



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**EL PRINCIPIO DE MÁXIMA ENTROPÍA  
Y VALUACIÓN DE OPCIONES**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**A C T U A R I A**

**P R E S E N T A:**

**ESTEFANÍA MONDRAGÓN AGUIRRE**



**DIRECTOR DE TESIS:**

**Dr. SERGIO IVÁN LÓPEZ ORTEGA**

CIUDAD UNIVERSITARIA, CD. MX., 2017



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

1. Datos del alumno

Apellido paterno  
Apellido materno  
Nombre(s)  
Teléfono  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Carrera  
Número de cuenta

2. Dator del tutor

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

3. Datos del sinodal 1

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

4. Datos del sinodal 2

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

5. Datos del sinodal 3

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

6. Datos del sinodal 4

Grado  
Nombre(s)  
Apellido paterno  
Apellido materno

7. Datos del trabajo escrito.

Título  
Subtítulo  
Número de páginas  
Año

1. Datos del alumno

Mondragón  
Aguirre  
Estefania  
25 92 79 53  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Actuaría  
310291688

2. Dator del tutor

Dr  
Sergio Iván  
López  
Ortega

3. Datos del sinodal 1

Dra  
Claudia Orquídea  
López  
Soto

4. Datos del sinodal 2

Dra  
Ana  
Meda  
Guardiola

5. Datos del sinodal 3

Act  
Germán  
Valle  
Trujillo

6. Datos del sinodal 4

Dra  
Bibiana  
Obregón  
Quintana

7. Datos del trabajo escrito

El principio de máxima entropía y la valuación de opciones  
64 p  
2017

# Índice general

<b>Agradecimientos</b>	<b>3</b>
<b>Introducción</b>	<b>5</b>
<b>1. Entropía</b>	<b>7</b>
1.1. Introducción . . . . .	7
1.2. Definición y propiedades elementales . . . . .	10
1.3. Divergencia de Kullback-Leibler . . . . .	14
<b>2. Principio de máxima entropía</b>	<b>21</b>
<b>3. Supuestos principales</b>	<b>27</b>
3.1. Teorema de arbitraje . . . . .	27
3.1.1. Árboles binomiales . . . . .	30
3.2. Entropía en finanzas . . . . .	32
3.2.1. Arbitraje . . . . .	33
3.2.2. Eficiencia de mercado . . . . .	34
<b>4. Teoría de precio entrópico</b>	<b>37</b>
4.1. Elementos deterministas en la valuación de una opción . . . . .	38
4.2. Teoría clásica de valuación de opciones (Black Scholes-Merton) . . . . .	39
4.3. Teoría del precio entrópico . . . . .	41
4.3.1. Principio de máxima entropía en la valuación de opciones . . . . .	42
4.3.2. Fórmula de valuación de opciones entrópica . . . . .	45
<b>5. Simulación</b>	<b>49</b>
5.1. Aplicación de la fórmula de valuación de opciones entrópica . . . . .	49
5.2. Movimiento Browniano Geométrico . . . . .	50
5.3. Proceso Poisson Compuesto . . . . .	51
5.4. Proceso de Cox-Ingersoll-Ross . . . . .	53
<b>Conclusiones generales</b>	<b>57</b>
<b>Apéndice</b>	<b>59</b>
Función de Costos . . . . .	59
Movimiento Browniano . . . . .	60
Proceso Poisson Compuesto . . . . .	61
Proceso Cox-Ingersoll-Ross . . . . .	61
<b>Bibliografía</b>	<b>63</b>



# Agradecimientos

Mis agradecimientos sinceros a la máxima casa de estudios, la *Universidad Nacional Autónoma de México*, especialmente a la *Facultad de Ciencias*, por ser mi segunda casa, por ser lugar de vivencias, aprendizaje y conocimiento. Por permitirme conocer personas importantes para mi y sobre todo por forjar la persona que soy hoy en día.

A mis padres *Romelia* y *Alejandro*, a mi hermano *César* por ser parte de esta aventura que decidí emprender hace algún tiempo. Por su apoyo y por creer en mi siempre. Son las personas que le dan sentido a este logro. A mis abuelos, a mis tías *Montse*, *Moni* y *Mari* y a mi tío *Horacio* por ser mi ejemplo a seguir, por su apoyo y consejos a lo largo de todo este camino.

A mis amigos, ellos que fueron parte de cada paso, decisión, desvelo, lágrima y risa que viví durante la carrera. Sé que no es necesario escribir sus nombres, pues saben lo agradecida que estoy con cada uno de ellos por todo lo que me dieron, por dejarme aprender de ellos y por su apoyo incondicional.

A las doctoras *Ana Meda*, *Claudia López* y *Bibiana Obregón*, al actuario *Gérman Valle* por sus observaciones y comentarios a este trabajo que fueron de suma importancia para mi y para la conclusión satisfactoria de este proyecto. Así mismo, les doy las gracias por enseñarme que cuando uno ama lo que hace, el transmitir conocimiento es más fácil.

Para finalizar, gracias al doctor *Sergio Iván López Ortega*, quien dirigió esta tesis. Gracias, por su apoyo, ayuda, conocimiento, consejos y risas compartidos en la elaboración del mismo. Quiero que sepa, que sin usted este trabajo no hubiera salido adelante. Y sobre todo, gracias por ser ahora un gran amigo.

*¡MUCHAS GRACIAS!*



# Introducción

Las finanzas constituyen una disciplina cotidiana en el quehacer diario. Junto con la economía, proporciona una cuantificación y una interpretación de la realidad, que va desde la economía del hogar hasta la planificación a nivel país.

Un aspecto importante a contemplar en esta área, es la existencia de agentes (personas o instituciones) llamadas inversores, que al invertir en algún activo financiero, buscan hacer crecer su dinero. Para dar fluidez a tales inversiones, existen instrumentos financieros; un ejemplo de estos instrumentos son los conocidos productos derivados. Entre ellos, uno de las más importantes son las opciones, pues este es un contrato que nos permite fijar un precio sobre algún activo y adquirir el compromiso de comprar o vender tal activo a ese precio.

Gracias a la flexibilidad de este producto, en los últimos años se ha desarrollado continuamente teoría que permita valuarlo y ayudar al inversionista a saber si le conviene o no realizar un contrato de este tipo. Entre este desarrollo, se encuentra la fórmula de Black-Scholes-Merton, cuya funcionalidad es tal que, es probablemente, la forma de valuar por excelencia estos productos. Sin embargo, la búsqueda de nuevos métodos que permitan una mejor valuación y ayudar a la toma de decisiones, nos lleva a otros campos como la teoría de la información.

El objetivo de esta tesis, es justamente seguir el análisis sobre el uso de la entropía y la teoría de la información en la valuación de opciones desarrollado por el economista Less Gulko (véase [6]). Tras ese desarrollo, se buscará comparar ambas teorías para poder contestar la pregunta de que este método pueda reemplazar o al menos ser igual de satisfactorio que el desarrollado por Black-Scholes-Merton.

Para ello, en el primer y segundo capítulo, se presentan todas los conceptos previos relacionados a la entropía. En el capítulo dos, se muestra el principio de máxima entropía que es crucial en capítulos posteriores. En el tercer capítulo, veremos algunos conceptos conocidos de la teoría clásica de valuación de opciones como las oportunidades de arbitraje, y además, se mostrará la manera en que la teoría entrópica puede ser aplicada a estos mismos conceptos.

En el capítulo 4, veremos la aplicación de la teoría entrópica, buscando una fórmula aplicable, parecida a la dada por Black-Scholes-Merton, que nos permita valuar opciones europeas. Para finalizar, en el capítulo 5, por medio de simulaciones, se hará un análisis para contrastar la teoría clásica de valuación de opciones con la teoría entrópica de valuación de opciones.



# Capítulo 1

## Entropía

### 1.1. Introducción

A lo largo de la historia, el concepto de entropía ha estado presente en muchas áreas del conocimiento; abarca desde la filosofía, la física y hasta la llamada teoría de la información. Este concepto proviene de la termodinámica, siendo utilizada por primera vez por Rudolf Clausius (1850) y en 1877, Ludwig Boltzmann encontró la manera matemática de expresarla; dicho boceto se presentará a continuación.

Considérese un conjunto  $X$  finito y  $p$  una medida de probabilidad definida en  $X$ . En el modelo de Maxwell-Boltzmann,  $X$  es el conjunto de todos los posibles *niveles de energía* de un sistema de partículas, y  $p$  corresponde a un histograma específico de energías que describen algún sistema microscópico. Suponga que el conjunto de estados microscópicos está dado por  $X^n$ . Un estado microscópico  $w \in X^n$ , representa los niveles de energía de las partículas  $1, 2, \dots, n$ . Suponga, por el momento, que cada  $p(x)$ ,  $x \in X$ , es un múltiplo de  $1/n$ , es decir,  $p$  es un histograma con  $n$  ensayos; a este histograma se le llama estado macroscópico del sistema. Note que dos estados microscópicos pueden tener el mismo estado macroscópico. Por ejemplo, consideremos los siguientes estados microscópicos:

$$w_1 = (A, A, A, B, A, B, C)$$

$$w_2 = (B, A, A, B, C, A, A)$$

donde para ambos conjuntos, su estado macroscópico está dado por la figura (1.1):

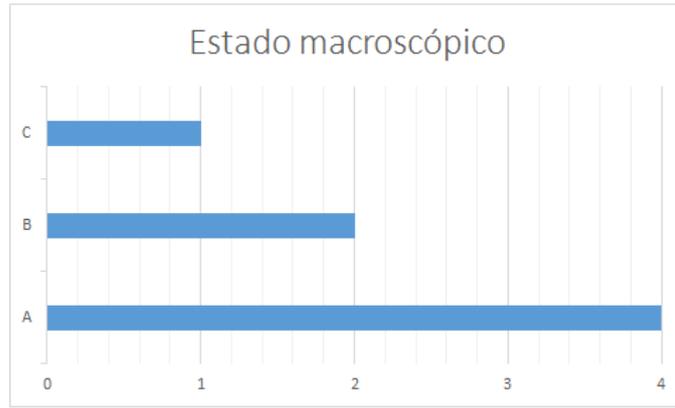


Figura 1.1: Histograma de estados macroscópicos.

La idea de Boltzmann es la siguiente:

*La entropía de un estado macroscópico  $p$  corresponde al grado de incertidumbre, y puede ser medida como  $\log(N_n(p))$ , donde  $N_n(p)$  es el número de estados microscópicos  $w$  que tienen asociado al macroscópico  $p$ .*

Explícitamente, para un estado microscópico  $w \in X^n$ , definimos

$$L_n^w = \frac{1}{n} \sum_{i=1}^n \delta_{w_i}^{(x)},$$

donde,

$$\delta_{w_i}^{(x)} = \begin{cases} 1 & \text{si } x=w_i \\ 0 & \text{si no} \end{cases}.$$

El estado macroscópico que describe cómo las partículas son distribuidas sobre los niveles de energía  $L_n^w$  es llamado la distribución empírica. Entonces

$$N_n(p) = \{w \in X^n : L_n^w = p\} = \frac{n!}{\prod_{x \in X} (np(x))!},$$

donde obtenemos el coeficiente multinomial.

Una aproximación para una medida de probabilidad general  $p$  se da por una sucesión de medidas de probabilidad  $p_n$ , donde  $p_n$  está asociada a un espacio con  $n$  partículas. Se define la incertidumbre  $H(p)$  del estado macroscópico  $p$  asociado a configuraciones infinitas por el límite de tales aproximaciones.

Más explícitamente, si definimos  $p$  y  $p_n$  medidas de probabilidad sobre  $X$  tales que  $p_n(x)$  converge a  $p(x)$  para toda  $x$  en  $X$  y  $np_n(x)$  es un número entero, entonces el límite

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log N_n(p_n).$$

existe y es igual a

$$-\sum_{x \in X} p(x) \log(p(x)).$$

Esto motivará después la definición formal de entropía para una variable aleatoria. Antes, considere el siguiente ejemplo,

**Ejemplo 1.1** Sean A, B y C tres tipos de automóviles en una agencia. Nos gustaría saber el orden en el cual fueron vendidos, sabiendo que en total se vendieron 6 autos en cada día. Contamos con las siguientes gráficas que proporcionan la cantidad de cada tipo de auto que se vendió en los últimos tres días:

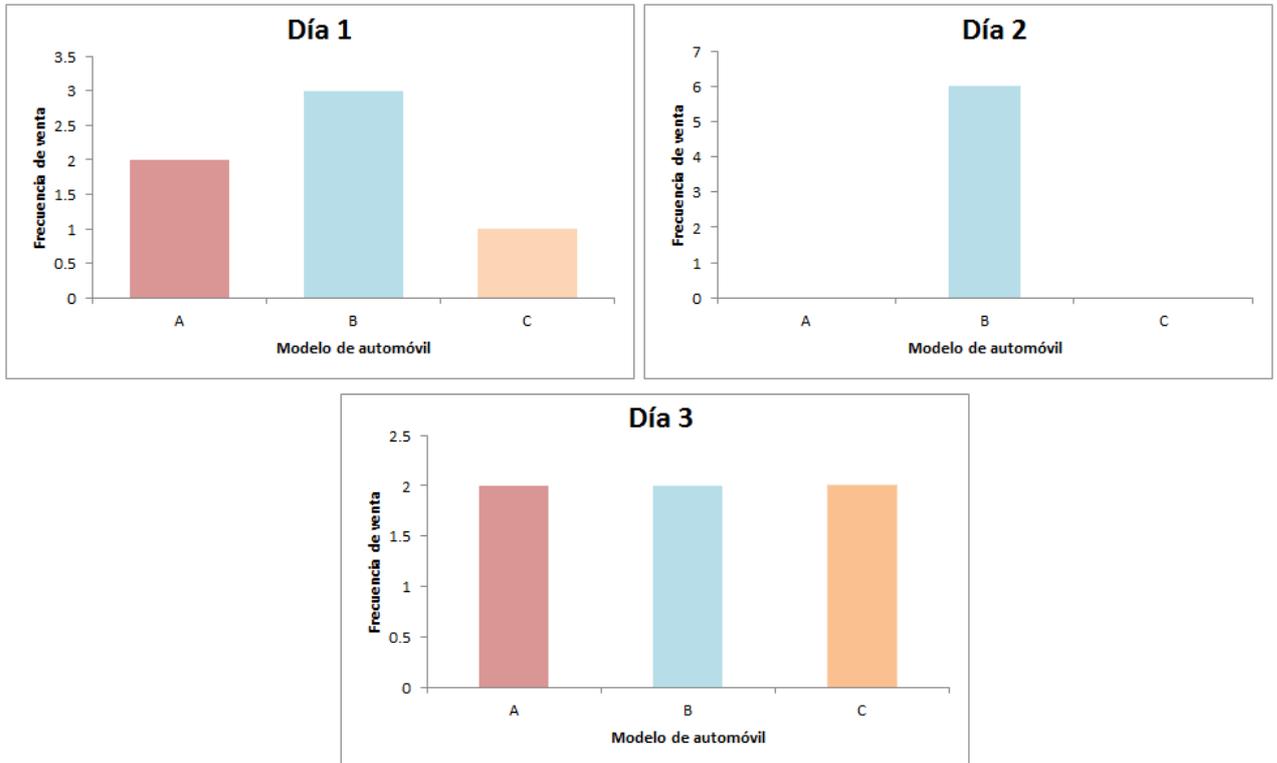


Figura 1.2: Histograma de venta diaria de automóviles.

