciones no favorables en los precios del activo subyacente. Por este derecho, el poseedor del contrato paga una prima.

Para ejemplificar esto suponga que una fábrica industrial tiene planes de comprar un tablero de automatización que un proveedor le ha ofrecido pero piensa hacerlo dentro de seis meses. Dicho tablero tiene un precio actual de \$500,000.00 pesos mexicanos.

La fábrica quiere asegurar que el precio actual sea el mismo (o lo más cercano posible) al que tendrá en la fecha de compra, por lo que necesita llegar a un acuerdo con el proveedor, con la intención de que, a cambio de pagar por un seguro hoy, supóngase de \$10,000.00, la fábrica podrá adquirir el tablero por este precio en seis meses.

Esta suposición tiene dos posibles escenarios, si al llegar a la fecha establecida el precio del tablero ha subido, digamos a \$580,000.00 pesos mexicanos, entonces a la fábrica le conviene ejercer la opción que ha adquirido, pagando así un total de \$510,000.00 por él. Si en cambio, el precio del tablero en el mercado resulta ser de \$450,000.00, entonces a la fábrica no le conviene ejercer el contrato, porque si lo compra a su precio actual estaría pagando por él un total de \$460,000.00, mucho más conveniente que el precio establecido.

Ahora, es evidente que el problema principal que presentan estos tipos de contratos es asignar el valor de la opción, es decir, cuál debería ser el precio de la prima que el poseedor debe pagar para poder adquirir el derecho sobre el activo, de forma que éste sea justo para las partes involucradas. En esta sección se presentarán dos de los métodos utilizados para ello, *el modelo de Cox-Ross-Rubinstein* y el modelo de *Black-Scholes*.

6.1. El modelo de Cox-Ross-Rubinstein

El modelo de Cox-Ross-Rubinstein se utiliza para realizar una estimación teórica del precio de las opciones europeas en períodos de tiempo discretos, basándose en un árbol binomial, donde las ramas representan las trayectorias posibles que el precio del activo puede tomar. Este modelo parte de la construcción de un portafolio compuesto del activo subyacente y la acción

correspondiente.

Para entender esta idea primero se analizará el modelo binomial conformado de un solo período. Además, para poder realizar el desarrollo de este modelo de una forma más simple se considera un mercado financiero ideal con las siguientes suposiciones:

- Todas las transacciones realizadas tienen costo nulo.
- Las tasas de interés son simétricas, es decir, el valor de la tasa de interés es el mismo tanto para prestar como para tomar prestado.
- No hay oportunidad de arbitraje, es decir, no existe la posibilidad de generar ganancias sin inversión de capital propio.

Considere dos activos financieros: uno de ellos con riesgo, por ejemplo una acción, y uno más sin riesgo, por ejemplo un bono, además de una opción europea tipo call sobre el activo con riesgo.

Suponga que en este escenario, la acción tiene un precio p_a cuando $t = 0$ y un precio aleatorio $S(t)$ para $t > 0$. El bono tiene un precio p_b al tiempo $t = 0$ y un precio determinístico $B(t)$ para $t > 0$. La opción tiene un precio de ejercicio K al finalizar el periodo de tiempo T .

El valor del activo sin riesgo evoluciona con respecto a la siguiente ecuación:

$$B(t) = p_b e^{rt} \quad t \in [0, T]$$

donde r es el valor del interés, el cual se supone constante.

Además, el modelo binomial supone que el rendimiento de la acción al finalizar el período T se comporta como un proceso binomial multiplicativo [20], es decir, el precio puede haber aumentado con probabilidad p o por el contrario, haber disminuido con probabilidad $(1 - p)$. Esto es,

$$S(T) = \begin{cases} up_a & \text{con probabilidad } p \\ dp_a & \text{con probabilidad } (1 - p) \end{cases} \quad (6.1)$$

donde $u > d$

El movimiento del valor de la acción obtenidos de la ecuación (6.1), puede representarse en un diagrama como el mostrado en de Figura 6.1.

Entonces, si una persona decide invertir una cantidad b en activos sin riesgo y un número a de acciones del activo con riesgo, para generar un portafolios (b, a) entonces su fortuna al tiempo T sería igual a:

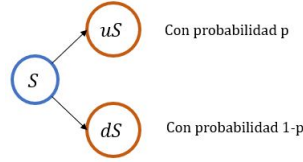


Figura 6.1: Árbol binomial uniperiodo

$$F(T) = \begin{cases} be^{rT} + ap_a u & \text{con probabilidad } p \\ be^{rT} + ap_a d & \text{con probabilidad } (1 - p) \end{cases} \quad (6.2)$$

Para evitar la posibilidad de arbitraje se debe cumplir que $d < e^{rT} < u$. Analicemos lo que pasaría si por ejemplo $e^{rT} < d < u$, esto significaría que el rendimiento del bono es menor que el de la acción, lo que haría más conveniente invertir en ella, es decir, se podría partir de cero, prestar del banco el valor de una acción, comprarla para venderla al finalizar el período de tiempo, de forma que pueda pagarse el crédito y además haber generado una ganancia sin haber invertido capital propio, lo que es arbitraje.

Ahora, si $e^{rT} > u > d$ entonces quien invierte podría asegurar generar ingresos vendiendo la acción e invirtiendo ese dinero en el bono, de forma que al finalizar el período tendrá el suficiente dinero para reponerla y además habrá obtenido ganancias, esto sin invertir capital propio, presentándose así nuevamente el escenario con oportunidad de arbitraje.

Considere que al expirar la opción, el valor de la acción cumple que $S(T) < K$, por lo tanto, como se ha mencionado anteriormente, no es conveniente para el poseedor de la opción reclamarla. En caso contrario, si $S(T) > K$ entonces es evidente que la opción se ejercerá. Entonces, el precio de la opción al momento de vencer está dado por:

$$(S(T) - K)^+ := \max\{0, S(T) - K\} = \begin{cases} S(T) - K & \text{si } S(T) > K \\ 0 & \text{si } S(T) \leq K \end{cases}$$

Defina C como el valor actual de la opción, con lo que, dependiendo de cómo se comporte el valor del activo S , se tienen las dos posibilidades siguientes:

$$C = \begin{cases} C_p := \max\{0, up_a - K\} & \text{con probabilidad } p \\ C_q := \max\{0, dp_a - K\} & \text{con probabilidad } 1 - p \end{cases} \quad (6.3)$$

Ahora, para calcular el valor de la opción se propone un portafolio autofinanciable que contiene un número de acciones a y una cantidad b de dinero en bonos, que además cumpla con los siguientes puntos:

- Es autofinanciado.
- Es una réplica de la opción.

Se dice que un portafolio es autofinanciado si cualquier cambio en el valor del portafolio (b, a) para $t > 0$ unicamente depende del aumento en el precio de la acción. El valor inicial V de este portafolio es igual a $V = b + ap_a$

Entonces, el valor del portafolio (b, a) al finalizar el período T está dado por:

$$V(T) = \begin{cases} be^{rT} + ap_a u & \text{con probabilidad } p \\ be^{rT} + ap_a d & \text{con probabilidad } 1 - p \end{cases} \quad (6.4)$$

El hecho de que el portafolio sea una réplica de la opción, implica que

$$C_p := \text{máx}\{0, up_a - K\} = be^{rT} + ap_a u \quad (6.5)$$

$$C_q := \text{máx}\{0, dp_a - K\} = be^{rT} + ap_a d \quad (6.6)$$

Lo que se desea encontrar son las variables que componen el valor del portafolio, a y b , para ello se puede obtener del sistema de ecuaciones formado por (6.5) y (6.6) lo siguiente:

De (6.6),

$$be^{rT} = C_q - ap_a d \quad (6.7)$$

Sustituyendo en (6.5)

$$\begin{aligned} C_p &= (C_q - p_a ad) + ap_a u \\ &= C_q + ap_a(u - d) \\ C_p - C_q &= ap_a(u - d) \\ ap_a &= \frac{C_p - C_q}{u - d} \end{aligned} \quad (6.8)$$

De (6.8) en (6.5) se obtiene,

$$\begin{aligned}
 be^{rT} &= C_p - \left(\frac{C_p - C_q}{u - d} \right) u \\
 &= \frac{C_p(u - d) - (C_p u - C_q u)}{u - d} \\
 &= \frac{C_p u - C_p d - C_p u + C_q u}{u - d} \\
 &= \frac{C_q u - C_p d}{u - d}
 \end{aligned}$$

Por lo tanto

$$b = e^{-rT} \frac{C_q u - C_p d}{u - d} \quad (6.9)$$

El precio justo de la opción al tiempo T está dado por el valor del portafolio V , es decir,

$$C = V = b + ap_a$$

De (6.9) y (6.8), se obtiene que el valor de la opción para el modelo uniperiódico es

$$\begin{aligned}
 C &= b + ap_a \\
 &= e^{-rT} \left(\frac{C_q u - C_p d}{u - d} \right) + \frac{C_p - C_q}{u - d} \\
 &= \frac{e^{rT}(C_p - C_q) + C_q u - C_p d}{e^{rT}(u - d)} \\
 &= \frac{e^{rT} C_p - e^{rT} C_q + C_q u - C_p d}{e^{rT}(u - d)} \\
 &= \frac{C_p(e^{rT} - d) + C_q(u - e^{rT})}{e^{rT}(u - d)} \\
 &= \frac{C_p}{e^{rT}} \left(\frac{e^{rT} - d}{u - d} \right) + \frac{C_q}{e^{rT}} \left(\frac{u - e^{rT}}{u - d} \right) \\
 &= e^{-rT} \left[C_p \left(\frac{e^{rT} - d}{u - d} \right) + C_q \left(\frac{u - e^{rT}}{u - d} \right) \right]
 \end{aligned}$$

Si se define $p^* = \frac{e^{rT} - d}{u - d}$ entonces

$$\begin{aligned}
C &= e^{-rT} \left[C_p \left(\frac{e^{rT} - d}{u - d} \right) + C_q \left(\frac{u - e^{rT}}{u - d} \right) \right] \\
&= e^{-rT} [C_p p^* + C_q (1 - p^*)]
\end{aligned} \tag{6.10}$$

Es importante notar algunos hechos de la fórmula obtenida en (6.10), por ejemplo, ésta no depende de la probabilidad p sino de p^* . Además, la única variable aleatoria de que influye en el valor de la opción de venta es el precio de la acción [20].