Veamos que en la primera gráfica tenemos que el tipo de auto A se vendió 2 veces, el tipo B se vendió 3 y C una vez. En la segunda tenemos que el único auto que se vendió fue B, mientras que en la tercera cada auto se vendió la misma cantidad de veces. En el día uno, un orden posible es por ejemplo la combinación A-B-B-B-A-C o A-A-B-B-B-C. Mientras que en el día dos sólo tenemos un orden posible: B-B-B-B-B-B. Por su parte, en el día tres podemos tener B-B-A-A-C-C o B-C-A-A-C-B.

Notemos que el histograma asociado al día tres nos proporciona muy poca información sobre el orden de cómo se vendieron los autos, pues todos se vendieron en la misma cantidad. En el día uno tenemos mayor cantidad de información pues sabemos cuál se vendió más y podríamos predecir mejor el orden en que se vendieron. El histograma del día dos nos da absolutamente toda la información sobre el orden pues todos los autos vendidos fueron del tipo B, y entonces el único orden posible de venta en ese caso es B - B - B - B - B - B.

A la *falta de información* que se tiene sobre el orden de la venta de los autos se le llama *entropía*. Si consideramos al tipo de auto resultante en la venta de autos como una variable aleatoria, podríamos decir que la entropía es la *cantidad de incertidumbre* con la que cuenta dicha variable aleatoria.

## 1.2. Definición y propiedades elementales

Motivados por la sección anterior, podemos definir a la entropía para variables aleatorias discretas como:

**Definición 1.1** Sea  $X$  una variable aleatoria discreta con función de densidad  $p(x)$ . Definimos a la entropía de  $X$  como:

$$H(X) = - \sum_x p(x) \log(p(x)), \quad (1.1)$$

donde el logaritmo está en la base natural y se considera la convención  $0 \cdot \log(0) = 0$ . Para una variable aleatoria continua, la definición es análoga sustituyendo la suma por integral.

Note que  $H(X)$  no depende de los valores específicos que toma  $X$ , sino únicamente de las probabilidades asociadas a ellos. Una de las consecuencias inmediatas de la definición es:

**Proposición 1.1**  $H(x) \geq 0$ .

**Demostración:**

Como  $X$  es una variable aleatoria, sabemos que su función de densidad satisface  $0 \leq p(x) \leq 1$  y que  $\log(y) \leq 0$  para toda  $y$  en  $[0, 1]$ . Entonces

$$p(x) \log(p(x)) \leq 0 \quad \forall x,$$

por lo que

$$- \sum_x p(x) \log(p(x)) \geq 0.$$

Q.E.D.

**Ejemplo 1.2** Para el ejemplo (1.1), definimos a la variable aleatoria  $X_i$  con  $i = 1, 2, 3$  como el tipo de auto obtenido al elegir un auto al azar entre los vendidos el día  $i$ .

Para el día 1, tenemos que:

$$H(X_1) = -\frac{2}{6} \log\left(\frac{2}{6}\right) - \frac{3}{6} \left(\log\frac{3}{6}\right) - \frac{1}{6} \log\left(\frac{1}{6}\right) = 0,4392$$

Para el día 2:

$$H(X_2) = -\frac{6}{6} \log\left(\frac{6}{6}\right) = 0.$$

Para el día 3:

$$H(X_3) = -\frac{2}{6} \log\left(\frac{2}{6}\right) - \frac{2}{6} \log\left(\frac{2}{6}\right) - \frac{2}{6} \log\left(\frac{2}{6}\right) = 0,4771.$$

Tal como se escribió anteriormente, la variable aleatoria del día 3 es la que tiene mayor entropía, mientras que la entropía de la variable aleatoria del día dos es cero pues este experimento no tiene incertidumbre asociada.

**Ejemplo 1.3** Sea  $X$  una v.a con distribución Bernoulli( $p$ ), es decir, es 1 con probabilidad  $p$  y 0 con probabilidad  $(1-p)$ . Entonces su entropía es:

$$H(X) = -p \log(p) - (1-p) \log(1-p).$$

Notemos que en los casos  $p = 0$  o  $p = 1$  la variable aleatoria se convierte en determinista y  $H(X)$  es cero en estos casos, lo que corresponde a que no exista incertidumbre asociada a la variable. Además, nuestra incertidumbre será máxima cuando  $p = \frac{1}{2}$ .

**Ejemplo 1.4** Sea  $X$  una variable aleatoria con distribución uniforme en el intervalo  $[0, a]$ , con  $p(x) = \frac{1}{a} 1_{[0, a]}(x)$ . Entonces, podemos calcular su entropía:

$$\begin{aligned} H(X) &= - \int_{-\infty}^{\infty} p(x) \log(p(x)) dx \\ &= - \int_0^a \frac{1}{a} \log\left(\frac{1}{a}\right) dx \\ &= - \int_0^a \frac{1}{a} [\log(1) - \log(a)] dx \\ &= - \int_0^a \frac{1}{a} \log(1) dx + \int_0^a \frac{1}{a} \log(a) dx \\ &= \log(a). \end{aligned}$$

**Observación 1.2** La entropía de una variable aleatoria  $X$  puede ser interpretada como el valor esperado de una variable aleatoria  $Y = \log\left(\frac{1}{p(X)}\right)$  donde  $X$  tiene la función de densidad  $p(x)$ . Es decir

$$H(X) = E_p \left( \log\left(\frac{1}{p(x)}\right) \right) = E_p[Y]. \quad (1.2)$$

Es posible definir a la entropía conjunta de dos variables aleatorias:

**Definición 1.2** Sea  $X$  y  $Y$  dos variables aleatorias discretas con función de densidad conjunta  $p(x,y)$ . Definimos a la entropía conjunta por:

$$H(X, Y) = - \sum_x \sum_y p(x, y) \log(p(x, y)). \quad (1.3)$$

De manera similar a la Observación 1.2, tenemos que

$$H(X, Y) = -E[\log(p(X, Y))]. \quad (1.4)$$

**Definición 1.3** Sea  $X$  y  $Y$  variables aleatorias con respectiva función de densidad conjunta  $p(x,y)$ . Definimos a la entropía condicional de  $Y$  dado  $X$  por:

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

donde:

$$H(Y|X = x) = - \sum_y p_{Y|X}(y|x) \log p_{Y|X}(y|x).$$

Tal definición también coincide con una esperanza.

**Observación 1.3**

$$H(Y|X) = -E[\log(p(Y|X))]$$

**Demostración:**

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

Q.E.D.

A continuación presentamos una importante relación entre la entropía y la entropía condicional.

**Teorema 1.4 (Regla de la cadena)**

$$H(X, Y) = H(X) + H(Y|X) \quad (1.5)$$

**Demostración:**

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

Intuitivamente, la regla de la cadena nos dice que la información (o su carencia) conjunta de dos variables, puede verse como la información de una de ellas, más la información que representa la otra contemplando a la información que nos da la primera variable como conocida.

**Ejemplo 1.5** Sean  $X$  y  $Y$  variables aleatorias con función de densidad conjunta  $p(x,y)$  representada en la tabla siguiente.

Y/X	1	2	3	4
1	1/8	1/16	1/32	1/32
2	1/16	1/8	1/32	1/32
3	1/16	1/16	1/16	1/16
4	1/4	1	1	1

Notemos que la densidad marginal es  $p(x) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right)$  y  $p(y) = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ . Y por la Ecuación (1.1) tenemos que  $H(X) = \frac{7}{4}$  y  $H(Y) = 2$ , entonces:

$$\begin{aligned}
 H(X|Y) &= \sum_{i=1}^4 P(Y=i)H(X|Y=i) \\
 &= \frac{1}{4}H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right) + \frac{1}{4}H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, \frac{1}{8}\right) + \frac{1}{4}H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) + \frac{1}{4}H(1, 0, 0, 0) \\
 &= \frac{1}{4}\left(\frac{7}{4}\right) + \frac{1}{4}\left(\frac{7}{4}\right) + \frac{1}{4}(2) + \frac{1}{4}(0) \\
 &= \frac{11}{8}.
 \end{aligned}$$

Análogamente, podemos calcular  $H(Y|X) = \frac{13}{8}$  y usando la regla de la cadena, concluimos que  $H(X, Y) = \frac{27}{8}$ .

### 1.3. Divergencia de Kullback-Leibler

Hasta ahora, hemos mostrado que la entropía de una variable aleatoria se puede asociar a la cantidad de incertidumbre que ésta posee, es decir, es la cantidad de información faltante para poder describirla. Definiremos a continuación a la llamada entropía relativa, ésta se asocia a la ineficiencia de suponer que una variable aleatoria tiene una distribución  $q$  cuando en realidad su distribución es  $p$ . En este sentido, la divergencia nos da una noción de distancia (aunque no lo es en sentido estricto).

**Definición 1.4** La entropía relativa o distancia de Kullback-Leibler entre dos funciones de probabilidad  $p(x)$  y  $q(x)$  se define como

$$D(p||q) = \sum_x p(x) \log \left( \frac{p(x)}{q(x)} \right). \quad (1.6)$$

La divergencia es un concepto importante relacionado a la entropía. Mostraremos algunas propiedades de ella. Por lo que es necesario que recordemos algunos conceptos.

**Definición 1.5 Función convexa**

Se dice que una función  $f$  es convexa en un intervalo  $(a, b)$  si para todo  $x_1, x_2$  en  $(a, b)$  y  $\lambda$  en  $[0, 1]$  se cumple que

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

Decimos que es estrictamente convexa si se cumple lo anterior con la desigualdad estricta.

**Función cóncava**

Una función  $f$  es cóncava si  $-f$  es convexa.

**Teorema 1.5 (Desigualdad de Jensen)** Para  $f$  una función convexa y  $X$  una variable aleatoria se cumple que

$$E[f(X)] \geq f(E[X]).$$

Si se da la igualdad y  $f$  es estrictamente convexa, entonces  $X = E[X]$  con probabilidad 1.

La prueba de este teorema, se presenta en los cursos básicos de probabilidad; puede consultarse en [15], por ejemplo. Enunciemos ahora algunas propiedades importantes sobre la divergencia de Kullback-Lebner.

**Teorema 1.6 (Desigualdad de la información)** Ocurre que  $D(p||q) = 0$  si y sólo si  $p = q$ .

**Demostración:**

$\Rightarrow$ ) Sean  $p$  y  $q$  dos funciones de densidad y supongamos que  $D(p||q) = 0$ , entonces

$$D(p||q) = E \left( \log \frac{q(x)}{p(x)} \right) = 0.$$

Por otro lado, tenemos que

$$\begin{aligned} -\log \left( E \left( \frac{q(x)}{p(x)} \right) \right) &= -\log \sum_x p(x) \frac{q(x)}{p(x)} \\ &= -\log(1) = 0. \end{aligned}$$

De lo anterior concluimos que

$$E \left( -\log \left( \frac{q(X)}{p(X)} \right) \right) = -\log \left( E \left( \frac{q(X)}{p(X)} \right) \right)$$

y por la aclaración al final del Teorema 1.5 ocurre que

$$\begin{aligned} \frac{q(X)}{p(X)} &= E \left( \frac{q(X)}{p(X)} \right) \\ &= \sum_x p(x) \left( \frac{q(x)}{p(x)} \right) \\ &= 1, \end{aligned}$$

por lo que,

$$q(X) = p(X) \quad \forall x \in \mathbb{R}.$$

$\Leftrightarrow$ ) Supongamos ahora que  $p = q$ , es decir  $p(x) = q(x)$ ,  $\forall x \in \mathbb{R}$ . Entonces

$$D(p||q) = \sum_x p(x) \log \left( \frac{p(x)}{p(x)} \right) = 0.$$

Q.E.D.

**Teorema 1.7 (Desigualdad log-sum)** Sean  $a_1, \dots, a_n$  y  $b_1, \dots, b_n$  números no negativos. Entonces, se cumple

$$\sum_{i=1}^n a_i \log \frac{a_i}{b_i} \geq \left( \sum_{i=1}^n a_i \right) \log \left( \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \right) \quad (1.7)$$

Nuevamente, tenemos las convenciones  $0 \cdot \log(0) = 0$ ,  $a \cdot \log \left( \frac{a}{0} \right) = \infty$  si  $a \neq 0$ .

**Demostración:**

Sea  $X$  una v.a que toma valores en  $\chi = \left\{ x_i = \frac{a_i}{b_i} : i = 1, \dots, n \right\}$ . Como la función  $f(x) = x \log(x)$  es estrictamente convexa para  $x$  no negativa, entonces, por el Teorema 1.5, tenemos que

$$E[X \log(X)] \geq E[X] \log(E[X]).$$

En particular, en el caso en que  $X$  toma el valor  $\frac{a_i}{b_i}$  con probabilidad  $p_i$  dada por  $p_i = \frac{b_i}{\sum_j b_j}$ , obtenemos

$$\begin{aligned} \sum_i \left( \frac{b_i}{\sum_j b_j} \right) \frac{a_i}{b_i} \log \left( \frac{a_i}{b_i} \right) &\geq \left( \sum_i \left( \frac{b_i}{\sum_j b_j} \right) \frac{a_i}{b_i} \right) \log \left( \sum_i \left( \frac{b_i}{\sum_j b_j} \right) \frac{a_i}{b_i} \right) \\ &\Rightarrow \sum_i a_i \log \left( \frac{a_i}{b_i} \right) \geq \left( \sum_i a_i \right) \log \left( \frac{\sum_i a_i}{\sum_i b_i} \right). \end{aligned}$$

Q.E.D.

**Teorema 1.8 (Convexidad de la entropía relativa o divergencia)** El funcional  $D(p||q)$  es convexo en el par  $(p, q)$ , es decir, para  $0 \leq \lambda \leq 1$  y  $p$  y  $q$  medidas de probabilidad se cumple que:

$$D(\lambda p_1 + (1 - \lambda)p_2 || \lambda q_1 + (1 - \lambda)q_2) \leq \lambda D(p_1 || q_1) + (1 - \lambda)D(p_2 || q_2)$$

Además,  $D(p||q)$  es convexa en  $p$  para  $q$  fija y viceversa.

**Demostración:**

Recordando la Definición 1.4 tenemos que

$$D(\lambda p_1 + (1 - \lambda)p_2 || \lambda q_1 + (1 - \lambda)q_2) = - \sum_x (\lambda p_i(x) + (1 - \lambda)p_2(x)) \log \left( \frac{\lambda p_i(x) + (1 - \lambda)p_2(x)}{\lambda q_i(x) + (1 - \lambda)q_2(x)} \right).$$

definimos ahora a  $a_x = \lambda p_1(x) + (1 - \lambda)p_2(x)$  y  $b_x = \lambda q_1(x) + (1 - \lambda)q_2(x)$  con  $i = 1, 2$ . Entonces, utilizando la Desigualdad de log-sum y multiplicando por  $-1$ , obtenemos que

$$\begin{aligned} D(\lambda p_1 + (1 - \lambda)p_2 || \lambda q_1 + (1 - \lambda)q_2) &= \sum_x (\lambda p_1(x) + (1 - \lambda)p_2(x)) \log \left( \frac{\lambda p_1(x) + (1 - \lambda)p_2(x)}{\lambda q_1(x) + (1 - \lambda)q_2(x)} \right) \\ &\leq \lambda \sum_x p_1(x) \log \left( \frac{\lambda p_1(x)}{\lambda q_1(x)} \right) + (1 - \lambda) \sum_x p_2(x) \log \left( \frac{\lambda p_2(x)}{\lambda q_2(x)} \right) \\ &\leq \lambda D(p_1 || q_1) + (1 - \lambda) D(p_2 || q_2). \end{aligned}$$

Para la parte extra del teorema, note que para  $q$  fija tomamos  $q_1 = q_2 = q$  y obtenemos la concavidad en  $p$ . Análogamente para  $p$  fija, tomamos  $p_1 = p_2 = p$ .

Q.E.D.

**Observación 1.9** Denotemos por  $\mathbb{P}$  al espacio de medidas de probabilidad para un conjunto  $\Omega$  y una  $\sigma$ -álgebra  $\mathcal{F}$  fijos. Entonces  $\mathbb{P}$  es un conjunto convexo.

**Demostración:**

Sea  $\lambda$  en  $[0, 1]$  y  $p_1, p_2$  medidas de probabilidad en  $\mathbb{P}$ , probaremos que la función  $p(\cdot) := \lambda p_1(\cdot) + (1 - \lambda)p_2(\cdot)$  es una medida de probabilidad. En efecto,

- $p(\emptyset) = 0$

$$p(\emptyset) = \lambda p_1(\emptyset) + (1 - \lambda)p_2(\emptyset).$$

y como  $p_1$  y  $p_2$  son medidas de probabilidad

$$\begin{aligned} p(\emptyset) &= \lambda \cdot 0 + (1 - \lambda) \cdot 0 \\ &= 0. \end{aligned}$$

- $p(\Omega) = 1$

$$p(\Omega) = \lambda p_1(\Omega) + (1 - \lambda)p_2(\Omega),$$

pero  $p_1$  y  $p_2$  son medidas de probabilidad, por lo que

$$\begin{aligned} p(\Omega) &= \lambda \cdot 1 + (1 - \lambda) \cdot 1 \\ &= \lambda + 1 - \lambda \\ &= 1. \end{aligned}$$

- $p\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_i p\left(\bigcup_{i=1}^{\infty} A_i\right)$  con  $A_i$  conjuntos disjuntos 2 a 2, donde  $A_i \in \mathcal{F}$  para todo  $i = 1, 2, \dots, \infty$ .

$$p\left(\bigcup_{i=1}^{\infty} A_i\right) = \lambda_1 p_1\left(\bigcup_{i=1}^{\infty} A_i\right) + \lambda_2 p_2\left(\bigcup_{i=1}^{\infty} A_i\right);$$

como  $p_1$  y  $p_2$  son medidas de probabilidad, tenemos

$$p\left(\bigcup_{i=1}^{\infty} A_i\right) = \lambda_1 p_1\left(\sum_{i=1}^{\infty} A_i\right) + \lambda_2 p_2\left(\sum_{i=1}^{\infty} A_i\right),$$

y como las sumas son positivas y el límite de ellas existe, se tiene

$$\begin{aligned} p\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{i=1}^{\infty} \lambda_1 p_1(A_i) + \lambda_2 p_2(A_i) \\ &= \sum_{i=1}^{\infty} p(A_i). \end{aligned}$$

Q.E.D.

El siguiente teorema, es un resultado importante, ya que nos deja ver como el funcional  $H(p)$  es cóncavo dentro de un espacio convexo.

**Teorema 1.10 (Concavidad de la Entropía)** *El funcional  $H(p)$  es cóncavo, es decir, para cualquier par de medidas de probabilidad  $p_1$  y  $p_2$  ocurre que*

$$H(\lambda p_1 + (1 - \lambda)p_2) \geq \lambda H(p_1) + (1 - \lambda)H(p_2).$$

**Demostración:**

Denotemos a la densidad uniforme en  $(0, 1)$  por  $u(x)$ . Por el Teorema 1.8 sabemos que para  $q$  fija, ocurre que la divergencia es convexa por lo que, para  $p_1$  y  $p_2$  medidas de probabilidad ocurre

$$H(\lambda p_1 + (1 - \lambda)p_2) \geq \lambda H(p_1) + (1 - \lambda)H(p_2).$$

Por el resultado del Teorema 1.8, con  $u = q$ , concluimos que

$$\begin{aligned} D(\lambda p_1 + (1 - \lambda)p_2 || u) &\leq \lambda D(p_1 || u) + (1 - \lambda)D(p_2 || u) \\ \log(x) - H(\lambda p_1 + (1 - \lambda)p_2) &\leq \lambda(\log(x) - H(p_1)) + (1 - \lambda)(\log(x) - H(p_2)) \\ -H(\lambda p_1 + (1 - \lambda)p_2) &\leq -\lambda H(p_1) - (1 - \lambda)H(p_2). \end{aligned}$$

De esta manera, multiplicando por por -1, obtenemos el resultado.

Q.E.D.

Para lo siguiente es necesario definir una distancia en el espacio de medidas de  $\mathbb{P}$  para un conjunto  $\mathbf{P}$  y una  $\sigma$ -álgebra  $\mathcal{F}$ . Nos contentaremos con decir que tal construcción es posible en muchos casos y referimos al lector al libro de Billingsley [3] para ello, además de utilizar la notación  $B_\epsilon(u)$  como una vecindad en un bola abierta en el espacio  $P$ .

De acuerdo al teorema anterior, y la Observación 1.9 podemos, ahora enunciar un resultado sobre la unicidad de la densidad de probabilidades que para algún  $\epsilon > 0$ , maximizan al funcional de la entropía.

**Teorema 1.11** *Sea  $u$  una medida de probabilidades en  $\mathbb{P}$  tal que es un máximo local del funcional  $H$ , es decir para algún  $\epsilon > 0$ ,*

$$H(u) \geq H(v) \quad \forall v \in B_\epsilon(u) \cap \mathbb{P},$$

*Entonces  $u$  es un máximo global, es decir  $H(u) \geq H(v)$ ,  $\forall v \in \mathbb{P}$ .*

**Demostración:**

Supongamos que  $u$  no es un máximo global, entonces existe una medida de probabilidad  $p$  en  $\mathbb{P}$  tal que  $H(p) > H(u)$ . Pero recordemos que  $\mathbb{P}$  es convexo, entonces

$$(1 - \theta)u + \theta p \in P, \quad \forall \theta \in [0, 1].$$

Debido a que  $H$  es cóncava por el Teorema (1.5), ocurre que

$$\begin{aligned} H((1 - \theta)u + \theta p) &\geq (1 - \theta)H(u) + \theta H(p) \\ &> (1 - \theta)H(u) + \theta H(u) \\ &\geq H(u). \end{aligned}$$

Sea  $\theta_0$  lo suficientemente pequeño, entonces  $(1 - \theta_0)u + \theta_0 p$  está en  $B_\epsilon(u)$ , y ocurre que:

$$H((1 - \theta_0)u + \theta_0 p) > H(u)$$

Esto es una contradicción pues  $u$  es un máximo local en  $B_\epsilon(u)$ .

Q.E.D.

De esta manera, hemos llegado a establecer todas las principales propiedades de la entropía, donde destacará a lo largo de los capítulos posteriores que nos encontramos frente a un funcional cóncavo dentro de un espacio conexo. Esto permitirá desarrollar el principio de máxima entropía y más aún, asegurará que de existir el máximo entonces será único.

# Capítulo 2

## Principio de máxima entropía

El principio de máxima entropía fue primero formulado por Gibbs, quien argumentó que la entropía en termodinámica puede y tiende a incrementarse hasta el máximo permitido por cualesquiera que sean las restricciones del problema. Más tarde, Boltzmann dió una explicación completa del principio. A mediados del siglo pasado con los trabajos de Shannon, se interpretó a la entropía como una medida de *falta de información* o como el aumento de incertidumbre de un sistema sometido a diferentes cambios.

El principio de máxima entropía, establece lo siguiente: la distribución de probabilidad menos sesgada que se le puede atribuir a un sistema estadístico es aquella en la que, dadas ciertas condiciones o restricciones fijas, maximiza la entropía.

Con base en lo establecido por el principio, el objetivo al que nos enfrentaremos será el de generar distribuciones de probabilidad bajo la información dada (que se escribirá en términos de restricciones específicas). Es decir, entre todas las distribuciones de probabilidad que son factibles en un conjunto de estados finito bajo cierta información, se desea encontrar la que tenga mayor nivel de incertidumbre.

Definamos a una variable aleatoria continua  $X$  en  $\mathbf{P}$ , donde  $\mathbf{P}$  denota el espacio de funciones de densidad asociadas a un espacio medible  $(\Omega, \mathcal{F})$ . Nosotros deseamos estimar una función de densidad  $f^*$  de  $X$  que satisfaga el siguiente programa:

$$\begin{aligned} \text{Maximizar} \quad & H(p) = - \int_{-\infty}^{\infty} f(x) \log(f(x)) dx \\ \text{sujeto a} \quad & \int_{-\infty}^{\infty} f_i(x) r_i(x) dx = a_i \quad i = 1, 2, \dots, m \\ & \int_{-\infty}^{\infty} f_i(x) dx = 1 \\ & f_i > 0 \text{ en } \mathbf{P} \end{aligned}$$

donde  $r_i(x)$  es una función cualquiera dada, y  $X$  es una variable aleatoria con soporte fijo e integrable en  $\mathbb{R}$ .

Lo que da paso al siguiente teorema:

**Teorema 2.1 (Teorema de máxima entropía)** Sea  $f^*(x) = f_\lambda(x) = e^{\lambda_0 + \sum_i^m \lambda_i r_i(x)}$ , para  $x$  real donde  $\lambda_0, \dots, \lambda_m$  son escogidas de tal modo que  $f^*$  satisface las restricciones

del programa anterior. Entonces existe  $f^*$  que maximiza  $H(f)$  de forma única sobre todas las probabilidades de densidad  $f$  que satisfagan las restricciones del programa anterior.

### Demostración:

Note que para esta prueba, se usará la optimización de funcionales de cálculo en variaciones, ver por ejemplo el Capítulo 7, p. 169, de [11]. De este modo podemos escribir al funcional como:

$$J(f) = - \int f \log f + \lambda_0 \int f + \sum_{i=1}^m \lambda_i \int f r_i \text{ para } \lambda_0, \lambda_1, \dots, \lambda_m \in \mathbb{R}$$

derivando con respecto a  $f(x)$ , se tiene:

$$\frac{\partial J}{\partial f(x)} = - \log f(x) - 1 + \lambda_0 + \sum_{i=1}^m \lambda_i r_i(x)$$

igualando a cero:

$$\begin{aligned} - \log f(x) - 1 + \lambda_0 + \sum_{i=1}^m \lambda_i r_i(x) &= 0 \\ \log f(x) &= -1 + \lambda_0 + \sum_{i=1}^m \lambda_i r_i(x) \\ f(x) &= e^{-1 + \lambda_0 + \sum_{i=1}^m \lambda_i r_i(x)} \end{aligned}$$

Por lo que la  $f^*$  que buscamos será:

$$f^*(x) = e^{-1 + \lambda_0 + \sum_{i=1}^m \lambda_i r_i(x)} \quad (2.1)$$

Donde  $\lambda_i, i = 0, 1, \dots, m$  son elegidas tales que  $f^*$  satisface las restricciones del problema. La ecuación de Euler-Lagrange (Teorema 1.1, p. 68, de [9]) nos asegura que tal  $f^*$  es la que minimiza al funcional  $J(f)$ . Recordemos que  $H$  es una función cóncava sobre un conjunto convexo (ver Teorema (1.10)). Por lo tanto, el máximo local que encontramos con los teoremas anteriores es de hecho un máximo global.

Q.E.D.

**Observación 2.2** Para variables aleatorias discretas, tenemos un teorema análogo que garantiza la existencia única de una densidad discreta que maximiza la entropía sobre un espacio restringido.

Ahora, considerando el teorema anterior, en su versión discreta, desarrollemos algunos ejemplos.

**Ejemplo 2.1** Entre todas las distribuciones en  $\{1, 2, 3, \dots, n\}$  la que alcanza la máxima entropía es  $q^* = (\frac{1}{n}, \dots, \frac{1}{n})$ .

El Teorema (2.1) esta enunciado para variables aleatorias continuas, sin embargo, utilizaremos su versión para densidades discretas. Escribimos,

$$\begin{aligned} p(x) &= \frac{1}{n} 1_{\{1,2,\dots,n\}}^{(x)} \\ &= e^{\log\left(\frac{1}{n} 1_{\{1,2,\dots,n\}}^{(x)}\right)} \\ &= e^{\log\left(\frac{1}{n}\right) + \log\left(1_{\{1,2,\dots,n\}}^{(x)}\right)}. \end{aligned}$$

Por el teorema de máxima entropía concluimos que:

$$p^* = e^{-1+\lambda_0}.$$

Dado que la única restricción que se tiene es  $\sum_{i=1}^{\infty} p(x) = 1$  y tener soporte en  $\{1, 2, \dots, n\}$ , entonces

$$\begin{aligned} \sum_x p^* &= \sum_x e^{-1+\lambda_0} 1_{\{1,2,\dots,n\}}^{(x)} \\ &= 1. \end{aligned}$$

Por lo que  $e^{-1+\lambda_0} = \frac{1}{n}$ .

**Ejemplo 2.2** Suponga que  $n$  dados son lanzados sobre una tabla y tomamos el número total de puntos como  $n\alpha$ . ¿ En qué proporción la cara  $i=1,2,3,\dots,6$  es mostrada?