Si se calcula la esperanza con respecto a p^* , se tiene que:

$$\begin{aligned}
E_{p^*}(e^{-rT}S(T)) &= e^{-rT} (up_a p^* + dp_a(1 - p^*)) \\
&= e^{-rT} \left[up_a \left(\frac{e^{rT} - d}{u - d} \right) + dp_a \left(1 - \frac{e^{rT} - d}{u - d} \right) \right] \\
&= \frac{p_a}{e^{rT}} \left[\left(\frac{ue^{rT} - ud}{u - d} \right) + \left(\frac{ud - e^{rT}d}{u - d} \right) \right] \\
&= \frac{p_a}{e^{rT}} \left[\frac{ue^{rT} - ud + ud - e^{rT}d}{u - d} \right] \\
&= \frac{p_a}{e^{rT}} \left[\frac{e^{rT}(u - d)}{(u - d)} \right] \\
&= p_a
\end{aligned}$$

Note que $e^{-rT}S(T)$ cumple con la propiedad de una martingala. Observe también que $0 < p^* < 1$, lo que implicaría que p^* tiene propiedades de una probabilidad, la cual puede expresarse como $p^* = P^*\{S(T) = up_a\}$ y que se conoce como *medida martingala*, *medida neutral con respecto al riesgo* o *probabilidad libre de riesgo*.

Ahora, suponga el escenario donde la valuación de la opción involucra dos períodos en lugar de uno. Considerando lo visto anteriormente, el valor del activo subyacente, después del primer período también tendría dos posibilidades, lo cual podría verse de forma gráfica representado en la Figura 6.2. El valor de la opción se vería afectado de la misma manera, lo cual se muestra en el diagrama de la Figura 6.3.

Del análisis anterior, se tiene que el valor de la opción es:

$$C = e^{-rT} [C_p p^* + C_q (1 - p^*)] \tag{6.11}$$

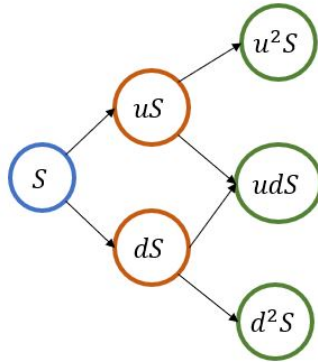


Figura 6.2: Árbol binomial de dos períodos para el valor de activo subyacente

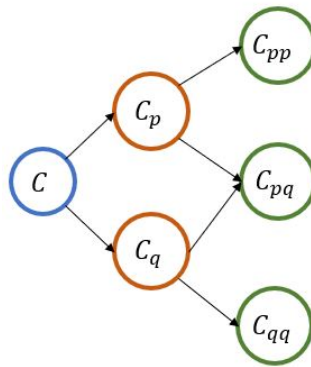


Figura 6.3: Árbol binomial de dos períodos para el valor de la opción

Para este caso, al finalizar el primer período se tendrá un período restante, quedando entonces un problema uniperiódico [20], idéntico al que se resolvió al principio de este apartado.

Considere que para este nuevo caso se tienen períodos de longitud $\frac{T}{2}$, lo cual implicaría que el valor del activo sin riesgo estaría dado como $e^{\frac{rT}{2}}$.

Por lo que, la ecuación (6.11) se escribiría como:

$$C = e^{-\frac{rT}{2}} [C_p p^* + C_q (1 - p^*)] \quad (6.12)$$

Para resolver (6.12) se necesitan conocer los valores de C_p y C_q , los cuales

están dados como:

$$\begin{aligned} C_p &= e^{-\frac{rT}{2}} [C_{pp}p^* + C_{pq}(1 - p^*)] \\ C_q &= e^{-\frac{rT}{2}} [C_{pq}p^* + C_{qq}(1 - p^*)] \end{aligned}$$

donde

$$p^* = \frac{e^{\frac{rT}{2}} - d}{u - d}$$

y

$$\begin{aligned} C_{pp} &= \text{máx} [0, u^2 p_a - K] \\ C_{pq} &= \text{máx} [0, u d p_a - K] \\ C_{qq} &= \text{máx} [0, d^2 p_a - K] \end{aligned}$$

Con lo que el valor de la opción C quedaría como:

$$\begin{aligned} C &= e^{-\frac{rT}{2}} \left[\left(e^{-\frac{rT}{2}} [C_{pp}p^* + C_{pq}(1 - p^*)] \right) p^* \right] \\ &\quad + e^{-\frac{rT}{2}} \left[\left(e^{-\frac{rT}{2}} [C_{pq}p^* + C_{qq}(1 - p^*)] \right) (1 - p^*) \right] \\ &= e^{-\frac{rT}{2}} \left[e^{-\frac{rT}{2}} (C_{pp}p^{*2} + 2C_{pq}p^*(1 - p^*) + C_{qq}(1 - p^*)^2) \right] \\ &= e^{-rT} [C_{pp}p^{*2} + 2C_{pq}p^*(1 - p^*) + C_{qq}(1 - p^*)^2] \end{aligned}$$

Si se repite este procedimiento incrementando el número de períodos, iniciando desde el período de expiración y sustituyendo en cada paso hasta llegar al período cero, se obtendrá [20]:

$$C = e^{-rT} \left(\sum_{i=0}^n \left(\frac{n!}{i!(n-i)!} \right) p^{*i} (1 - p^*)^{n-i} \text{máx}[0, u^i d^{n-i} p_a - K] \right) \quad (6.13)$$

donde

$$p^* = \frac{e^{\frac{rT}{n}} - d}{u - d}$$

El modelo multiperiodico será libre de arbitraje si se cumple que $d < e^{\frac{rT}{n}} < u$ [20].

La ecuación (6.13) puede escribirse como

$$C = e^{-rT} \left[\sum_{i=a}^n \left(\frac{n!}{i!(n-i)!} \right) p^{*i} (1-p^*)^{n-i} [u^i d^{n-i} p_a - K] \right] \quad (6.14)$$

donde a se define como el entero no negativo más pequeño que cumple con la desigualdad $u^a d^{n-a} p_a > K$, de forma que el $\max [u^a d^{n-a} p_a - K, 0] = u^a d^{n-a} p_a - K$ [20], en otras palabras, a es el mínimo valor para el cual al expirar el contrato éste termina *in the money*, lo que significa que al momento de ejercer la opción ésta genera beneficios al poseedor.

Note que

$$a > \frac{\log \left(\frac{K}{p_a d^n} \right)}{\log \left(\frac{u}{d} \right)}$$

Entonces (6.14) quedaría como

$$\begin{aligned} C &= e^{-rT} \left[\sum_{i=a}^n \left(\frac{n!}{i!(n-i)!} \right) p^{*i} (1-p^*)^{n-i} (u^i d^{n-i} p_a) \right] \\ &\quad - e^{-rT} \left[\sum_{i=a}^n \left(\frac{n!}{i!(n-i)!} \right) p^{*i} (1-p^*)^{n-i} K \right] \\ &= p_a \left[\sum_{i=a}^n \left(\frac{n!}{i!(n-i)!} \right) p^{*i} (1-p^*)^{n-i} \left(\frac{u^i d^{n-i}}{e^{rT}} \right) \right] \\ &\quad - e^{-rT} K \left[\sum_{i=a}^n \left(\frac{n!}{i!(n-i)!} \right) p^{*i} (1-p^*)^{n-i} \right] \end{aligned} \quad (6.15)$$

Observe que

$$\sum_{i=a}^n \left(\frac{n!}{i!(n-i)!} \right) p^i (1-p)^{n-i}$$

es la expresión de una función de distribución binomial (ver Definición 1.8).

La ecuación (6.15) puede ser expresada en términos de funciones de distribución binomial, para ello, se tiene que:

$$\begin{aligned}
 p^{*i}(1-p^*)^{n-i} \left(\frac{u^i d^{n-i}}{e^{rT}} \right) &= \frac{p^{*i} u^i (1-p^*)^{n-i} d^{n-i}}{e^{rT}} \\
 &= \frac{(up^*)^i [(1-p^*)d]^{n-i}}{e^{rT}} \\
 &= \left(\frac{up^*}{e^{\frac{rT}{n}}} \right)^i \left(\frac{(1-p^*)d}{e^{\frac{rT}{n}}} \right)^{n-i}
 \end{aligned}$$

Definiendo

$$\begin{aligned}
 p' &= \frac{up^*}{e^{\frac{rT}{n}}} \\
 1-p' &= 1 - \frac{up^*}{e^{\frac{rT}{n}}} \\
 &= \frac{e^{\frac{rT}{n}} - up^*}{e^{\frac{rT}{n}}} \\
 &= e^{-\frac{rT}{n}} \left[e^{\frac{rT}{n}} - u \left(\frac{e^{\frac{rT}{n}} - d}{u-d} \right) \right] \\
 &= e^{-\frac{rT}{n}} \left[\frac{e^{\frac{rT}{n}}(u-d) - u(e^{\frac{rT}{n}} - d)}{u-d} \right] \\
 &= e^{-\frac{rT}{n}} \left[\frac{d(u - e^{\frac{rT}{n}})}{u-d} \right] \\
 &= \frac{d}{e^{\frac{rT}{n}}} \left[\frac{u - e^{\frac{rT}{n}}}{u-d} \right] \\
 &= \frac{d}{e^{\frac{rT}{n}}} [1 - p^*]
 \end{aligned}$$

Entonces la ecuación (6.15) quedaría como

$$\begin{aligned}
 C &= p_a \left[\sum_{i=a}^n \left(\frac{n!}{i!(n-i)!} \right) p' (1-p')^{n-i} \right] \\
 &\quad - e^{-rT} K \left[\sum_{i=a}^n \left(\frac{n!}{i!(n-i)!} \right) p^{*i} (1-p^*)^{n-i} \right]
 \end{aligned} \tag{6.16}$$

Por lo tanto, de (6.16) se obtiene una expresión de la forma:

$$C = p_a \phi [a; n, p'] - e^{-rT} K \phi [a; n; p^*] \quad (6.17)$$

donde

$$p^* = \frac{e^{\frac{rT}{n}} - d}{u - d}$$

$$p' = e^{\frac{-rT}{n}} p^* u$$

$$\phi [a; n; x] = \sum_{i=a}^n \binom{n}{i} x^i (1-x)^{n-i}$$

Para visualizar el método anterior, se ha propuesto el Ejemplo 6.1

Ejemplo 6.1 *El dueño de una cadena de pastelerías chiapaneca sabe que uno de los productos básicos para la producción de su empresa es el azúcar y durante los últimos meses el precio de ésta ha variado considerablemente (véase Tabla 6.1), lo cual afecta tanto sus ventas como sus márgenes de ganancia, por lo que ha llegado a un acuerdo con su proveedor de la Central de Abastos para firmar una opción de compra con un plazo de vencimiento de un año, en la cual él se compromete a venderle el saco de azúcar a \$750.00 pesos mexicanos, en tres períodos. El precio actual del saco al día de la firma del contrato es de \$742.37. Se desea establecer el precio de la prima a pagar por el dueño de la empresa pastelera. Suponga que la tasa libre de riesgo es del 2%.*

De acuerdo a la Tabla 6.1 y aplicando el procedimiento de *volatilidad histórica* mostrado en [7], la volatilidad mensual de este producto es de 5.45% para más detalle ver Apéndice B.

Ahora, para el desarrollo del Ejemplo 6.1 se tienen los siguientes datos:

- Precio del activo subyacente en el día de la firma: \$742.37
- Precio de ejercicio: \$750.00
- Duración de la opción: 1 año en tres periodos ($n = 3$).
- Volatilidad $\sigma = 5.45\%$
- Tasa sin riesgo $r = 2\%$

Mes	Precio
Enero	\$608.85
Febrero	\$718.62
Marzo	\$712.34
Abril	\$747.01
Mayo	\$783.53
Junio	\$770.63
Julio	\$764.53
Agosto	\$752.13
Septiembre	\$736.17
Octubre	\$772.55
Noviembre	\$759.69
Diciembre	\$742.37

Tabla 6.1: Precio del azúcar en C.A. Tuxtla Gutiérrez, Chiapas

De acuerdo a [7], los valores de u y d pueden ser calculados aplicando la siguiente relación:

$$u = e^{\sigma\sqrt{\frac{T}{n}}}$$

$$d = \frac{1}{u}$$

esto con la finalidad de que tanto u como d dependan del número de períodos n y puedan obtenerse resultados realistas conforme la cantidad de períodos aumenta en un intervalo de tiempo fijo. Para más detalle de esta relación puede consultar [20].

En este caso u y d quedarían como:

$$u = e^{5.45\% \sqrt{\frac{1}{3}}} = 1.0320$$

$$d = \frac{1}{1.0320} = 0.9690$$

Se calculan las probabilidades p y q asociadas a los eventos u y d respectivamente.

$$p^* = \frac{e^{\frac{rT}{n}} - d}{u - d} = \frac{1.0067 - 0.9690}{1.0320 - 0.9690} = 0.5983$$

$$q = 1 - p^* = 1 - 0.5983 = 0.4017$$

La evolución del precio del activo se muestra en la Figura 6.4.

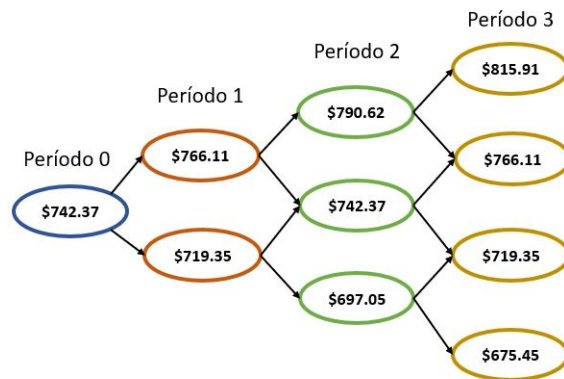


Figura 6.4: Diagrama de evolución del precio del activo subyacente.

Para el período de expiración, i.e, el período 3 se tiene que

$$C_{ppp} = \text{máx} [815.9194 - 750, 0] = 65.9194$$

$$C_{ppq} = \text{máx} [766.1186 - 750, 0] = 16.1186$$

$$C_{qpq} = \text{máx} [719.3575 - 750, 0] = 0$$

$$C_{qqq} = \text{máx} [675.4506 - 750, 0] = 0$$

Para el período 2, calculamos recursivamente

$$\begin{aligned}
 C_{pp} &= \frac{C_{ppp}p^* + C_{ppq}(1 - p^*)}{e^{\frac{rT}{n}}} \\
 &= \frac{65.9194(0.5983) + 16.1186(0.4017)}{1.0067} \\
 &= 45.6104 \\
 C_{pq} &= \frac{C_{ppq}p^* + C_{ppq}(1 - p^*)}{e^{\frac{rT}{n}}} \\
 &= \frac{16.1186(0.5983) + 0(0.4017)}{1.0067} \\
 &= 9.5800 \\
 C_{qq} &= \frac{C_{ppq}p^* + C_{ppq}(1 - p^*)}{e^{\frac{rT}{n}}} \\
 &= \frac{0(0.5983) + 0(0.4017)}{1.0067} \\
 &= 0
 \end{aligned}$$

Para el período 1, se calcula

$$\begin{aligned}
 C_p &= \frac{C_{pp}p^* + C_{pq}(1 - p^*)}{e^{\frac{rT}{n}}} \\
 &= \frac{45.6104(0.5983) + 9.5800(0.4017)}{1.0067} \\
 &= 30.9308 \\
 C_q &= \frac{C_{pq}p^* + C_{qq}(1 - p^*)}{e^{\frac{rT}{n}}} \\
 &= \frac{9.5800(0.5983) + 0(0.4017)}{1.0067} \\
 &= 5.6938
 \end{aligned}$$

Finalmente, para el período acutal

$$\begin{aligned}
 C &= \frac{C_p p^* + C_q (1 - p^*)}{e^{\frac{rT}{n}}} \\
 &= \frac{30.9308(0.5983) + 5.6938(0.4017)}{1.0067} \\
 &= 20.6555
 \end{aligned}$$

Por lo tanto el precio de la opción de compra es 20.6555

La evolución del precio de la opción puede verse en el diagrama de la Figura 6.5.

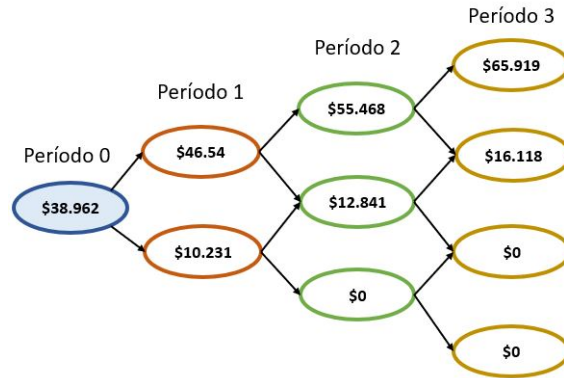


Figura 6.5: Diagrama de evolución del precio de la opción de compra.

El procedimiento anterior se desarrolló para tener una mejor idea de la deducción de la fórmula, ahora si se desea aplicar la fórmula directamente, entonces se realizan los pasos mostrados a continuación.