Una manera de saber esto es contando el número de formas en que los  $n$  dados pueden caer, tal que el dado  $n_i$  mostró la cara  $i$ ; esto consiste en  $\binom{n}{n_1, n_2, \dots, n_6}$  posibles maneras. Definamos a un estado macroscópico por el vector  $(n_1, \dots, n_6)$  correspondiente a  $\binom{n}{n_1, n_2, \dots, n_6}$  estados microscópicos (las salidas exactas en el dado) donde cada estado microscópico tiene probabilidad  $\frac{1}{6^n}$ . Para encontrar el estado macroscópico más probable, desearíamos maximizar  $\binom{n}{n_1, n_2, \dots, n_6}$  sobre el total de estados macroscópicos asociados que cumplen que

$$\sum_{i=1}^6 i n_i = n\alpha. \quad (2.2)$$

Usando la aproximación de Stirling,  $n! \simeq \left(\frac{n}{e}\right)^n \sqrt{n} = \frac{n^{n+\frac{1}{2}}}{e^n}$ , entonces

$$\begin{aligned} \binom{n}{n_1, n_2, \dots, n_6} &\simeq \left(\frac{n}{e}\right)^n \sqrt{n} = \frac{n^{n+\frac{1}{2}}}{e^n} \\ &= \prod_{i=1}^6 \left(\frac{n}{n_i}\right)^{n_i+\frac{1}{2}} \\ &= e^{nH\left(\frac{n_1}{n}, \dots, \frac{n_6}{n}\right)} \end{aligned}$$

Esto maximiza  $\binom{n}{n_1, n_2, \dots, n_6}$  sobre la restricción (2.2) lo cual es equivalente a maximizar  $H(p_1, \dots, p_6)$  sobre la restricción  $\sum_i i p_i = \alpha$ . De acuerdo al Teorema (2.1), encontramos que la máxima entropía está dada por la función de masa:

$$p_i^* = \frac{e^{\lambda_i}}{\sum_{i=1}^6 e^{\lambda_i}},$$

donde  $\lambda_i$  es escogida de tal forma que  $\sum_i i p_i^* = n$ . El estado más probable es  $(np_1^*, \dots, np_6^*)$  y esperamos que el dado muestre la cara  $i$   $n_i^* = np_i^*$  veces.

**Ejemplo 2.3** Considere las densidades en  $\{1, 2, 3, \dots\}$  con media  $\mu > 1$  fija. Entonces  $p_k = \left(\frac{1}{\mu}\right) \left(1 - \frac{1}{\mu}\right)^{k-1}$ , para  $k = 1, 2, 3, \dots$  es la de máxima entropía.

Del mismo modo que en el ejemplo anterior, escribamos a  $p(x)$  como

$$\begin{aligned} p(x) &= \left(\frac{1}{\mu}\right) \left(1 - \frac{1}{\mu}\right)^{x-1} 1_{\{1,2,\dots\}}^{(x)} \\ &= e^{\log\left(\left(\frac{1}{\mu}\right)\left(1-\frac{1}{\mu}\right)^{x-1} 1_{\{1,2,\dots\}}^{(x)}\right)} \\ &= e^{(x-1)\log\left(1-\frac{1}{\mu}\right) + \log\left(\frac{1}{\mu}\right) + \log\left(1_{\{1,2,\dots\}}^{(x)}\right)} \\ &= e^{x\log\left(1-\frac{1}{\mu}\right)} e^{\log\left(\frac{1}{\mu}\right) - \log\left(1-\frac{1}{\mu}\right) + \log\left(1_{\{1,2,\dots\}}^{(x)}\right)}. \end{aligned}$$

Según las siguientes restricciones

$$\begin{aligned} \sum_x p(x) &= 1, \\ \sum_x xp(x) &= \mu, \end{aligned}$$

y por el teorema de máxima entropía concluimos que

$$e^{-1+\lambda_0+\lambda_1x} = e^{x\log\left(1-\frac{1}{\mu}\right)} e^{\log\left(\frac{1}{\mu}\right) - \log\left(1-\frac{1}{\mu}\right) + \log\left(1_{\{1,2,\dots\}}^{(x)}\right)},$$

donde

$$\begin{aligned} \lambda_0 &= \log\left(\frac{1}{\mu}\right) - \log\left(1 - \frac{1}{\mu}\right), \\ \lambda_1 &= \log\left(1 - \frac{1}{\mu}\right). \end{aligned}$$

**Ejemplo 2.4** Consideremos al conjunto de todas las densidades en  $[0, \infty)$  con media fija  $\mu > 0$ , entonces  $p(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}$  tiene la máxima entropía.

Para este caso, contamos con una restricción similar al ejemplo anterior. Note que trabajaremos con densidades continuas, por lo que aplicaremos el teorema de máxima entropía en versión continua. Entonces por (2.1),

$$e^{\lambda_1 x} = e^{-\frac{x}{\mu}} \Rightarrow \lambda_1 = -\frac{1}{\mu},$$

y

$$e^{-1} = \mu,$$

por lo que  $\frac{1}{\mu} e^{-\frac{x}{\mu}}$  es la que maximiza a la entropía.

**Ejemplo 2.5** Definimos  $S = (-\infty, \infty)$ ,  $E(X) = \alpha_1$  y  $E(X^2) = \alpha_2$ . La distribución de máxima entropía es una densidad normal  $N(\alpha_1, \alpha_2 - \alpha_1^2)$ .

Recordemos, la función de distribución de una normal está dada por

$$f(x) = \frac{1}{\sqrt{2\pi\alpha_2 - \alpha_1^2}} e^{-\frac{(x-\alpha_1)^2}{2\alpha_2 - \alpha_1^2}},$$

mientras que  $f^*(x)$  tiene la forma

$$f^*(x) = e^{\lambda_0 + \lambda_1 x + \lambda_2 x^2}$$

Entonces, igualando exponentes, tenemos la solución explícita:

$$\begin{aligned}\lambda_0 &= \frac{\alpha_1^2}{2\alpha_2 - \alpha_1^2}, \\ \lambda_1 &= \frac{-\alpha_1}{\alpha_2 - \alpha_1^2}, \\ \lambda_2 &= \frac{-1}{2\alpha_2 - \alpha_1^2}.\end{aligned}$$

En los capítulos posteriores, aplicaremos el programa de máxima entropía sujeto a restricciones dadas por el ámbito financiero. Obtendremos densidades  $f^*$  tales que maximizan la incertidumbre dada la información disponible en el mercado para un activo financiero, en un contexto simple.



# Capítulo 3

## Supuestos principales

En este capítulo discutiremos los supuestos financieros que serán importantes en los análisis posteriores. Es decir, examinaremos los conceptos relacionados con el quehacer financiero; como por ejemplo el concepto de arbitraje. Introduciremos ideas de teoría económica que permitirán más adelante contextualizar la teoría entrópica del precio de una opción. Además, buscamos enlazar a la teoría financiera con los conceptos de entropía.

### 3.1. Teorema de arbitraje

A través de los años, la importancia de tener alguna posibilidad de ganar en una operación financiera dada, es una de las cosas que más preocupa a las personas cuya actividad principal es invertir dinero. De esta manera, se creó el término *arbitraje* que es una palabra francesa que denota una decisión de arbitraje o tribunal y que apareció por primera vez con un sentido financiero en 1704 en *La science des négocians et teneurs de livres*, un tratado escrito por Mathieu de la Porte. Con este término hacía referencia a las consideraciones de los diferentes tipos de cambio, para reconocer los lugares más rentables de emisión de letras de cambio.

La definición formal de arbitraje es la siguiente:

**Definición 3.1** *Se le llama arbitraje a la posibilidad de realizar una operación financiera con probabilidad positiva de ganancia y en donde con probabilidad cero no perderemos.*

Para ilustrar lo anterior, considere el conjunto  $S = \{1, 2, \dots, m\}$  que tiene a los posibles resultados que podría tener un experimento, y suponga que existen  $n$  posibles apuestas disponibles. Sea  $x$  el monto apostado en la estrategia de apuesta  $i$ . Entonces  $xr_i(j)$  será el dinero recibido si el resultado de nuestro experimento es  $j$  en el espacio de estados  $S$ . De esta manera  $r_i(\cdot)$  es la función de rendimiento por unidad apostada en la estrategia  $i$ . Dicho monto  $x$  podría ser negativo, cero o positivo, esto porque se puede perder lo apostado, ganar o simplemente no ganar ni perder en la apuesta.

Sea  $x = (x_1, x_2, \dots, x_n)$  un vector de estrategias de apuestas, donde  $x_i$  es el monto apostado en la estrategia  $i$ . Si tenemos que el resultado del experimento es  $j$ , entonces la ganancia del vector de estrategias  $\mathbf{x}$  será

$$x = \sum_{i=1}^n x_i r_i(j).$$

Así, existe un vector de probabilidades  $p = (p_1, \dots, p_m)$  que está relacionado con el espacio de estados  $S$  de nuestro experimento, el cual nos dirá la ganancia esperada de cada una de las apuestas que podría ser igual a cero o darnos la posibilidad de tener una estrategia de apuesta que tenga una probabilidad positiva de ganar en cada resultado de nuestro experimento. Relacionado a lo anterior postulamos uno de los teoremas fundamentales en la teoría de arbitraje.

**Teorema 3.1 (Teorema del arbitraje)** *Exactamente una de las siguientes afirmaciones es verdadera:*

1. Existe un vector de probabilidades  $p = (p_1, p_2, \dots, p_m)$  que cumple

$$\sum_{j=1}^m p_j r_i(j) = 0 \quad \text{para toda } i = 1, \dots, n$$

2. Existe una estrategia de apuesta  $x = (x_1, x_2, \dots, x_n)$  para la cual se cumple que

$$\sum_{i=1}^n x_i r_i(j) > 0 \quad \text{para toda } j = 1, \dots, m$$

Antes de presentar la demostración del teorema anterior, es necesario recordar algunos conceptos de programación lineal.

Supongamos que, para constantes dadas  $c_i$ ,  $b_j$  y  $a_{ij}$  con  $i = 1, \dots, m$  y  $j = 1, \dots, n$ , se buscan valores  $x_1, \dots, x_n$  que solucionen el siguiente programa, conocido como *programa primal*.

$$\begin{aligned} \text{Maximizar} \quad & \sum_{i=1}^n c_i x_i, \\ \text{sujeto a} \quad & \sum_{i=1}^n a_{ij} x_i \leq b_j, \quad j = 1, 2, \dots, m. \\ & x_i \geq 0. \end{aligned}$$

Además, cada primal tiene un *programa dual* relacionado. Para poder solucionar el dual, se buscan valores  $y_1, \dots, y_m$  que solucionen

$$\begin{aligned} \text{Minimizar} \quad & \sum_{j=1}^m b_j y_j, \\ \text{sujeto a} \quad & \sum_{j=1}^m a_{i,j} y_j \geq c_i, \quad i = 1, \dots, n, \\ & y_j \geq 0, \quad j = 1, \dots, m. \end{aligned}$$

Ahora bien, se dice que un programa de programación lineal es *factible* si existen variables  $x_1, \dots, x_n$  y  $y_1, \dots, y_m$  ( $x_i$  en el primal y  $y_j$  para el dual) que satisfagan ambos problemas. Tenemos el siguiente teorema clásico de programación lineal. La prueba puede consultarse en [2].

**Proposición 3.2 (Teorema de dualidad fuerte)** *Sea  $P$  un problema primal de minimización y  $D$  el problema dual asociado. Si  $P(D)$  tiene solución óptima entonces  $D(P)$  tendrá solución óptima y el valor que minimiza  $P$  es el mismo que máxima  $D$ .*

Al teorema de arbitraje presentado previamente, y con las nuevas herramientas se dará una prueba formal del teorema.

### Demostración del Teorema de arbitraje:

Denotemos por  $x_{n+1}$  al monto seguro que puede ganar un jugador, se busca entonces maximizar dicho monto. Si el jugador usa la estrategia  $(x_1, \dots, x_n)$  entonces ganará  $\sum_{i=1}^n x_i r_i(j)$  en el caso de que el resultado del experimento sea  $j$ . Por lo que, el jugador buscará elegir su estrategia y maximizar  $x_{n+1}$ , es decir:

$$\begin{array}{ll} \text{maximizar} & x_{n+1} \\ \text{sujeto a} & \sum_{i=1}^n x_i r_i(j) \geq x_{n+1}, \quad j = 1, \dots, m. \end{array}$$

Definamos a  $a_{i,j} = -r_i(j)$ ,  $i = 1, \dots, n$  y  $a_{n+1,j} = 1$ . Entonces, se puede reescribir el problema anterior como:

$$\begin{array}{ll} \text{maximizar} & x_{n+1} \\ \text{sujeto a} & \sum_{i=1}^{n+1} a_{i,j} x_i \leq 0, \quad j = 1, \dots, m. \end{array}$$

Note que en el problema anterior las variables valen  $c_1 = c_2 = \dots = c_n = 0$  mientras que  $c_{n+1} = 1$  y  $b_j = 0$ . Así, el programa dual será encontrar variables  $(y_1, \dots, y_m)$  tales que satisfagan

$$\begin{array}{ll} \text{minimizar} & 0 \\ \text{sujeto a} & \sum_j a_{i,j} y_j = 0, \quad i = 1, \dots, n. \\ & \sum_{j=1}^m a_{n+1,j} y_j = 1, \\ & y_j \geq 0, \quad j = 1, \dots, m, \end{array}$$

donde  $a_{ij}$  cumple la definición de  $a_{i,j}$  dada. El programa anterior puede reescribirse como

$$\begin{array}{ll} \text{minimizar} & 0 \\ \text{sujeto a} & \sum_{j=1}^m r_i(j) y_j = 0, \quad i = 1, \dots, n. \\ & \sum_{j=1}^m y_j = 1, \\ & y_j \geq 0, \quad j = 1, \dots, m. \end{array}$$

Observe que el problema dual es factible, y que el valor que minimiza es el cero, si y sólo si existe un vector  $(y_1, \dots, y_m)$  con el cual la esperanza de retorno sea cero. Así mismo, el problema primal es factible porque  $x_i = 0$  para  $i = 1, \dots, n + 1$  debido a que de esta manera no se tienen estrategias que nos permitan ganar o perder, además se satisfacen las restricciones del programa, de donde se sigue, por el Teorema (3.2), que si ambos programas son factibles entonces el valor óptimo para el primal es cero y por lo tanto no existe una estrategia para ganar un monto positivo con probabilidad positiva.

En el caso de que el problema dual sea infactible, entonces el primal no tendrá una solución óptima, esto debido a que existirá una estrategia  $x$  que tenga un rendimiento mínimo positivo y al existir hace que el problema primal sea no acotado <sup>1</sup> es decir, entonces se puede garantizarla existencia de la una estrategia que nos permita ganar y no perder.