Dados los siguientes datos:

$$n = 3 \quad u = 1.0320 \quad d = 0.9690 \quad p_a = 742.37 \quad K = 750$$

Se calcula a , recordando que se trata del mínimo entero no negativo para el cual se cumple:

$$a > \frac{\log\left(\frac{K}{p_a d^n}\right)}{\log\left(\frac{u}{d}\right)} = \frac{\log(1.1104)}{\log(1.0650)} = 1.6628 \Rightarrow a = 2$$

Se obtienen los valores de p^* y p'

$$p^* = \frac{e^{\frac{rT}{n}} - d}{u - d} = \frac{1.0067 - 0.9690}{1.0320 - 0.9690} = 0.5983$$

$$p' = \frac{p'u}{e^{\frac{rT}{n}}} = \frac{0.5983(1.0320)}{1.0067} = 0.6133$$

Se prosigue desarrollando la sumatoria y calculando los valores de las probabilidades con distribución binomial

$$\begin{aligned}
 \phi [2; 3, p'] &= \sum_{i=2}^3 \binom{3}{2} p'^i (1 - p')^{3-i} \\
 &= \frac{3!}{2!(3-2)!} p'^2 (1 - p')^{3-2} + \frac{3!}{3!(3-3)!} p'^3 (1 - p')^{3-3} \\
 &= \frac{3!}{2!} p'^2 (1 - p') + p'^3 \\
 &= 0.6671
 \end{aligned}$$

$$\begin{aligned}
 \phi [2; 3, p^*] &= \sum_{i=2}^3 \binom{3}{2} p^{*i} (1 - p^*)^{3-i} \\
 &= \frac{3!}{2!(3-2)!} p^{*2} (1 - p^*)^{3-2} + \frac{3!}{3!(3-3)!} p^{*3} (1 - p^*)^{3-3} \\
 &= \frac{3!}{2!} p^{*2} (1 - p^*) + p^{*3} \\
 &= 0.6456
 \end{aligned}$$

Con lo que,

$$\begin{aligned}
 C &= p_a \phi [2, 3, p'] - e^{-\frac{rT}{n}} K \phi [2, 3, p^*] \\
 C &= 742.37(0.6671) - \frac{750(0.6456)}{1.0067} \\
 &= 20.6555
 \end{aligned}$$

Llegando así a que el valor de la opción de compra para el Ejemplo 6.1 es de \$20.6555

6.2. El modelo de Black-Scholes

El modelo de Black-Scholes, también conocido como el modelo de *Black-Scholes-Merton (BSM)* es una ecuación propuesta por los matemáticos Fischer Black y Myron Scholes, con aportaciones del economista Merton Miller, publicada en 1973 [19]. Dicha ecuación está basada en lo que se conocen como *procesos estocásticos* (ver Definición 1.16) y, en un principio, tenía el objetivo de calcular el precio justo de las opciones de compras europeas.

La publicación del modelo BSM coincidió con la apertura de *Chicago Board Options Exchange*, el primer mercado de opciones del mundo [19], por lo que comenzó a ser utilizado tanto por los creadores de este mercado como por cientos de inversionistas dentro de EE.UU. Este método tuvo tanto éxito, en ese entonces como ahora, que se convirtió en el principal para la valoración de diferentes activos financieros.

De la misma forma que el modelo de Ross-Cox-Rubinstein hace ciertas suposiciones sobre el mercado financiero, el modelo BSM supone los siguientes puntos:

- La volatilidad σ y la tasa libre de riesgo r permanecen constantes.
- Cualquier transacción realizada tiene costo nulo.
- No existe oportunidad de arbitraje.
- Las tasas de interés son simétricas, es decir, la tasa de interés tiene el mismo valor tanto para prestar como para tomar prestado.
- La opción es de tipo europeo.
- El comportamiento del precio del activo subyacente S sigue una distribución normal logarítmica (ver Definición 1.7).

Dicho de una forma muy general, el principio de este modelo se basó en construir un portafolio autofinanciado que replicara el valor de la opción al tiempo de ejecución, para lo cual se implementó una función derivable que relacionaba el valor de la acción con el tiempo de validez de la opción y el precio del activo subyacente.

Un punto interesante a resaltar es el hecho de que el modelo binomial, desarrollado en la sección anterior, converge a la fórmula de Black-Scholes-Merton cuando el tiempo de ejercicio de la opción es dividido en subintervalos lo suficientemente pequeños. Las relaciones entre ambos métodos pueden consultarse en [21].

En este apartado únicamente se expondrán las fórmulas que componen al modelo de Black-Scholes-Merton, para los detalles de la demostración se recomienda ver [21] y [33].

La ecuación de Black-Scholes-Merton está dada por:

$$C = N(d_1(t))S(t) - Ke^{-r(T-t)}N(d_2(t)) \quad (6.18)$$

$$d_1(t) = \frac{\ln\left(\frac{S(t)}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T - t)}{\sigma\sqrt{T - t}}$$

$$d_2(t) = d_1(t) - \sigma\sqrt{T - t}$$

donde:

t es el tiempo para el cual se inicia la validez de la opción.

T es el tiempo expresado en años hasta la fecha de expiración.

σ es la volatilidad de la opción.

r es la tasa libre de riesgo.

$S(t)$ es el precio del activo subyacente.

K es el precio de ejercicio determinado.

C el valor de la opción de compra.

$N(\cdot)$ es la probabilidad de una distribución normal estándar para el punto evaluado. La función normal estándar está dada por la expresión (1.4) cuando $\sigma = 1$ y $\mu = 0$.

Para ejemplificar el uso de este método se propone el Ejemplo 6.2

Ejemplo 6.2 *Un floricultor del Estado de México que distribuye a la Central de Abastos en Iztapalapa sabe que cada año el precio de las gladiolas por gruesa (144 piezas) baja considerablemente en el mes de agosto afectando el precio de venta de su producción y por consecuente sus ganancias. A inicios del año siguiente, ha decidido proponer a sus compradores firmar una opción de venta por un precio de ejercicio de \$450.00 por gruesa, por lo que necesita calcular cuál sería el precio justo a pagar por la prima para ambas partes involucradas. Supóngase una tasa libre de riesgo del 3%*

Los datos del Ejemplo 6.2 pueden resumirse como:

- Tiempo de ejercicio $T = 0.66$ años y $t = 0$, esto debido a que la opción se hará válida hasta agosto.
- Precio de ejercicio $K = \$450.00$ por gruesa.
- Precio al día de la firma del contrato $S(0) = \$581.00$
- Tasa libre de riesgo $r = 3\%$

Mes	Precio
Enero	\$422.10
Febrero	\$469.10
Marzo	\$44.38
Abril	\$541.67
Mayo	\$515.20
Junio	\$452.50
Julio	\$346.53
Agosto	\$606.60
Septiembre	\$369.00
Octubre	\$451.40
Noviembre	\$493.50
Diciembre	\$581.00

Tabla 6.2: Precio de la gruesa de Gladiolas en C.A. Iztapalapa, CDMX.

Como referencia para este problema, en la Tabla 6.2 pueden verse los precios mensuales de las gladiolas en 2019, los cuales fueron obtenidos de [49], dichos precios sirven para calcular la volatilidad anual σ . Esta volatilidad es conocida como *volatilidad histórica* y el procedimiento detallado para obtenerla puede consultarse en el Apéndice B. Con los valores de la Tabla 6.3, se obtiene que la volatilidad mensual es:

$$\sigma_m = \sqrt{\frac{0.2590}{11}} \times 100 = 15.34\%$$

y la volatilidad anual es:

$$\sigma = (\sqrt{12} \times 0.2590) \times 100 = 53.16\%$$

Ahora, si se supone que la opción de venta no puede ejercerse hasta llegar al tiempo de ejercicio, entonces existe una relación de paridad compra-venta de las opciones europeas que implica lo siguiente [21]:

$$P(t) = C(t) - S(t) + Ke^{-r(T-t)} \quad (6.19)$$

donde P es el valor de la opción de venta.

Cálculo de Volatilidad					
Mes	Precio Mensual	Pn/Pn-1	Ln	Desviación	Desviación ²
12	581	1.1773	0.1632	0.1341	0.01800
11	493.50	1.0932	0.08916	0.06001	0.0036
10	451.40	1.2233	0.2015	0.1725	0.0297
9	369.00	1.2035	0.1852	0.1562	0.02440
8	306.60	0.8847	-0.1224	-0.1514	0.0229
7	346.53	0.7658	-0.2668	-0.2958	0.0875
6	452.5	0.8782	-0.1297	-0.1588	0.0252
5	515.20	0.9511	-0.0501	0.0791	0.0062
4	541.67	1.2189	0.1979	0.1689	0.0285
3	444.38	0.9473	-0.0541	-0.0831	0.0069
2	469.10	1.1113	0.1055	0.0765	0.0058
1	422.10				
		Media Ln	0.0290	Suma	0.2590

Tabla 6.3: Cálculo de volatilidad para el Ejemplo 6.2

Entonces, se calculará el valor de la opción de compra para después obtener la de venta.

$$\begin{aligned}
 d_1(0) &= \frac{\ln\left(\frac{S(0)}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T)}{\sigma\sqrt{T}} \\
 &= \frac{\ln\left(\frac{581}{450}\right) + \left(0.03 + \frac{(0.5316)^2}{2}\right)(0.66)}{(0.5316)\sqrt{0.66}} \\
 &= 0.8534
 \end{aligned}$$

$$\begin{aligned}
 d_2(0) &= d_1(0) - \sigma\sqrt{T} \\
 &= 0.8534 - [(0.5316)(0.66)] \\
 &= 0.4215
 \end{aligned}$$

Una vez que se han calculado los valores d_1 y d_2 se necesitan calcular el valor de la función de distribución normal acumulada correspondiente. Para

ello pueden utilizarse las tablas que se encuentran usualmente en los libros de probabilidad para la función de distribución normal estándar, si lo desea puede consultar [41].

En este caso, estos valores se ha calculado con ayuda de la herramienta de cómputo MatLab. Obteniendo los valores siguientes:

$$N(d_1) = 0.8033$$

$$N(d_2) = 0.6633$$

El valor de la opción de compra es

$$\begin{aligned} C &= N(d_1)S(0) - Ke^{-r(T-t)}N(d_2) \\ &= (0.8033 \times 581) - (450 \times e^{-(0.03)0.66}) \times 0.6633 \\ &= 174.0669 \end{aligned}$$

Con los datos anteriores, el valor de la opción de venta quedaría como:

$$\begin{aligned} P &= C - S(0) + Ke^{-r(T-t)} \\ &= 174.0669 - 581 + \frac{450}{1.0200} \\ &= 34.2445 \end{aligned}$$

Para entender un poco más el comportamiento del modelo Black-Scholes consideremos el cambio en el valor de la opción de compra C conforme los parámetros que lo componen ($S(t)$, K , t , σ , r) varían.

Observe la Tabla 6.4, donde se ha fijado el valor del activo subyacente al día de la firma del contrato y se han variado los parámetros de volatilidad σ , precio de ejercicio K , tiempo de expiración T y la tasa libre de riesgo r .

Se han propuesto dos tiempos de ejercicio, el primero a mitad del año $T = 0.5$ y el segundo a un año $T = 1$. Note como el precio de las opciones, fijando todos los demás parámetros, es mayor conforme el período de tiempo incrementa. Es decir, una opción tendrá más valor entre mayor sea el período de vigencia.

Se tienen tres valores para el precio pactado de compra, $K = 400$, $K = 500$ y $K = 600$, estos valores se interpretarían como que la opción terminó *out the money*, *at the money* y *in the money* respectivamente. Observe que

S=500							
		r=3 %		r=10 %		r=20 %	
σ	K	Tiempo de expiración en años					
		0.5	1	0.5	1	0.5	1
15 %	400	106.15	113.06	119.58	138.41	138.07	172.55
	500	24.92	37.42	10.68	58.34	52.25	93.50
	600	1.41	6.53	2.87	14.29	6.92	34.21
30 %	400	112.10	126.41	123.81	147.15	140.56	176.82
	500	45.74	66.41	54.53	83.67	68.38	111.01
	600	14.09	31.45	18.32	43.03	25.85	63.57
40 %	400	13.92	139.68	130.40	158	145.60	184.41
	500	59.61	85.69	67.90	101.59	80.63	126.06
	600	26.15	50.69	31.25	62.9290	39.64	83.02

Tabla 6.4: Valores de opciones con parámetros variantes.

el precio de la opción disminuye cuando el precio de ejercicio aumenta. Esto puede interpretarse de la siguiente manera: entre mayor sea el valor que tengamos que pagar por el activo subyacente en el presente, lo más seguro es que al tiempo de ejercicio de la opción el precio de ésta en el mercado esté por debajo de lo pactado y por lo tanto la opción no se ejercerá. Por el contrario, si el precio de ejercicio decrece más y más, el valor de la opción tenderá al valor del activo subyacente.

Ahora, note el cambio que existe en el precio de la opción cuando la volatilidad varía, note que entre mayor sea la volatilidad mayor es el precio de la prima a pagar, esto porque una mayor volatilidad implica una mayor probabilidad de fluctuaciones en el valor del activo subyacente S y si se hace la suposición de que al finalizar el período el precio de S es mucho más bajo en el mercado que lo pactado en la opción entonces ésta no se ejercerá, aunque por el contrario, si el precio de la acción es muy alto entonces el contrato se utilizará.

Si fijamos todos los valores, con excepción de la tasa libre de riesgo r , entonces se podrá observar que el precio de la opción aumenta conforme r lo hace. Entre mayor sea el valor de r más se aproxima el precio de la opción al valor del activo subyacente [21].

Formalmente, los cambios en el valor de la opción son conocidos como la *sensibilidad* de la fórmula y pueden cuantificarse utilizando las derivadas parciales de C o P dependiendo si se calcula para la opción de compra o de venta respectivamente. Por ejemplo, si se sabe que existe un incremento n en el precio del activo subyacente S para una opción de compra entonces el valor de la opción cambiará con respecto a

$$n \left(\frac{\partial C}{\partial S} \right) = nN(d_1)$$

A estas derivadas parciales se les conoce como *griegas* debido a la representación utilizada para identificarlas, dentro de ellas se encuentran: *delta*, *rho*, *theta* y *vega*.

Delta es la medida de sensibilidad que se basa en las variaciones del precio del mercado del activo subyacente, para una opción de compra se obtiene derivando (6.18) con respecto a S . Esto es,

$$\delta = \frac{\partial C}{\partial S} = S \frac{\partial}{\partial S} N(d_1) + N(d_1) \frac{\partial}{\partial S} S - Ke^{r(T-t)} \frac{\partial}{\partial S} N(d_2) \quad (6.20)$$

Por la regla de la cadena se tiene que,

$$\frac{\partial}{\partial x} N(d(x)) = N'(d(x)) \frac{\partial}{\partial x} d(x) \quad (6.21)$$

donde $N'(x) = \frac{1}{\sqrt{2\pi}} e^{-(x)^2/2}$ es la función de densidad para una distribución normal estándar evaluada en el punto correspondiente y x representa la variable respecto a la cual se está derivando en cada caso.

Aplicando (6.21) en (6.20)

$$\begin{aligned} \frac{\partial C}{\partial S} &= SN'(d_1) \frac{\partial}{\partial S} d_1 + N(d_1) - Ke^{r(T-t)} N'(d_2) \frac{\partial}{\partial S} d_2 \\ &= SN'(d_1) \frac{\partial}{\partial S} d_1 + N(d_1) - Ke^{r(T-t)} N'(d_2) \frac{\partial}{\partial S} (d_1 - \sigma \sqrt{(T-t)}) \\ &= SN'(d_1) \frac{\partial}{\partial S} d_1 + N(d_1) - Ke^{r(T-t)} N'(d_2) \frac{\partial}{\partial S} d_1 \end{aligned}$$

Además,

$$\begin{aligned}
N'(d_2) &= N'(d_1 - \sigma\sqrt{T-t}) \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(d_1 - \sigma\sqrt{T-t})^2}{2}\right) \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{d_1^2 - 2d_1\sigma\sqrt{T-t} + \sigma^2(T-t)}{2}\right) \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{d_1^2}{2}\right) \exp\left(d_1\sigma\sqrt{T-t} - \frac{\sigma^2(T-t)}{2}\right) \\
&= N'(d_1) \exp\left(d_1\sigma\sqrt{T-t} - \frac{\sigma^2(T-t)}{2}\right) \\
&= N'(d_1) \exp\left(\left[\frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\right] \sigma\sqrt{T-t} - \frac{\sigma^2(T-t)}{2}\right) \\
&= N'(d_1) \exp\left(\ln\left(\frac{S}{K}\right) + r(T-t)\right)
\end{aligned}$$

Por lo tanto,

$$N'(d_2) = N'(d_1) \frac{S}{K} e^{r(T-t)} \quad (6.22)$$

Con (6.22), la derivada parcial quedaría como

$$\begin{aligned}
\delta &= \frac{\partial C}{\partial S} = SN'(d_1) \frac{\partial}{\partial S} d_1 + N(d_1) - Ke^{-r(T-t)} \left[N'(d_1) \frac{S}{K} e^{r(T-t)} \frac{\partial}{\partial S} d_1 \right] \\
&= SN'(d_1) \frac{\partial}{\partial S} d_1 + N(d_1) - SN'(d_1) \frac{\partial}{\partial S} d_1 \\
&= N(d_1)
\end{aligned}$$

Para una opción de venta P , recordando que $P = C - S(t) + Ke^{-r(T-t)}$ y aplicando un procedimiento similar al anterior, el valor de delta estaría dado como

$$\delta = \frac{\partial P}{\partial S} = \frac{\partial C}{\partial S} - 1$$

Rho, es la medida de la sensibilidad del precio de la opción con respecto a los cambios ocurridos en la tasa libre de riesgo. Su valor puede obtenerse

calculando la derivada parcial de C o P con respecto a r . Para una opción de compra, se tiene que

$$\begin{aligned}\rho &= \frac{\partial C}{\partial r} = S \frac{\partial}{\partial r} N(d_1) - K \left[e^{-r(T-t)} \frac{\partial}{\partial r} N(d_2) + N(d_2) (-(T-t)e^{-r(T-t)}) \right] \\ &= S \frac{\partial}{\partial r} N(d_1) - K \left[e^{-r(T-t)} N'(d_2) \frac{\partial}{\partial r} d_2 - N(d_2)(T-t)e^{-r(T-t)} \right] \\ &= S \frac{\partial}{\partial r} N(d_1) - K \left[e^{-r(T-t)} N'(d_2) \frac{\partial}{\partial r} (d_1 - \sigma\sqrt{T-t}) - N(d_2)(T-t)e^{-r(T-t)} \right]\end{aligned}$$