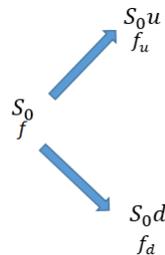
Q.E.D.

### 3.1.1. Árboles binomiales

Los árboles binomiales son diagramas que representan los diferentes caminos que podría seguir el precio de un activo a lo largo de la duración de la opción. El supuesto subyacente es que el precio de un activo sigue la distribución de una caminata aleatoria: en cada tiempo, se tendrá cierta probabilidad de subir o bajar por cierto porcentaje. Si dicho modelo se hace tender a infinito, este modelo nos llevará al supuesto de que el precio del activo se distribuye log-normal (es decir, nos llevará al modelo de Black-Sholes-Merton).

#### Árboles binomiales a un paso

Considere  $S_0$  el precio stock de una opción y sea  $f$  el precio actual de la opción. Supongamos que la opción termina en  $T$  y que mientras está vigente, el precio stock puede subir,  $S_0u$ , donde  $u > 1$  o bajar,  $S_0d$ , con  $d < 1$ . El porcentaje con el que se incrementa el precio stock cuando sube es  $u - 1$ ; así mismo, el porcentaje con que decrece cuando baja será  $1 - d$ . Si el precio stock sube a  $S_0u$ , supondremos que el payoff de la opción es  $f_u$ ; si baja a  $S_0d$ , supondremos que el payoff será  $f_d$ . (Véase la siguiente figura).



<sup>1</sup>Véase el Capítulo 6 de [2]

Entonces, imaginemos un portafolio que consiste de una posición larga en una cantidad delta en acciones y una posición corta en una opción. Se calculo el valor una cantidad delta que hace que el portafolio sea libre de riesgo. Si el precio sube, el valor del portafolio al final de la vigencia de la opción es

$$S_0u\delta - f_u.$$

Y si baja, entonces el valor se convierte en

$$S_0d\delta - f_d.$$

Y suponiendo que son iguales obtenemos que

$$S_0u\delta = S_0d\delta.$$

Entonces

$$\delta = \frac{f_u - f_d}{S_0u - S_0d}. \quad (3.1)$$

En este caso, el portafolio es libre de riesgo y no habrá oportunidades de arbitraje. Se ganará al menos la tasa de interés libre de riesgo. La ecuación (3.1) muestra que  $\delta$  es la tasa del cambio en el precio de la opción entre los nodos al tiempo  $T$ .

Ahora bien, denotemos a  $r$  como nuestra tasa libre de riesgo. El valor presente de nuestro portafolio está dado por

$$(S_0u\delta - f_u)e^{-rT}.$$

El costo al inicio del portafolio es

$$S_0\delta - f,$$

de donde se sigue que

$$f = S_0\delta(1 - ue^{-rT}) + f_ue^{-rT}.$$

Y sustituyendo en (3.1), obtenemos

$$f = e^{-rT}[pf_u + (1 - p)f_d], \quad (3.2)$$

con

$$p = \frac{e^{-rT} - d}{u - d}. \quad (3.3)$$

Cuando consideramos esta probabilidad  $p$  particular, podemos suponer que no existe arbitraje. La justificación matemática se debe al teorema de arbitraje.

En la sección anterior vimos que existirá una  $p$  que nos permitirá al momento de valuar saber que lo que se invierte es lo que se gana. Entonces, por la ecuación (3.2), tenemos que el único vector de probabilidad que resulta tener una esperanza de cero está dado por

$$P[X_i = 1 | X_1 = x_1, \dots, X_{i-1} = x_{i-1}] = \frac{e^{-rT} - d}{u - d},$$

donde los  $x_1, \dots, x_n$  son arbitrarios. Esto implicará que el único vector de probabilidad sobre el conjunto de resultados que se obtienen en todos los ensayos (al ser  $X_1, \dots, X_n$  variables aleatorias independientes) está definido por:

$$P[X_i = 1] = p = 1 - P[X_i = 0], \quad i = 1, \dots, n.$$

y despejando a  $p$  obtenemos la Ecuación (3.3).

### Valuación neutral al riesgo

Notemos que en la Ecuación (3.2) no es necesario hacer ningún supuesto sobre las probabilidades involucradas en los movimientos del precio. Consideremos ahora el supuesto de la ausencia de oportunidades de arbitraje. Se supone que el valor de  $p$  es la probabilidad de que el precio suba y que  $(1 - p)$  es la probabilidad del precio baje, entonces

$$pf_u + (1 - p)f_d$$

es el valor esperado del payoff. Con esta interpretación de  $p$ , la Ecuación (3.2) nos da el valor presente del valor futuro del payoff descontado a una tasa libre de riesgo, donde esta tasa viene dada por el rendimiento mínimo que podemos encontrar en el mercado al invertir en activos libres de riesgo (en el caso de México ésta suele ser la tasa de rendimiento de los CETES). Sin embargo, note que esta tasa puede por un lado, reducir el valor presente de los pagos futuros y por otro, puede hacer que el precio del subyacente aumente de igual manera.

De esta manera, el rendimiento esperado del precio stock al tiempo  $T$ , denotado por  $E(S_T)$ , será

$$E(S_T) = pS_0(u - d) + S_0d.$$

Sustituyendo en (3.3), obtenemos

$$E(S_T) = S_0e^{-rT}. \quad (3.4)$$

De modo que el precio stock crece en promedio igual a la tasa libre de riesgo. Entonces, suponer que la probabilidad de que el precio suba sea igual a  $p$ , es equivalente a suponer que el rendimiento generado por el stock será igual a la tasa libre de riesgo.

En el mundo neutral al riesgo todos los individuos son indiferentes al riesgo. Los inversionistas no requieren una compensación por el riesgo adherido al contrato, y el rendimiento esperado sobre todos los valores es igual a la tasa libre de riesgo.

## 3.2. Entropía en finanzas

En el ámbito de finanzas, no es de extrañar que se busquen maneras nuevas de medir aspectos de interés. Por ejemplo modelar de forma diferente el precio de un activo, considerar otros modelos en donde sí es posible hacer arbitraje, etc. La entropía, como se ha descrito en el Capítulo 1, tiene una interpretación dentro de la teoría de la información. Cover y Thomas (2006), muestran como las propiedades de la entropía física pueden ser aplicados a nuevos campos, entre ellos, la economía. Algunos otros autores se han dado a la tarea de hacer aplicaciones con la entropía. Entre ellos, Arrow (1971) examinó la idoneidad de la entropía como medida de información en economía y finanzas. Rubinstein (1994) sugirió que la entropía podría ser usada como un filtro estadístico para obtener distribuciones a *posteriori* de precios stock. Así mismo, Gulko (1997) transformó el principio de máxima entropía en una manera de hacer valuación de instrumentos financieros, aún sin el supuesto de la ausencia de arbitraje. En este Capítulo examinaremos a detalle los supuestos necesarios para dichos resultados.

### 3.2.1. Arbitraje

Una noción fundamental en finanzas es la de arbitraje. Decimos que hay posibilidad de arbitraje si es posible realizar una inversión con ganancia positiva segura. Esta noción está inspirada por el hecho de que, en el campo financiero lo que se busca es tener una ganancia sin tener que arriesgar demasiado. A continuación ejemplificamos la noción anterior y desarrollamos un contexto adecuado para escribir tal noción en términos matemáticos.

Considere un mercado competitivo en donde tenemos un portafolio de inversión constante, con un bono libre de riesgo y un número finito  $n$  de activos riesgosos sobre un intervalo de tiempo. Al inicio del contrato, el precio del bono será  $P$  y el precio de los activos serán  $p_1, \dots, p_n$ . Por lo que, el precio de nuestro portafolio al inicio del contrato podemos denotarlo por

$$p = [P \ p_1 \ p_2 \ \dots \ p_n].$$

Al término del contrato, al tiempo  $T$ , el bono libre de riesgo puede ser representado como un vector columna unitario de dimensión  $n$ , mientras que el precio de los activos riesgosos serán representados por un vector columna, donde en cada entrada tendremos el precio del activo  $i$ , y estos precios serán representados por  $A_i$ . Así al tiempo  $T$ , el valor de nuestro portafolio, está dado por la suma de sus activos:

$$A = [1 \ A_1 \ A_2 \ \dots \ A_n]$$

La regla del precio lineal de Ross (1976), nos garantiza que existe un vector estrictamente positivo  $q = [q_1 \ q_2 \ \dots \ q_m]$  tal que

$$p = q \cdot A.$$

Así, consideremos un vector renglón  $f = [f_1 \ f_2 \ \dots \ f_m]$  y definimos  $q = f \cdot P$ ; por construcción se sigue que el vector  $f$  debe ser positivo. Entonces,  $p = q \cdot A = P \cdot f \cdot A$ , y se sigue que  $P = P \cdot f \cdot 1$ . Por lo que, se concluye que  $f$  representa una posible distribución de probabilidad positiva, y por la regla del precio lineal, se tiene una alternativa dada por

$$p = P \cdot E(A),$$

donde  $E$  es la esperanza con respecto a la probabilidad dada por  $f$ . En esta ecuación, es posible ver que aparece el descuento de riesgo neutral, en la regla antes mencionada; se conoce a  $f$  como *la probabilidad de las creencias neutrales al riesgo*. El problema ahora, se verá reducido a encontrar a la distribución  $q$  o a la distribución  $f$ .

Sin embargo, no existe un camino ya definido a seguir para poder encontrar el espacio de precio  $q$  o la probabilidad de creencias  $f$ . La búsqueda de  $f$  ha sido discutida por Cox (1975), Cox y Ross (1976) y Merton (1976) donde se ofrecen algunos ejemplos con ayuda del método de máxima-verosimilitud.

Pues bien, con el fin de determinar la función de probabilidad neutral al riesgo podemos recordar que el funcional  $H(X)$  es también, en algún sentido, una función de verosimilitud, y considerando como  $g(i)$  a la probabilidad de que el activo tenga un precio  $i$ , una alternativa es considerar a la función  $f$  como aquella que maximice la función de máxima verosimilitud y la regla del precio lineal. Dicho de otra manera,  $f$  es la función que da

solución al siguiente programa

$$\begin{aligned}
 \text{Maximizar} \quad & - \sum_{i=1}^m g_i \log(g_i) \\
 \text{sujeto a} \quad & \sum_{i=1}^m g_i = 1 \\
 & \sum_{i=1}^m g_i S_i = \frac{S}{P} \\
 & g_i > 0 \text{ para todo } i.
 \end{aligned}$$

La función de verosimilitud  $H(g) = - \sum_{i=1}^m g_i \log(g_i)$  sujeta solamente a su primer momento tendrá siempre una única solución pues como vimos en el Capítulo 1, estamos considerando una función cóncava sobre un conjunto convexo. Y así mismo, la solución a este programa, está dada por la ecuación (2.1).

Así, en el caso de conocer de forma *a priori* la varianza y el precio,  $f$  será una distribución normal con soporte en el soporte permitido para tales densidades.

Como se vió anteriormente, el hecho de que se tome en cuenta la varianza, el precio o la curtosis de manera *a priori*, puede influenciar de manera directa en el análisis y el resultado. Esto quiere decir que la cantidad de información disponible altera significativamente a nuestra solución  $f$ . La función máxima verosímil  $f$  representa también las creencias colectivas del mercado expresadas matemáticamente a través de las restricciones del programa. Es decir cuando  $f$  está determinada por información *a priori*, estamos pensando que esto representa el conocimiento común de varios inversionistas respecto al activo.

### 3.2.2. Eficiencia de mercado

Como ya sabemos,  $f$  representa la función de máxima verosimilitud de la creencia colectiva de los inversionistas. Pueden ofrecerse dos interpretaciones de este hecho: una estadística y la referente a la teoría de la información.

La primera, es ver a la función de máxima verosimilitud como un medio para filtrar las creencias neutrales al riesgo, incluyendo el precio actual  $S$ ,  $P$  y la varianza, para así poder obtener el precio futuro  $S_T$ .

Mientras que la segunda, nos habla de cómo dicha función es informacionalmente eficiente para obtener los precios del mercado. Suponga que la función  $H(g) = - \sum_i g_i \log(g_i)$  mide el grado de incertidumbre de las creencias colectivas sobre los posibles cambios en el precio del activo. De esta manera, conociendo los valores de  $S$ ,  $P$  y  $\sigma^2$ , estamos encontrando el posible precio que contenga mayor incertidumbre, pues se tiene en cuenta a toda la información relevante influenciando dicho precio.

El precio del activo que refleje toda la información disponible sobre él, es llamado *precio*

*del mercado eficiente.* Este precio, mantendrá a los inversionistas en completa incertidumbre sobre el precio de un tiempo futuro, así  $H(g)$  se asumirá como un índice para medir la incertidumbre colectiva que arrojará como resultado la eficiencia en el máximo. Aún más, si  $H(g)$  es un índice de la incertidumbre del mercado, entonces el programa de máxima verosimilitud será una condición necesaria para el eficiencia informacional del mercado.

La interpretación económica que se deriva del teorema de máxima entropía es la siguiente: *la información eficiente de un activo maximiza la entropía de las creencias colectivas del mercado condicionadas a la información apriori con la que se cuente.* Así, en la teoría de la información, la creencia del mercado con máxima entropía debería prevalecer. Entonces, si la entropía mide el grado de incertidumbre del mercado, entonces la máxima entropía deberá ser una condición necesaria para obtener el precio de mercado informacionalmente eficiente.

El programa de máxima entropía con media y varianza conocidas tiene una única  $f$  que lo maximiza porque el funcional  $H(g)$  es cóncavo sobre un conjunto convexo. Si se supone como cierta la interpretación de la teoría de la información, y además, se supone también que el precio del mercado es informacionalmente eficiente, entonces la distribución  $f$  será la única creencia neutral al riesgo en el mercado. Por otra lado, si se considera la interpretación estadística,  $f$  es la creencia de mercado neutral al riesgo más probable aunque no es necesariamente la única.

La aproximación de máxima verosimilitud para la valuación de activos con la interpretación de la teoría de la información, es llamada *la teoría del precio entrópico.* En el siguiente capítulo veremos cómo se valúa bajo esta teoría en el contexto de un activo que se comporta de forma similar al modelo de Black-Scholes-Merton.



# Capítulo 4

## Teoría de precio entrópico

La palabra información, en teoría de valuación entrópica, se refiere a cualquier noticia que pueda determinar el precio de un activo aleatorio y que es posible conocer por adelantado.

Al trabajar con entropía en un mercado eficiente, es necesario saber que se trabajará bajo *probabilidad neutral al riesgo*, es decir, aquella que hace que los precios descontados ( $S_t$ ) de un activo financiero sean martingala<sup>1</sup>. Del mismo modo, se estudiará la existencia de *estrategias de arbitraje*, que son aquellas tales que con probabilidad positiva vamos a ganar. Todo esto se abordará tomando en cuenta las llamadas *creencias de mercado* sobre el precio ( $S_T$ ), las cuales veremos como una probabilidad  $g$  sobre  $S_T$ . Es decir, son todos los posibles escenarios que se pueden tener sobre el precio.

Cuando se habla de mercados eficientes de capital, nos referimos a que los precios de los activos reflejan totalmente la información disponible sobre ellos.

Una condición necesaria para que los inversionistas tengan un incentivo para negociar hasta que los precios reflejen totalmente toda la información relativa a ellos, es que el costo de adquisición de la información y negociación sea cero. Como estos costos son claramente positivos en la vida real, una exigencia más realista es que los precios reflejen la información tal que los costos de obtención del activo no excedan al beneficio dado.

Algunos autores exigen que los precios reflejen exactamente la información fundamental para que un mercado sea eficiente. Sin embargo, las hipótesis de los mercados eficientes se refieren simplemente a la rapidez de incorporación de la información, pero no si reflejan los precios corrientes. Entendemos la hipótesis de que los precios reflejen los valores fundamentales como racionalidad del mercado.

Dentro de la teoría de valuación de opciones se han cuestionado muchos de los supuestos del mercado. En las siguientes secciones presentamos dos formas de valuación. De esta

---

<sup>1</sup>Se dice que  $Z_n$  una sucesión de variables aleatorias es martingala con filtración  $f_n$  si:

- $E[|Z_n|] < \infty \forall n$ .
- $Z_n$  está adaptada a  $f_n$  entonces  $Z_n$  es  $f_n$  - medible  $\forall n$ .
- $E[Z_{n+1}|f_n] = Z_n$

manera, identificaremos las similitudes y diferencias entre cada teoría, y los supuestos necesarios sobre el mercado.

## 4.1. Elementos deterministas en la valuación de una opción

En la teoría, se contemplan dos tipos de opciones fundamentales:

- **Opción call**

Da al comprador el derecho, más no la obligación, a comprar un activo subyacente. El vendedor tiene la obligación de vender si el comprador así lo desea.

- **Opción put**

Da al comprador el derecho, más no la obligación, de vender el subyacente. El vendedor tiene la obligación de comprar si el comprador ejerce la opción.

Las ganancias y pérdidas de estas opciones, según las posiciones tomadas, se ven como:

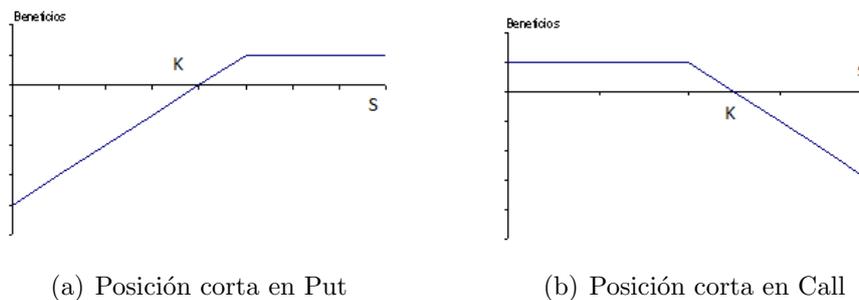


Figura 4.1: Funciones de pago de una opción Call y Put.

Para su valuación es necesario tomar en cuenta que el valor de la opción a su vencimiento esta compuesto por el *valor intrínseco* dado por<sup>2</sup>

$$\text{Call: } \max(S_T - K, 0),$$

$$\text{Put: } \max(K - S_T, 0).$$

Mientras que el *valor temporal*, es decir, el valor de la posibilidad de un movimiento de precios favorables antes del vencimiento de la opción<sup>3</sup>; estará en función del tiempo al vencimiento, volatilidad y tasa libre de riesgo.

Cuando la opción tiende al vencimiento, el valor temporal tiende a cero. Entre los métodos más destacados para la valuación de una opción se encuentran los llamados *árboles binomiales* y el *modelo de Black-Scholes*.

<sup>2</sup>En caso de que exista una prima, se suma o se resta según sea el caso

<sup>3</sup>Un ejemplo de esto, son los diferentes tipos de precios que se usan en los árboles binomiales

## 4.2. Teoría clásica de valuación de opciones (Black-Scholes-Merton)

La teoría clásica de valuación de opciones, fue publicada en 1973 por Robert C. Merton en *Theory of rational option pricing*. En él se hacía referencia a un modelo matemático desarrollado por Fisher Black y Myron Scholes. Dicho modelo es conocido como el modelo de *Black-Scholes-Merton* y actualmente, todavía es utilizado por entidades financieras para la valuación de estos derivados. Para presentar dicha teoría, comencemos por introducir a los árboles binomiales, cuyo refinamiento nos conduce a la teoría previamente mencionada.

El modelo de Black-Scholes-Merton, es uno de los métodos más importantes en la teoría de valuación. Supone que el comportamiento de los precios sigue una distribución lognormal y se basa en los modelos estocásticos de Wiener y el conocido lema de Itô. De esta manera, siguiendo los argumentos de arbitraje, se determina una ecuación diferencial parcial de segundo orden cuya solución representa el precio de la opción.

Los supuestos principales que maneja este modelo son los siguientes:

- La tasa libre de riesgo de corto plazo es conocida y es constante durante el tiempo de vigencia de la opción.
- El comportamiento del precio del subyacente sigue una caminata aleatoria en tiempo continuo, es decir, un movimiento browniano; y la distribución marginal de dichos precios es lognormal.
- La volatilidad de los rendimientos del valor del subyacente no cambia en el tiempo de vigencia de la opción.
- La opción solamente puede ser ejercida al termino del tiempo pactado<sup>4</sup>.
- No existe la oportunidad de arbitraje.
- El activo subyacente tiene liquidez en un mercado eficiente.

### Propiedad lognormal del precio stock

Supongamos que en promedio los cambios en el precio stock en un periodo corto de tiempo se distribuye normal. Entonces definamos los valores:

$\mu$ : Rendimiento esperado del stock

$\sigma$ : Volatilidad adherida al precio stock

La media del rendimiento al tiempo  $\Delta t$  es  $(\Delta t)\mu$  y la desviación estándar es  $\sigma\sqrt{\Delta t}$ , entonces

$$\frac{\Delta S}{S} \sim N(\mu\Delta t, \sigma^2\Delta t), \quad (4.1)$$

con  $\Delta S$  es el cambio en el precio stock al tiempo  $\Delta t$  y  $N$  denota una distribución normal estándar con media  $\mu\Delta t$  y varianza  $\sigma^2\Delta t$ .

<sup>4</sup>Este supuesto sólo es válido para el caso de una opción europea

Por medio de herramientas del cálculo estocástico (Veáse Hull, Capítulo 12) es posible probar que

$$\ln S_T - \ln S_0 \sim N \left[ \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right],$$

de lo que se sigue que

$$\ln S_T \sim N \left[ \ln S_0 + \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right], \quad (4.2)$$

donde  $S_T$  es el precio stock al tiempo  $T$  y  $S_0$  es el precio stock inicial. La Ecuación (4.2) muestra que  $S_T$  tiene una distribución lognormal.

### Distribución de la tasa de rendimiento

La distribución lognormal del precio stock puede ser usada para, obtener información de la distribución de probabilidad de la tasa de rendimiento continua asociada al precio stock entre el tiempo 0 y  $T$ . Si definimos a esta tasa anual como  $x$  entonces,

$$S_T = S_0 e^{xT},$$

de donde

$$x = \frac{1}{T} \ln \frac{S_T}{S_0}.$$

Y por la ecuación (3.2), se sigue que

$$x \sim N \left( \mu - \frac{\sigma^2}{2}, \frac{\sigma^2}{2} \right). \quad (4.3)$$

Así, la tasa de rendimiento anual es una distribución normal con media  $\mu - \frac{\sigma^2}{2}$  y desviación estándar  $\sigma\sqrt{T}$ . Cuando  $T$  crece, la desviación estándar de  $x$  crece.

### Fórmula de Black-Scholes-Merton

Las fórmulas de Black-Scholes-Merton para encontrar el precio de una opción europea al tiempo 0, para el precio de una opción call, son:

$$c = S_0 \Phi(d_1) - K e^{-rT} \Phi(d_2) \quad (4.4)$$

y para el precio de una opción put,

$$p = K e^{-rT} \Phi(-d_2) - S_0 \Phi(-d_1), \quad (4.5)$$

donde

$K$ : precio strike,

$S_0$ : precio stock,

$r$ : tasa de rendimiento,

$\sigma$ : volatilidad,

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

$$d_2: d_1 - \sigma\sqrt{T}.$$

La derivación correspondiente puede consultarse en [8]. La función  $\Phi(x)$  es la distribución de probabilidad acumulada para una normal estándar.

### 4.3. Teoría del precio entrópico

La teoría del precio entrópico (EPT por sus siglas en inglés) fue introducida por Gulko (1997) como un intento para formalizar y operar la hipótesis del mercado eficiente.

Se supondrá que el precio de un activo riesgoso refleja *las creencias del mercado* sobre el comportamiento del precio futuro. De esta manera, las creencias del mercado sobre un posible precio futuro están caracterizadas por la máxima incertidumbre que estas puedan desarrollar sobre dicho precio. Sin embargo, si el mercado tiene información privilegiada sobre dicho activo, entonces el precio futuro no será *informacionalmente eficiente*, es decir no contará con la misma cantidad de incertidumbre pues se tiene acceso a información que podría cambiar de manera poco aleatoria el precio.

Considere lo siguiente:

**Supuesto 4.1** *En economía matemática, el modelo Arrow-Debreu afirma que bajo los supuestos de preferencias convexas, competencias perfectas e independencia de la demanda; debe existir un conjunto de precios tal que las ofertas agregadas sean iguales a las demandas agregadas para cada bien en la economía.*

Entonces, suponga que se tiene un activo riesgoso en una economía que satisface el supuesto anterior en un intervalo  $[0, T]$ . Sea  $S_T$  la variable aleatoria definida por el precio del activo al tiempo  $T$  con valores en el espacio de estados  $P$ , un subconjunto de los reales. Definamos ahora  $g(S_T)$  como la función de probabilidad sobre  $S_T$  es decir,  $g(\cdot)$  representa a las posibles creencias que se tengan sobre el activo, pues otorga probabilidades a cada uno de los diferentes escenarios que este pueda tomar. Se está interesado en la eficiencia del mercado  $f(S_T)$ . Sea  $H(g)$  el operador de entropía, que nos representa la incertidumbre que presenta el mercado sobre  $S_T$  definido en el conjunto de creencias  $g(S_T)$ . De esta manera la eficiencia de mercado estará representada por una densidad que maximiza  $H(g)$ .

A la densidad  $f(S_T)$  que maximice al operador entrópico  $H(g)$  se le conoce como *creencia de máxima entropía*. De acuerdo a la EPT, la creencia de mercado de máxima entropía prevalece aún en un mercado informacionalmente eficiente, esto de acuerdo a [7]. Lo anterior le da gran importancia al índice entrópico:  $H(f(S_T))$ .

El operador  $H(g)$  se definirá de la misma manera que en el Capítulo 1 de este trabajo:

$$H(g) = - \int_P g(x) \log(g(x)) dx.$$

Así, se formaliza la propiedad de *máxima ignorancia* de las creencias de un mercado eficiente, es decir, se verá a  $H(g)$  como el ponderador de todas las posibles creencias que tengan los inversionistas sobre el activo. Una propiedad similar suele ser usada en finanzas continuas, donde la incertidumbre tiende a ser representada como una martingala, un proceso estocástico cuyo valor futuro es igual al valor de hoy. Un precio martingala, en efecto, expresa la incertidumbre del mercado sobre la esperanza de cambio sobre el tiempo.

En resumen, las creencias de máxima entropía  $f(S_T)$  maximizan a  $H(g)$  sujeta a toda la información relevante sobre  $S_T$ .

### 4.3.1. Principio de máxima entropía en la valuación de opciones

El desarrollo análogo a la fórmula de Black-Scholes-Merton se hará en el contexto usual, es decir, se considerará una opción europea que no paga dividendos. Esta teoría más allá que proponerse como un reemplazo de la actual, puede ser un soporte adicional para poder tomar una decisión, pues la teoría actual ofrece un precio libre de arbitraje, mientras que la teoría entrópica ofrece un precio de máxima entropía.

La aproximación de máxima entropía para el análisis de los precios de una opción es una extensión directa del teorema de arbitraje y el precio neutral al riesgo. Como se sabe, la ausencia de arbitraje nos da la posibilidad de garantizar una creencia neutral al riesgo, sin embargo, esta última carece de unicidad y un camino para construirla. Cabe resaltar, que en el caso de trabajar con mercados económicos incompletos<sup>5</sup>, la EPT resuelve este problema, pues por medio de mercados informacionalmente eficientes, las creencias de máxima entropía deben de prevalecer a pesar de las preferencias de los inversionistas hacia inversiones riesgosas. Es decir, no importa que tan riesgosa sea dicha inversión, la máxima entropía de los precios debe ser la misma para una inversión riesgosa o no. Esto se debe a que el operador de entropía no depende de la varianza (que es una medida de dispersión) sino exclusivamente de las probabilidades de que una variable aleatoria tome cierto valor.

Diremos que el precio de una opción será *carente de información* si el precio es una función determinada sobre el precio de un bono, el precio stock y la volatilidad, en otro caso el precio de la opción será un precio *informado*. Por ejemplo, tomando la Ecuación (4.4), se puede distinguir que el precio obtenido por Black-Scholes es un precio carente de información pues entre sus parámetros toma en cuenta lo antes mencionado. Los supuestos de falta de información en el precio y en particular, los utilizados en opciones, facilitan la valuación de estas últimas.

Considere un mercado de capital con tres activos: un bono libre de riesgo, un activo riesgoso (stock) y una opción sobre dicho activo. La razón para considerar dentro de nuestro mercado un bono libre de riesgo se explica a continuación.

Lo que se busca es que, el dinero invertido en un activo riesgoso se vea asegurado por el precio del bono. Esta relación nos será útil más adelante.

El precio del bono ( $P$ ) y el precio del activo ( $S$ ) son conocidos, se busca determinar el precio de la opción. La opción tendrá un tiempo de duración  $T$ . En  $T$ , el bono pagará un peso y el precio del activo  $S_T$  será una variable aleatoria dentro del espacio de estados, mientras que la opción paga  $h(S_T)$ . Para una opción call europea  $h(S_T)$  será  $\max\{0, S_T - K\}$ , donde  $K$  es el precio strike marcado de la opción. Sea  $\sigma^2$  la varianza de  $S_T$ .