Usando la igualdad (6.22)

$$\begin{aligned}\rho &= S \frac{\partial}{\partial r} N(d_1) - K \left[e^{-r(T-t)} \left(N'(d_1) \frac{S}{K} e^{r(T-t)} \right) \frac{\partial}{\partial r} d_1 - N(d_2)(T-t)e^{-r(T-t)} \right] \\ &= S \frac{\partial}{\partial r} N(d_1) - SN'(d_1) \frac{\partial}{\partial r} d_1 + KN(d_2)(T-t)e^{-r(T-t)} \\ &= KN(d_2)(T-t)e^{-r(T-t)}\end{aligned}$$

Para una opción de venta, ρ está definida como

$$\rho = \frac{\partial P}{\partial r} = \frac{\partial C}{\partial r} - K(T-t)e^{-r(T-t)}$$

Prosiguiendo con las griegas, se tiene que Theta es la sensibilidad de la opción con respecto a los movimientos en el plazo de vencimiento, es decir, con respecto a $T-t$ y por consiguiente para obtener su valor se calcula la derivada parcial de C o P , dependiendo el tipo de opción, con respecto al período de tiempo. Llámese $\tau = T-t$, para una opción de compra, se tiene que

$$\begin{aligned}
\Theta &= \frac{\partial C}{\partial \tau} = S \frac{\partial}{\partial \tau} N(d_1) - K \left[e^{-r\tau} \frac{\partial}{\partial \tau} N(d_2) + N(d_2)(-re^{-r\tau}) \right] \\
&= SN'(d_1) \frac{\partial}{\partial \tau} d_1 - K \left[e^{-r\tau} N'(d_2) \left(\frac{\partial}{\partial \tau} d_1 - \frac{\partial}{\partial \tau} \sigma \sqrt{\tau} \right) - N(d_2) re^{-r\tau} \right] \\
&= SN'(d_1) \frac{\partial}{\partial \tau} d_1 - K \left[e^{-r\tau} N'(d_2) \left(\frac{\partial}{\partial \tau} d_1 - \frac{\sigma}{2\sqrt{\tau}} \right) - N(d_2) re^{-r\tau} \right] \\
&= SN'(d_1) \frac{\partial}{\partial \tau} d_1 - K \left[e^{-r\tau} \left(N'(d_1) \frac{S}{K} e^{r\tau} \right) \left(\frac{\partial}{\partial \tau} d_1 - \frac{\sigma}{2\sqrt{\tau}} \right) - N(d_2) re^{-r\tau} \right] \\
&= SN'(d_1) \frac{\partial}{\partial \tau} d_1 - SN'(d_1) \left(\frac{\partial}{\partial \tau} d_1 - \frac{\sigma}{2\sqrt{\tau}} \right) + KN(d_2) re^{-r\tau} \\
&= SN'(d_1) \frac{\sigma}{2\sqrt{\tau}} + KN(d_2) re^{-r\tau} \\
&= SN'(d_1) \frac{\sigma}{2\sqrt{T-t}} + KN(d_2) re^{-r(T-t)}
\end{aligned}$$

Para la opción de venta, el valor de Theta quedaría como:

$$\Theta = \frac{\partial P}{\partial \tau} = \frac{\partial C}{\partial \tau} - Kre^{-r\tau}$$

La última griega de la que se hablará, aunque realmente no es una letra griega ni se encuentra en el alfabeto griego, es conocida como Vega, la cual es la medida de la sensibilidad de la opción debido a las fluctuaciones de la volatilidad del activo subyacente. Para una opción de compra, el valor de Vega está dado como

$$\begin{aligned}
\frac{\partial C}{\partial \sigma} &= S \frac{\partial}{\partial \sigma} N(d_1) - Ke^{-r(T-t)} \left[\frac{\partial}{\partial \sigma} N(d_2) \right] \\
&= SN'(d_1) \frac{\partial}{\partial \sigma} d_1 - Ke^{-r(T-t)} \left[N'(d_2) \frac{\partial}{\partial \sigma} d_2 \right] \\
&= SN'(d_1) \frac{\partial}{\partial \sigma} d_1 - Ke^{-r(T-t)} \left[N'(d_1) \frac{S}{K} e^{r(T-t)} \frac{\partial}{\partial \sigma} (d_1 - \sigma \sqrt{T-t}) \right] \\
&= SN'(d_1) \frac{\partial}{\partial \sigma} d_1 - \left[N'(d_1) S \frac{\partial}{\partial \sigma} d_1 - N'(d_1) S \sqrt{T-t} \right] \\
&= SN'(d_1) \frac{\partial}{\partial \sigma} d_1 - N'(d_1) S \frac{\partial}{\partial \sigma} d_1 + N'(d_1) S \sqrt{T-t} \\
&= N'(d_1) S \sqrt{T-t}
\end{aligned}$$

Para el caso de una opción de venta, se tiene que

$$\frac{\partial C}{\partial \sigma} = \frac{\partial P}{\partial \sigma}$$

Para dar un ejemplo breve, retome los datos del Ejemplo 6.2, el valor de ρ para la opción de venta es -98.0353 . Lo que significaría que si el valor de la tasa libre de interés sube, por ejemplo, a 5% entonces el valor de la opción de venta podría llegar a ser $P = 34.2445 + (-98.0353 \times 2\%) = 32.2838$

Para más detalles sobre los procedimientos y demostraciones, se aconseja consultar [21].

Conclusiones

En la información que ha sido expuesta a lo largo de esta tesis se ha podido apreciar la utilidad de la probabilidad en diversas áreas de estudio. Se presentó el desarrollo de problemáticas que conciernen a economía, como lo son las compañías de seguros y cómo por medio del uso de la esperanza condicional se puede conocer la posibilidad de que ésta llegue a estar en la quiebra. Además, se explicó cómo la probabilidad influye en la valuación de las opciones financieras, las cuáles pueden evitar pérdidas para algunos comerciantes, generadas por los cambios en las tasas de interés, la volatilidad, entre otros aspectos más, considerados en los mercados financieros.

Se estimó la probabilidad de que a una cierta hora, en un lugar específico, la radiación solar sea mayor a un valor establecido analizando los registros que se tenían durante días pasados, esto por medio de la aproximación de funciones de densidad usando histogramas de frecuencia, mostrando así el uso de la probabilidad junto con la estadística.

Además, se expuso una de sus aplicaciones en las tecnologías actuales, explicando cómo se basaron los creadores de *Google* en las cadenas de Markov y la *Teoría de Grafos* para asignarle un valor de importancia, *PageRank*, a cada resultado obtenido después de una búsqueda y así poder ordenarlos, de forma que el usuario pueda conseguir la información más relevante del tema, de otra manera las búsquedas en la Web resultarían inservibles. También se habló sobre el algoritmo que emplea el *Filtro de Kalman* y cómo se utiliza para estimar parámetros.

Por mucho tiempo se ha subestimado la importancia del estudio de la probabilidad, limitando su uso a los juegos de azar, con todo lo visto anteriormente, se puede enfatizar que la relevancia de esta rama de las matemáticas radica en la capacidad de estimar resultados de sucesos que poseen incertidumbre.

Cabe resaltar que, en el desarrollo de esta tesis, se han utilizado conceptos

de la teoría de probabilidad que se estudian en una maestría en matemática, sin embargo, ésta cuenta con otras herramientas más, las cuales permiten analizar problemas que no son considerados en el cálculo estocástico clásico en el sentido de Itô, tales como el cálculo de Malliavin, que permite analizar problemas que no se encuentran adaptados a la información conocida o el estudio de aplicaciones de la teoría de la probabilidad utilizando diversas interpretaciones de la integral estocástica o incluso los procesos de Lévy, los cuales permiten considerar precios con saltos. Tampoco se ha tratado uno de los temas más estudiados acualmente, el movimiento browniano fraccionario, que se utiliza en mercados financieros para dar modelos más certeros que incluyan la volatilidad.

El objetivo de hablar de estos temas, tanto los estudiados como los que no se han expuesto en este trabajo, es mostrar lo importante que es la teoría de la probabilidad en el conocimiento humano. No hay duda de que, con el paso del tiempo, se seguirán descubriendo más aplicaciones para esta materia en diversos temas más.

Apéndice A

Ecuaciones en diferencias.

Para una función $y(t)$, el cambio en ella ocasionado por un incremento h en su argumento t se conoce como *primera diferencia* [37].

Una *ecuación en diferencias* es una expresión la cual relaciona los valores de una función $y(t)$ con sus respectivas diferencias finitas y variables independientes. Es decir,

$$F(y(t+n), y(t+n-1), y(t+n-2), \dots, y(t+1), y(t), t) = 0 \quad \forall t \in \mathbb{Z} \quad (\text{A.1})$$

Toda sucesión y que satisface una ecuación de la forma de (A.1) es una solución de la misma. A la familia de todas las soluciones se le conoce como *solución general*, la cual presenta ciertos parámetros que pueden determinarse a partir de las condiciones iniciales dadas, si éstos son calculados entonces se tiene lo que se conoce como *solución particular*. La solución general tiene tantas constantes arbitrarias como sea el orden de la ecuación [37].