Supongamos que el precio de dicha opción es carente de información. De esta manera, se busca encontrar la densidad libre de riesgo  $f(S_T)$ . La densidad neutral al riesgo es inferida por medio de la información conjunta que contengan  $P$ ,  $S$  y la varianza. El

<sup>5</sup>Un mercado incompleto es un tipo de fallo de mercado que ocurre cuando no existe una demanda suficiente como para producir. Esto se debe a que el nivel de la demanda hace que el mercado no sea rentable y por ello las empresas no tienen interés en invertir en el sector.

precio de la opción no es incluido en este conjunto de información pues este no es más informativo que las variables anteriores combinadas.

Así, es posible entablar un programa parecido al presentado en el Capítulo 2 de esta tesis, sin embargo, para esta ocasión se contará con una variante:

**Proposición 4.1** *Dado un conjunto de información  $I = \{\mathcal{P}, S, P, \sigma^2\}$ , la densidad neutral al riesgo  $f(S_T)$  resuelve el siguiente programa:*

$$\begin{aligned} \text{Max} \quad & - \int_{\mathcal{P}} g \log(g) dx, \\ & \int_{\mathcal{P}} g(x) dx = 1, \\ & \int_{\mathcal{P}} xg(x) dx = \frac{S}{P}, \\ & \int_{\mathcal{P}} x^2g(x) dx = \sigma^2 + \left(\frac{S}{P}\right)^2, \\ & \mathcal{C} = \{x : g(x) > 0\}. \end{aligned}$$

La demostración se presenta en el Teorema 2.1, del Capítulo 2 de este documento.

Las diferentes restricciones se dan por como buscamos la distribución neutral al riesgo y el precio del bono. La primera restricción nos habla sobre tomar densidades que representen medidas de probabilidad. La segunda sobre la positividad de  $g(S_T)$  definida sobre el precio neutral al riesgo. La tercera se refiere al valor de la varianza.

**Proposición 4.2** *La única solución al programa de la Proposición (4.1) es*

$$f(S_T) = \frac{\exp(\lambda_1 x + \lambda_2 x^2)}{\int_{\mathcal{P}} \exp(\lambda_1 x + \lambda_2 x^2) dx}$$

donde  $\lambda_1$  y  $\lambda_2$  son elegidos de tal manera que  $f(S_T)$  satisface el precio y la varianza construidas.

**Demostración:**

Recordemos que, por la proposición anterior, la  $f$  que maximiza el sistema es

$$f(x) = \exp(\lambda_0 + \lambda_1 x + \lambda_2 x^2 - 1). \quad (4.6)$$

Ahora necesitamos encontrar los valores de  $\lambda_0$ ,  $\lambda_1$  y  $\lambda_2$ . De acuerdo al Ejemplo 2.5, que nos habla de la distribución normal (presentado en el Capítulo 2), sabemos que para este caso  $f$  será la densidad normal. Así, que igualamos:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x - \frac{S}{P})^2}{2\sigma^2}\right) = \frac{\exp(\lambda_1 x + \lambda_2 x^2)}{\int_{\mathcal{P}} \exp(\lambda_1 x + \lambda_2 x^2) dx}$$

De aquí se desprende fácilmente que  $\lambda_0 = 0$ . Entonces:

$$\begin{aligned} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x - \frac{S}{P})^2}{2\sigma^2}\right) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-x^2 + 2x\frac{S}{P} - (\frac{S}{P})^2}{2\sigma^2}\right) \\ &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(\frac{S}{P})^2}{2\sigma^2}\right)\right) \exp\left(\frac{-x^2}{2\sigma^2} + \frac{x\frac{S}{P}}{\sigma^2}\right) \\ &= c \cdot \exp(\lambda_1 x + \lambda_2 x^2) \end{aligned}$$

De este modo,  $c = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\frac{S}{P}\right)^2}{2\sigma^2}\right)$ ,  $\lambda_1 = -\frac{1}{2\sigma^2}$ , y  $\lambda_2 = \frac{S}{P\sigma^2}$ .

Q.E.D.

Debido a que el precio puede tomar exclusivamente valores no negativos, podemos elegir el soporte  $\mathcal{C}$  sobre el cual se maximizó para que sea  $[0, \infty)$ . La solución correspondiente de máxima entropía  $f(x)$  es una normal truncada cuya cola izquierda es cortada en cero<sup>6</sup>. Explícitamente, la densidad de  $f(x)$  es:

$$N\left(x \mid \frac{S}{P}, \sigma^2; 0, \infty\right) = \begin{cases} \frac{\exp-\left(\frac{(x-\lambda_1)^2}{2\lambda_2^2}\right)}{\int_0^\infty \exp-\left(\frac{(x-\lambda_1)^2}{2\lambda_2^2}\right) dx} & x \in [0, \infty] \\ 0 & x \notin [0, \infty]. \end{cases} \quad (4.7)$$

$\lambda_1$  y  $\lambda_2$  tal como en la demostración.

Ahora, considere la valuación de la opción sobre el activo con falta de información, cuando el proyecto de los inversionistas se basa en el espacio de rendimientos  $\mathfrak{R}$  que pudieran ser generados por el activo. El rendimiento del activo será  $r = \log\left(\frac{S_T}{S}\right)$ . Recuerde que el espacio de  $S_T$  es  $\mathcal{C} = [0, \infty)$ , entonces el espacio de estados de rendimientos será  $\mathfrak{R} = (-\infty, \infty)$ . Al igual que con el precio, es posible construir un programa de máxima entropía para el rendimiento  $r$ :

$$\begin{aligned} \text{Max} & \quad - \int_{\mathfrak{R}} u \log(u) dx \\ \text{sujeto a} & \quad \int_{\mathfrak{R}} u(x) dx = 1 \\ & \quad \int_{\mathfrak{R}} x u(x) dx = \mu \\ & \quad \int_{\mathfrak{R}} x^2 u(x) dx = \sigma_r^2 + \mu^2 \\ & \quad u(x) > 0 \end{aligned}$$

donde  $\mu$  es el rendimiento esperado, y  $\sigma_r^2$  la varianza asociada al rendimiento. El valor neutral al riesgo es determinado por la regla del precio lineal. Denotaremos por  $f_R$  a la densidad de máxima entropía que resuelve el programa anterior.

La densidad con máxima entropía es -de manera similar que con el precio- una densidad normal truncada:  $f_R(x) = n(x \mid \mu, \sigma_r^2; -\infty, \infty)$ . Si se hace el cambio de variable  $R = \log\left(\frac{S_T}{S}\right)$ , el resultado será la densidad lognormal para la variable aleatoria  $S_T$ . La densidad lognormal es considerada de riesgo neutral para el valor neutral al riesgo asociado al rendimiento esperado.

Por definición de  $f_{S_T}$ , sabemos que ésta cumple la regla del precio lineal, es decir, su esperanza es proporcional a lo invertido en el bono y el activo riesgoso. Entonces, al

<sup>6</sup>Las densidades normales truncadas ser consultadas en [10]

esperar un rendimiento neutral al riesgo, tenemos que:

$$\mu = \log\left(\frac{1}{P}\right) - \frac{\sigma_r^2}{2}.$$

Y así, la densidad de probabilidad neutral al riesgo, es:

$$f_{S_T} = \text{Lognormal}\left(\log\left(\frac{1}{P}\right) - \frac{\sigma_r^2}{2}, \sigma_r^2, -\infty, \infty\right).$$

### 4.3.2. Fórmula de valuación de opciones entrópica

Después de haber estudiado una manera diferente de encontrar el precio de una opción, se utilizará este procedimiento para obtener una fórmula similar a la de Black-Scholes-Merton. Para ello, se tomará en cuenta una opción europea de tipo call y un bono libre de riesgo.

Antes de iniciar este análisis, es necesario enumerar algunas propiedades de la distribución normal truncada (ver [10] p.146).

**Definición 4.1** Sea  $X$  una variable aleatoria normal con media  $m$  y varianza  $s^2$  y definida es tal que  $X \in (a, b)$  con  $-\infty \leq a < b \leq \infty$ . Entonces, si condicionamos  $a < X < b$  se dice que  $X$  tiene una distribución normal truncada, con función de densidad:

$$f(x; \mu, \sigma^2, a, b) = \begin{cases} \frac{\frac{1}{\sigma} \phi\left(\frac{x-m}{s}\right)}{\Phi\left(\frac{b-m}{s}\right) - \Phi\left(\frac{a-m}{s}\right)} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}$$

Con  $\phi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$  y  $\Phi(t)$  es la función acumulada de una normal estándar. Así mismo, si  $b = \infty$  entonces  $\Phi\left(\frac{b-m}{s}\right) = 1$ , y de manera similar, si  $a = -\infty$  se tiene que  $\Phi\left(\frac{a-m}{s}\right) = 0$

**Observación 4.3** Si la variable aleatoria  $X$  tiene una distribución truncada, entonces su esperanza está dada por:

$$\mu = E(X | a < X < b) = m + s \frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)}$$

Y su varianza,

$$\sigma^2 = \text{Var}(X | a < X < b) = s^2 \left[ 1 + \frac{\alpha\phi(\alpha) - \beta\phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} - \left( \frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} \right)^2 \right].$$

Con  $\alpha = \left(\frac{a-m}{s}\right)$  y  $\beta = \left(\frac{b-m}{s}\right)$ ,

La demostración puede ser consultada en [10].

Recordemos, cómo se ve la posible función de pago de una opción europea de tipo call:

$$\text{Call} = P \int_K^\infty (x - K) f(x) dx.$$

Y recordando que  $f(x)$ , la función de máxima-entropía, es una normal truncada, que resuelve el programa de la Proposición (4.1) y cuya función de densidad se puede escribir como (4.7). Entonces,

$$\begin{aligned} Call &= P \int_K^\infty (x - K) f(x) dx \\ &= P \int_K^\infty (x - K) \frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{\sigma \left(1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right)} dx \\ &= \frac{P}{\sigma \left(1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right)} \int_K^\infty (x - K) \phi\left(\frac{x - \mu}{\sigma}\right) \end{aligned}$$

Se aplica el cambio de variable:

$$\begin{aligned} y = \frac{x - \mu}{\sigma} &\implies x = \sigma y + \mu \\ \partial y = \frac{\partial x}{\sigma} &\implies \partial x = \sigma \partial y, \end{aligned}$$

Entonces, se tiene:

$$\begin{aligned} Call &= \frac{P}{\sigma \left(1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right)} \int_{\frac{K-\mu}{\sigma}}^\infty \sigma(\sigma y + \mu - K) \phi(y) \\ &= \frac{P}{\sigma \left(1 - \Phi\left(\frac{-\mu}{\sigma}\right)\right)} \left[ \int_{\frac{K-\mu}{\sigma}}^\infty \sigma^2 y \phi(y) dy + \int_{\frac{K-\mu}{\sigma}}^\infty \sigma(\mu - K) \phi(y) dy \right]. \end{aligned}$$

Para el cálculo de la primera integral de la expresión anterior, se hará por partes, entonces se tiene:

$$\begin{aligned} \int_{\frac{K-\mu}{\sigma}}^\infty \sigma^2 y \phi(y) dy &= \sigma^2 \int_{\frac{K-\mu}{\sigma}}^\infty y \phi(y) dy \\ &= \sigma^2 \int_{\frac{K-\mu}{\sigma}}^\infty y \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \\ &= -\frac{\sigma^2}{\sqrt{2\pi}} \left[ e^{-\frac{y^2}{2}} \Big|_{\frac{K-\mu}{\sigma}}^\infty \right] \\ &= -\frac{\sigma^2}{\sqrt{2\pi}} \left( 0 - e^{-\frac{\left(\frac{K-\mu}{\sigma}\right)^2}{2}} \right). \end{aligned}$$

Entonces

$$\int_{\frac{K-\mu}{\sigma}}^\infty \sigma^2 y \phi(y) dy = \frac{\sigma^2}{\sqrt{2\pi}} e^{-\frac{\left(\frac{K-\mu}{\sigma}\right)^2}{2}}.$$

Ahora, calculamos la segunda integral del valor de la opción Call:

$$\int_{\frac{K-\mu}{\sigma}}^\infty \sigma(\mu - K) \phi(y) dy = \sigma(\mu - K) \int_{\frac{K-\mu}{\sigma}}^\infty \phi(y) dy,$$

por la definición de distribución acumula se tiene que

$$\int_{\frac{K-\mu}{\sigma}}^\infty \sigma(\mu - K) \phi(y) dy = \sigma(\mu - K) \left[ 1 - \Phi\left(\frac{K - \mu}{\sigma}\right) \right].$$

Así, uniendo las ecuaciones anteriores, se obtiene:

$$Call = \frac{P}{1 - \Phi\left(-\frac{\mu}{\sigma}\right)} \left[ \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{\left(\frac{K-\mu}{\sigma}\right)^2}{2}} + (\mu - K) \left[ 1 - \Phi\left(\frac{K - \mu}{\sigma}\right) \right] \right]. \quad (4.8)$$

Así, finalmente llegamos a una fórmula cerrada, que es en espíritu similar a la fórmula de Black-Scholes-Merton, pero recurre a otros medios para encontrar el precio futuro de un activo. Por otro lado, es necesario recalcar que no solo se toma en cuenta la manera la incertidumbre o las probabilidades que puede tomar el activo, pues como ya se ha mencionado, esta fórmula tiene en su construcción aspectos subjetivos de la teoría financiera. Dando así, una fórmula matemática, posiblemente más completa.

Una observación importante es que, la Fórmula (4.8) está en términos de  $\mu$  y  $\sigma^2$ . En términos computacionales nos gustaría que estuviese en términos de  $m$  y  $s^2$ , los parámetros de una normal asociada (no truncada) tal que la variable normal truncada tenga media  $\mu$  y varianza  $\sigma^2$ . Sin embargo, no es posible encontrar un procedimiento analítico para reescribir la Fórmula (4.8) en términos de  $m$  y  $s^2$ . En el siguiente capítulo damos una alternativa computacional para resolver tal problema.

Por otro lado, es posible dar algunas diferencias visibles entre la fórmula de Black-Scholes-Merton y la Fórmula (4.8). Éstas son:

- El factor de descuento  $e^{-r(T-t)}$  es reemplazado por el precio del bono, y
- El término de volatilidad  $\sigma\sqrt{T-t}$  es simplemente reemplazado por  $\sigma$ , es decir, es posible reducirlo a la volatilidad que se espera para el activo en tiempo  $t$ .