Una ecuación en diferencias se dice lineal si puede expresarse como:

$$c_0y(t+n) + c_1y(t+n-1) + \dots + c_ny(t) = g(t) \quad (\text{A.2})$$

Además, se dice *homogénea* cuando $g = 0$, en caso contrario se conoce como *no homogénea*.

El orden de una ecuación en diferencias se obtiene a partir de la diferencia dada entre el mayor y el menor argumento de y que se encuentra en la ecuación.

En este apartado únicamente se expondrá la solución de ecuaciones homogéneas de coeficientes constantes, para más detalles y casos particulares se pueden consultar [26] y [37].

A.1. Solución de una ecuación en diferencias

La ecuación lineal en diferencias puede representarse como,

$$c_0 E^n y(t) + c_1 E^{n-1} y(t) + \dots + c_n y(t) = g(t)$$

donde $Ey(t) = y(t+1)$, se conoce como el *operador corrimiento* E . De forma que,

$$\begin{aligned} Ey(t) &= y(t+1) \\ E^2 y(t) &= E[Ey(t)] = y(t+2) \\ E^3 y(t) &= E[E^2 y(t)] = y(t+3) \\ &\vdots \\ E^m y(t) &= E[E^{m-1} y(t)] = y(t+m) \end{aligned}$$

Se representará por $\phi(E)$ al polinomio de E , es decir

$$\phi(E) = c_0 E^n + c_1 E^{n-1} + \dots + c_{n-1} E + c_n \quad (\text{A.3})$$

Usando (A.3), la ecuación (A.2) puede reescribirse como,

$$\phi(E)y(t) = g(t) \quad (\text{A.4})$$

Por ejemplo, si se tiene la ecuación $y(t+4) + 3y(t+2) - y(t+1) - 2y(t)$, ésta puede representarse con el operador E como,

$$\begin{aligned} y(t+4) + 3y(t+2) - y(t+1) - 2y(t) &= E^4 y(t) + 3E^2 y(t) - Ey(t) - 2y(t) \\ &= (E^4 + 3E^2 - E - 2)y(t) \end{aligned}$$

Dada una ecuación en diferencias lineal y homogénea con coeficientes constantes de orden n , es decir,

$$\phi(E)y(t) = 0 \quad (\text{A.5})$$

Se tienen los siguientes teoremas que presentan las propuesta de solución dependiendo de las formas que posean las raíces de la ecuación característica de (A.5).

La *ecuación característica* es un polinomio de grado n del que depende la solución de la ecuación en diferencias de orden n y la cual sólo puede encontrarse cuando la ecuación es de coeficientes constantes, lineal y homogénea.

Al proponer como solución de (A.5) la ecuación $y(t) = \alpha^t$, donde α es una constante, entonces esta función debe satisfacer a (A.5). De tal forma que,

$$(c_0\alpha^n + c_1\alpha^{n-1} + \dots + c_{n-1}\alpha + c_n)\alpha^t = 0 \quad (\text{A.6})$$

donde,

$$y(t+n) = \alpha^{t+n} = \alpha^n \alpha^t$$

(A.6) es la ecuación característica.

La información anterior y los teoremas siguientes pueden encontrarse con más detalle y demostraciones en [37]. Cabe aclarar que, en este apartado no se expondrá lo relacionado al caso de raíces complejas.

Teorema A.1 *Cuando la ecuación característica de la ecuación lineal en diferencias de coeficientes constantes y de orden n , $\phi(E)y(t) = 0$ tiene n raíces reales y diferentes denominadas como r_1, r_2, \dots, r_n , entonces la solución general de la ecuación en diferencias se propone de la forma:*

$$y(t) = c_1 r_1^t + c_2 r_2^t + \dots + c_n r_n^t \quad (\text{A.7})$$

Teorema A.2 *Para la ecuación característica de la ecuación en diferencias lineal, de coeficientes constantes de orden n , $\phi(E)y(t) = 0$ que tiene n raíces iguales denominadas como $r_1 = r_2 = \dots = r_n = r$, entonces la solución general se propone de la forma:*

$$y(t) = (c_1 + c_2 t + c_3 t^2 + \dots + c_n t^{n-1}) r^t \quad (\text{A.8})$$

Ahora, se hablará sobre el caso donde $g \neq 0$. La solución de una ecuación lineal en diferencias no homogénea (A.4) tiene la forma

$$y(t) = y_g(t) + y_p(t) \quad (\text{A.9})$$

donde $y_g(t)$ se conoce como la solución complementaria y es la solución general de la ecuación homogénea asociada [26]

$$\phi(E)y(t) = 0 \quad (\text{A.10})$$

$y_p(t)$ es la solución particular, esta solución dependerá de la forma que tenga $g(t)$.

Para obtener la solución particular de una ecuación en diferencias con las características descritas anteriormente, se puede utilizar el método de coeficientes indeterminados. Cabe aclarar que éste no es el único método que existe pero será el único que se expondrá en este apartado, además de presentar los casos más usuales.

Dada una ecuación en diferencias lineal no homogénea $\phi(E)y(t) = g(t)$ donde g es solución de alguna ecuación en diferencias lineal y homogénea, con lo que g es combinación lineal de funciones [27], las cuales pueden ser de las formas propuestas en los siguientes casos y para los cuales se presenta la propuesta de solución correspondiente:

- Si g es constante y $r = 1$ no es una solución de la ecuación característica, entonces $y_p = c$, donde c es una constante.
- Si g es una constante y $r = 1$ es una raíz de la ecuación característica con multiplicidad n entonces la solución tiene la forma $y_p = ct^n$
- Si $g = r^t$ donde r no es una raíz de la ecuación característica, $y_p = cr^t$
- Si $g = r^t$ donde r es una raíz de la ecuación característica, $y_p = ct^n r^t$
- Para $g = \sin(\beta t)$ o $\cos(\beta t)$ donde $r = \cos(\beta t) + i \sin(\beta t)$ no es raíz de la ecuación característica entonces, $y_p = c_1 \cos(\beta t) + c_2 \sin(\beta t)$

Apéndice B

Volatilidad histórica

En un mercado financiero, la volatilidad es la velocidad con la que se mueve el precio de un activo subyacente [7]. Es decir, es una medida de la frecuencia con la que un precio cambia y usualmente representa riesgo e incertidumbre. En el caso de las opciones de compra y venta, entre más alto sea el valor de la volatilidad de un activo subyacente, mayor será el precio de la opción.

En los mercados de opciones se hace la suposición de que el activo subyacente está en un mercado eficiente, lo que implica que la distribución de sus precios se aproxima a una distribución normal [7]. De forma que la varianza o la desviación típica pueden ser utilizadas para medir el cambio en los precios del activo. Uno de los métodos para realizar ello es conocido como *volatilidad histórica*.

La volatilidad histórica, como su nombre lo indica, utiliza los datos pasados para predecir la volatilidad actual. Cabe resaltar que este método no puede utilizarse en situaciones que impliquen un inestabilidad económica alta.

Los pasos para calcular la volatilidad presentados a continuación pueden verse con más detalle en [7].

Para hacer el cálculo de la volatilidad histórica se parte de los precios registrados anteriormente, ya sea por día, semana, mes, etc. Con lo cual, el procedimiento sería:

- Calcular el cociente $P_n = \frac{\text{precio al tiempo } n}{\text{precio al tiempo } n - 1}$
- Calcular el logaritmo natural de cada uno de los cocientes anteriores.

- Obtener la media de los logaritmos naturales, es decir, se calcula el cociente

$$Media = \frac{\sum_{n=1}^{m-1} \ln(P_n)}{m - 1}$$

donde m es la cantidad de meses, semanas, días, etc. de los que se tienen registros.

- Encontrar la desviación de cada uno de los logaritmos con respecto la media de los mismos, es decir

$$Desviación = \ln(P_n) - Media$$

- Calcular el cuadrado de cada una de las desviaciones del paso anterior para posteriormente sumarlas, al valor obtenido lo llamaremos s
- Calcular

$$Volatilidad = \sqrt{\frac{s}{m - 1}} \times 100$$

Aplicando el método descrito anteriormente para calcular la volatilidad del precio del azúcar en 2019 obtenida de [49], utilizada en el Capítulo 6, se obtienen los valores mostrados en la Tabla B.1.

Po lo tanto,

$$Volatilidad\ mensual = \sqrt{\frac{0.0327}{11}} \times 100 = 5.45\%$$

Cálculo de Volatilidad					
Mes	Precio Mensual	P_n/P_{n-1}	\ln	Desviación	Desviación ²
12	742.37	0.9772	-0.0230	-0.0410	0.0016
11	759.69	0.9833	-0.01678	-0.03481	0.0012
10	772.55	1.0494	0.04823	0.0302	0.0009
9	736.17	0.9787	-0.0214	-0.03947	0.0015
8	752.13	0.9837	-0.0163	-0.0343	0.0011
7	764.53	0.9920	-0.0079	-0.0259	0.0006
6	770.63	0.9835	-0.0166	-0.0346	0.00119
5	783.53	1.0488	0.0477	0.0297	0.0008
4	747.01	1.0486	0.0475	0.0294	0.0008
3	712.34	0.9912	-0.0087	-0.0268	0.0007
2	718.62	1.1802	0.1657	0.1477	0.0218
1	608.85				
		Media \ln	0.01802	Suma	0.0327

Tabla B.1: Cálculo de volatilidad para el Ejemplo 6.1

Bibliografía

- [1] G.J. ALEXANDER, J.V. BAILEY Y W. F. SHARPE, *Fundamentals of Investments*, Third Edition, Pearson Education, (2010).
- [2] L.J.S. ALLEN, *An Introduction to Stochastic Processes with Applications to Biology*, CRC Press, (2010).
- [3] M.A. ALLER, J. ARIAS, J.I. ARIAS Y L. LORENTE, *Fisiopatología Quirúrgica*, Tébar,(1999).
- [4] A. ANDREWS Y M.S. GREWAL, *Kalman Filtering: Theory and Practice Using MATLAB*, New York: John Wiley and Sons **14**, (2001).
- [5] T.M. APOSTOL, *Análisis Matemático*, Segunda Edición, Reverté, España, (1976).
- [6] L. ARNOLD, *Stochastic Differential Equations: Theory and Applications*, Wiley-Interscience, (1973).
- [7] J. AZNAR, A.A. LÓPEZ PERALES, C.A. TEODOSIO Y J.L. VIVANCOS BONO, *Introduction to Stochastic Integration*, Editorial Universitat Politècnica de Valencia, (2018).
- [8] J.M. BARRIOLA Y M. DOTTA, *¿Cómo funciona Google?: el algoritmo pagerank, diagramas de grafos y cadenas de Markov*, Revista de Investigación en Modelos Matemáticos aplicados a la Gestión y la Economía **03**, (2019), 3-20.
- [9] G.P. BASHARIN, A.N. LANGVILLE Y V.A. NAUMOV, *The life and work of A.A. Markov*, Linear Algebra and its Applications **386**, (2004), 3-26.

- [10] G. R. BELLO, *Operaciones Bancarias en Venezuela: Teoría y Práctica*, Universidad Católica Andrés Bello, (2004).
- [11] M.L. BERENSON Y D.M LEVINE, *Estadística Básica en Administración: Conceptos y Aplicaciones*, Sexta Edición, Prentice-Hall Hispanoamericana, México, (1996).
- [12] L. BLANCO CASTAÑEDA , *Breve historia de la probabilidad y la estadística*, Aportaciones Matemáticas: Memorias **44**, (2011), 69-82.
- [13] L. BLANCO CASTAÑEDA Y J.A. LEÓN, *Algunos ejemplos de aplicaciones de la teoría de la probabilidad*, Aportaciones Matemáticas **44**, (2011), 86.
- [14] L. BLANCO CASTAÑEDA Y M. MUÑOZ DE ÖZAK, *Introducción a la Teoría Avanzada de la Probabilidad*, Colección Textos, Universidad Nacional de Colombia, Colombia, (2002).
- [15] B. BOLLOBAS, R.M. DUDLEY Y W. FULTON, *Real Analysis and Probability*, First Edition, Cambridge University Press, (2002).
- [16] A. BONDY Y U.S. MURTY, *Graph Theory*, Graduate Texts in Mathematics, Springer London, (2007).
- [17] M.M. BUTT, B. GE, L. JIANG, Z. LI Y H. ZHANG, *Adaptive Unscented Kalman Filter for Target Tracking with Unknown Time-Varying Noise Covariance*, Sensors **19**, (2019), 1371.
- [18] G.C. CANAVOS, *Probabilidad y Estadística: Aplicaciones y Métodos*, Cuarta Edición, McGraw-Hill,(1991).
- [19] A. CORREDOR VILLALBA, G.H. CUETO Y C. MARTÍNEZ ALDANA, *Estado del Arte de Las Finanzas*, Primera Edición, Universidad Santo Tomas, Colombia, (2007).
- [20] J.C. COX, S.A. ROSS Y M. RUBINSTEIN, *Option pricing: A simplified approach*, Journal of Financial Economics **07**, (1979), 229-263.
- [21] J.C COX Y M. RUBINSTEIN, *Options Markets*, Prentice-Hall, (1985).
- [22] M.J. EVANS Y J.S. ROSENTHAL, *Probabilidad y Estadística*, Primera Edición, Reverté, España, (2005).

- [23] I.V. EVSTIGNEEV, T. HENS Y K.R. SCHENK-HOPPÉ, *Mathematical Financial Economics: A Basic Introduction*, First Edition, Springer International Publishing, (2015).
- [24] W. FELLER, *An Introduction to Probability Theory And Its Applications Vol.1*, Second Edition, Wiley, (1968).
- [25] F. GAGO COUSO, M.L. GONZÁLEZ, C.V. MARTÍN, F. MARTÍN, G. PÉREZ VERA, Y A.M. RODRÍGUEZ, *Teoría de Grafos: Ejercicios y Problemas Resueltos*, Ediciones Paraninfo S.A., (2014).
- [26] J.A. GARCÍA, *Matemáticas Financieras con Ecuaciones de Diferencia Finita*, Pearson Education, (2000).
- [27] P. GARCÍA Y C. DE LA LANZA, *Ecuaciones Diferenciales y en Diferencias*, Primera Edición, Limusa, México, (1984).
- [28] GOBIERNO DE LA CDMX, *Índice de radiación ultravioleta*, www.aire.cdmx.gob.mx y www.sedema.cdmx.gob.mx, (2019).
- [29] J. GONZÁLEZ L. Y L. RESTREPO B., *La historia de la probabilidad*, Revista Colombiana de Ciencias Pecuarias **16**, (2003), 83-87.
- [30] L. GOROSTIZA, *La probabilidad en el siglo XX*, Miscelánea Matemática **33**, (2001), 62-92.
- [31] G.R. GRIMETT Y D. STIRZAKER, *Probability and Random Processes*, Third Edition, OUP Oxford, (2001).
- [32] J. JACOD Y P. PROTTER, *Probability Essentials*, Second Edition, Springer Berlin Heidelberg, (2012).
- [33] E. KORN Y R. KORN, *Option Pricing and Portfolio Optimization: Modern Methods of Financial Mathematics*, American Mathematical Society, (2001).
- [34] P.R. KRUGMAN Y R. WELLS, *Introducción a la Economía. Macroeconomía*, Reverté, (2007).
- [35] R. KUMAR P., A. KUMAR SINGH Y A.G. KWANG LENG, *Application of Markov Chain in the PageRank Algorithm*, (2011), 1-6.

- [36] H.H. KUO, *Introduction to Stochastic Integration*, Universitext, Springer New York, (2006).
- [37] E.C. DE LA LANZA, *Ecuaciones Diferenciales y en Diferencias*, Universidad Nacional Autónoma de México, Facultad de Ingeniería, (1988).
- [38] J.A. LEÓN, *Integración estocástica con respecto al movimiento browniano*, Matemáticas: Enseñanza Universitaria vol.XIV **2**, Escuela Regional de Matemáticas, Colombia, (2006), 71-109.
- [39] J. LI, *Markov chain interpretation of Google page rank*, Preimpreso, (2005).
- [40] J.L. MARÍN Y A.T. PONCE, *Manual de Mercados Financieros*, Paraninfo, España,(2004).
- [41] W. MENDENHALL III, R.L SCHEAFFER Y D.D. WACKERLY, *Estadística Matemática con Aplicaciones*, Sexta Edición, CENGAGE, (2008).
- [42] K. OGATA, *Ingeniería de Control Moderna*, Quinta Edición, Pearson Education, México, (2003).
- [43] B. ØKSENDAL, *Stochastic Differential Equations: An Introduction with Applications*, Springer Berlin Heidelberg, (2010).
- [44] W. PINZON Y J. VARGAS CALEÑO, *Implementación de Sistemas de Control a partir de Modelos Sistema-Dinámicos*, American Mathematical Society, (2013).
- [45] A. POZNYAK, *Advanced Mathematical Tools for Automatic Control Engineers, Volume 2: Stochastic Systems*, First Edition, Elsevier Science, (2009).
- [46] L. RINCÓN, *Construyendo la integral estocástica de Itô*, Aportaciones Matemáticas, Serie Comunicaciones **35**, (2005), 265-283.
- [47] J.M. ROCHA MARTÍNEZ, *Un Primer Curso de Integración de Lebesgue en R^n* , Primera edición.
- [48] A.F. SARASOLA, *Control de los Sistemas Continuos*, Textos Universitarios, Ediciones de la Universidad de Oviedo, (2007).

- [49] SISTEMA NACIONAL DE INFORMACIÓN E INTEGRACIÓN DE MERCADOS (SNIIM), *Promedio del Precio Frecuente Mensual en Centrales de Abasto*, [www.economia – sniim.gob.mx/AzucarAnualPorMes](http://www.economia-sniim.gob.mx/AzucarAnualPorMes), (2019).
- [50] M.S. SOLÉ, *Probabilitats*, Col·lecció UB Series, Publicacions i Edicions de la Universitat de Barcelona, España, (1999).
- [51] M.R. SPIEGEL , *Probabilidad y Estadística*, Primera Edición, McGraw-Hill,(1992).
- [52] D. WILLIAMS, *Probability with Martingales*, Cambridge University Press, (1991).