Sin embargo, es posible que algunos supuestos como lo son la continuidad del precio, las tasas de interés continuas o que la trayectoria del activo se comporte como un proceso movimiento geométrico pueden ser ignorados cuando se trabaja con la EPT. En contraste, el modelo de máxima entropía será utilizado con la varianza final al momento de finalizar el ejercicio.



# Capítulo 5

## Simulación

### 5.1. Aplicación de la fórmula de valuación de opciones entrópica

Para la aplicación del precio entrópico, es necesario hacer un contraste entre la densidad de una normal truncada y la densidad lognormal (que es la ley que se tiene bajo el modelo de Black-Scholes-Merton, el usual). Para esto, fue necesario hacer uso de un programa computacional de búsqueda.

Note que la definición de la normal truncada (Definición (4.3)) involucra una normal no truncada asociada. Es decir, toma a la normal truncada como una transformación de la densidad normal no truncada, sin tener una relación analítica exacta entre los parámetros de la normal no truncada y la esperanza y varianza de la normal truncada. Por esa razón, dada la esperanza y varianza de la normal truncada, buscaremos la esperanza y varianza de la normal no truncada que corresponde a tal asociación. Para ello se utilizó un método computacional de búsqueda estocástico bajo la función de costo:

$$c(x, y) := \left[ x + y \left( -\frac{\phi\left(-\frac{x}{y}\right)}{1 - \Phi\left(-\frac{x}{y}\right)} \right) - \mu \right]^2 + \left[ y^2 \left( 1 + \frac{\frac{-x}{y}}{1 - \Phi\left(-\frac{x}{y}\right)} - \left( 1 + \frac{\frac{-x}{y}}{1 - \Phi\left(-\frac{x}{y}\right)} \right)^2 \right) - \sigma^2 \right]^2,$$

que corresponde a la distancia entre la esperanza y varianza de la normal truncada y la esperanza y varianza de la normal no truncada asociada. La idea es encontrar  $(x, y)$  en  $\mathbb{R}^2$  tal que  $c(x, y)$  sea igual a cero. Esto se resuelve de manera aproximada (con un error pequeño). El código establecido para este propósito puede consultarse en el apéndice.

Veremos ahora en qué casos la Fórmula de valor de la opción (4.8) obtenida mediante la propuesta de Gulko, puede competir con la propuesta por Black-Scholes-Merton.

Como ejemplo utilizaremos tres dinámicas diferentes para el precio de un activo: el Movimiento Browniano Geométrico, el proceso de Poisson Compuesto y el proceso de Cox-Ingersoll-Roll.

## 5.2. Movimiento Browniano Geométrico

El Movimiento Browniano Geométrico se utiliza frecuentemente para modelar los cambios que tienen en el tiempo los precios de algún activo determinado, por ejemplo las acciones de una empresa, el tipo de cambio, el precio del petróleo, etc. Por ello, el modelo de Black-Scholes-Merton toma como base en su formulación el supuesto de que el precio del activo con el que se negocia se mueve durante el tiempo con la trayectoria de un Movimiento Browniano Geométrico<sup>1</sup>, con lo cual todas sus características (esperanza, varianza, etc.) son conocidas.

Para la valuación de una opción Call, consideremos los siguientes datos:<sup>2</sup>

- $k= 50, 100, 150, 200$ ,
- $r= 3\% \text{ y } 60\%$ ,
- $S_0=125$ ,
- $t=0.25$  años,
- $\sigma=0.3$ .

Se verificará el precio de la opción bajo diferentes precios strike y rendimientos, con el fin de comprobar si se cumple o no la fórmula del precio entrópico.

A continuación se muestra la simulación de algunas trayectorias en R del Movimiento Browniano geométrico<sup>3</sup> con la volatilidad y los rendimientos antes mencionados. Tenemos:

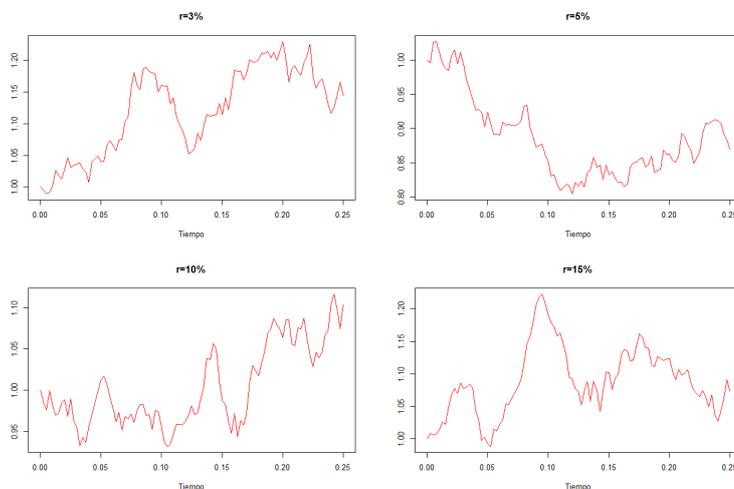


Figura 5.1:  
Trayectorias  
simuladas de  
un proceso  
Browniano  
geométrico  
con base en  
los paráme-  
tros antes  
mencionados.

<sup>1</sup>Vease el Capítulo 14 de Hull [8]

<sup>2</sup>Datos propios, no se toma en cuenta ningún contrato de opción real.

<sup>3</sup>Vease el código expuesto en el Apéndice

Con la ayuda del método Monte-Carlo, se utilizó el siguiente estimador:

$$f = \frac{\sum_{i=1}^n \exp(-rt_i) \cdot \text{máx}(0, S_{T_i} - k)}{n}.$$

La siguiente tabla presenta los resultados obtenidos de la simulación, la aplicación directa de la fórmula de Black-Scholes y la fórmula entrópica, utilizando un rendimiento del 3%:

Strike/Fórmula	Monte-Carlo	Black-Sholes-Merton	Gulko
50	73.28	75.87	80.19
100	39.04	37.27	38.71
150	17.19	16.70	16.97
200	8.57	7.40	9.49

Mientras que los resultados obtenidos con un rendimiento igual al 6% son:

Strike/Fórmula	Monte-Carlo	Black-Sholes-Merton	Gulko
50	83.50	82.18	80.69
100	44.30	45.66	46.26
150	22.65	23.09	25.47
200	12.87	11.39	12.84

Dado lo anterior, es posible afirmar que la fórmula de Black-Scholes-Merton es muy cercana a los precios simulados de la opción. Éste es un resultado poco sorprendente ya que, la fórmula de Black-Scholes-Merton se obtiene bajo el supuesto de un activo que se comporta como un Movimiento Browniano Geométrico. Por otra parte, tenemos la fórmula entrópica que proporciona soluciones similares al precio simulado. Si se hace una comparación directa entre ambas fórmulas, se puede ver que no existe gran discrepancia entre ellas aunque la fórmula clásica es mejor.

### 5.3. Proceso Poisson Compuesto

En esta sección supondremos que el precio del activo se comporta como un un proceso Poisson compuesto. La decisión del uso de este modelo, debe al hecho de querer contrastar la fórmula utilizando un proceso con parámetros conocidos (esperanza y varianza) y que involucra saltos.

El proceso Poisson compuesto tiene la forma:

$$X_t = \sum_{i=1}^{N_t} Y_i$$

donde  $N_t$  es un proceso Poisson Homogéneo y  $Y_i$  son variables idénticamente distribuidas e independientes (entre ellas y del Proceso  $N_t$ ). Supondremos, además, que las  $Y_{i's}$  tienen distribución exponencial. Los parámetros que se utilizarán para la simulación son:

- $k= 1, 2, 3$  y  $4$ ,
- $r= 3\%$ ,
- $S_0=2.5$ ,
- $t=3$  años,
- $\sigma=0.3$ ,
- $\lambda_1=0.06$ ,
- $\lambda_2=0.09$ ,

donde  $\lambda_1$  es la tasa del Proceso Poisson Homogéneo y  $\lambda_2$  es la tasa de la distribución exponencial. Al simular el proceso, con el código programado en R (ver Apéndice), se obtuvo la trayectoria siguiente.

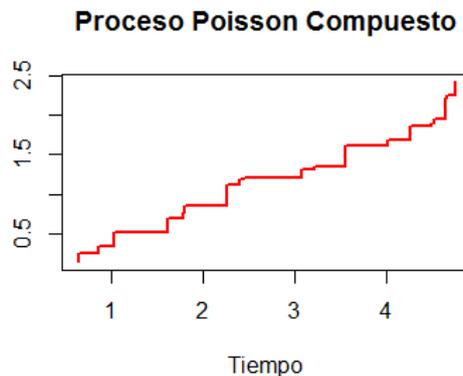


Figura 5.2: Trayectoria simulada de un proceso Poisson compuesto.

Para obtener el precio debemos tomar en cuenta que la fórmula de Black-Scholes-Merton toma sus parámetros de la distribución lognormal, por lo que se necesita igualar las esperanzas del proceso Poisson compuesto y la esperanza de la lognormal, en el tiempo  $t$ .

La esperanza y varianza de la distribución lognormal son:

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right)t, \quad (5.1)$$

$$\text{Var}(X) = \exp^{2\mu t + \frac{\sigma^2}{2}t} (\exp^{\sigma^2 t} - 1). \quad (5.2)$$

Mientras que para el proceso de Poisson compuesto con  $Y_i$  variables aleatorias independientes e idénticamente distribuidas con distribución exponencial, se tiene:

$$E(X) = \lambda_1 t E(Y) = \lambda_1 t \left( \frac{1}{\lambda_2} \right),$$

$$Var(X) = \lambda_1 t E(Y^2) = \lambda_1 t \left( \frac{\lambda_2}{\lambda_2 - 2} \right).$$

igualando las esperanzas y varianzas,

$$\exp^{\mu + \frac{\sigma}{2}} = \lambda_1 t \left( \frac{1}{\lambda_2} \right),$$

$$\exp^{2\mu + \frac{\sigma}{2}} (\exp^{\sigma^2} - 1) = \lambda_1 t \left( \frac{\lambda_2}{\lambda_2 - 2} \right).$$

Y tomando en cuenta que  $\lambda_1$  y  $\lambda_2$  son parámetros conocidos, se obtiene:

$$\sigma^2 = \frac{\log \left( \frac{\lambda_2 \lambda_1 t}{\lambda_2 - 2} \right)}{t}$$

$$\mu = \frac{2 \left( \log \frac{\lambda_1 t}{\lambda_2} \right) - \log \left( \frac{\lambda_2 \lambda_1 t}{\lambda_2 - 2} \right)}{t}$$

Ahora bien, para la Fórmula (4.8) del Capítulo 4, la esperanza y varianza utilizadas dentro de ésta, fue la obtenida mediante la aproximación de los momentos correspondientes al proceso Poisson compuesto. Por lo que, los resultados fueron:

Strike/Fórmula	Monte-Carlo	Black-Sholes-Merton	Gulko
1	2.71	2.499203	1.14
2	2.41	2.498864	0.2722
3	1.49	2.49860	0.2688
4	0.96	2.498394	0.07745

A comparación de los resultados obtenidos en la sección anterior, estos son pocos favorables para ambas formulaciones. Por un lado, la fórmula de Black-Scholes-Merton no logra acercarse al valor que se toma como absoluto (el precio simulado), la fórmula no es sensible a la disminución en el precio strike y presenta precios casi constantes en las diferentes circunstancias presentes.

Por otro lado, la Fórmula (4.8), si disminuye conforme el precio strike baja pero subpondera demasiado el valor al que se quiere llegar. Esto puede deberse posiblemente a la naturaleza del proceso.

## 5.4. Proceso de Cox-Ingersoll-Ross

El proceso de Cox-Ingersoll-Roll es también de los procesos estocásticos populares en el mundo financiero, pues muchos modelos de tasas de interés involucran a este proceso. Por ello, resulta interesante utilizarlo como una posible trayectoria de un activo. Este proceso se define a través de la ecuación diferencial estocástica:

$$dr_t = a(b - r_t)dt + \sigma \sqrt{r_t} dW_t,$$

donde  $a$  es un parámetro positivo no acotado (fuerza de la deriva hacia la media) y  $b$  parámetro positivo no acotado (media a largo plazo del activo). Al simular un proceso de este tipo en R con el método de Euler para ecuaciones estocásticas, se obtiene una trayectoria:

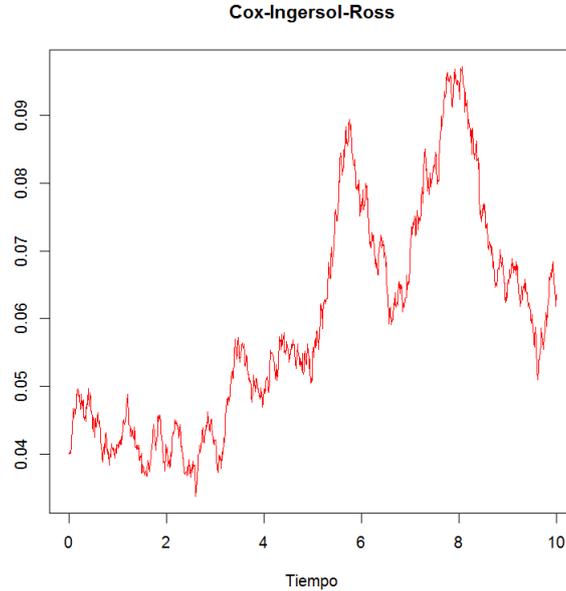


Figura 5.3: Trayectoria simulada de un proceso CIR.

Al igual que en nuestro proceso de Poisson, es necesario igualar la varianza y esperanza del CIR a la esperanza de una distribución lognormal, esto con el fin de poder ser contrastados con la fórmula de Black-Scholes-Merton.

Considere que el Proceso Cox-Ingersol-Ross, cuenta con esperanza y varianza como sigue<sup>4</sup>

$$E(r_t|r_0) = r_0 \exp^{-at} + b(1 - \exp^{-at}),$$

$$Var(r_t|r_0) = r_0 \frac{x^2}{a} (\exp^{-at} - \exp^{-2at}) + \frac{bx^2}{2a} (1 - \exp^{-at})^2.$$

igualando lo anterior con las Ecuaciones (5.1) se tiene:

$$\exp^{\mu + \frac{\sigma}{2}} = r_0 \exp^{-at} + b(1 - \exp^{-at}),$$

$$\exp^{2\mu + \frac{\sigma}{2}} (\exp^{\sigma^2} - 1) = r_0 \frac{x^2}{a} (\exp^{-at} - \exp^{-2at}) + \frac{bx^2}{2a} (1 - \exp^{-at})^2.$$

Note que  $r_0$ ,  $x^2$ ,  $a$  y  $b$  son conocidas<sup>5</sup>, por lo que bastará con despejar a  $\mu$  y  $\sigma^2$ . Se tiene entonces que:

$$\sigma^2 = \frac{\log \left( r_0 \frac{x^2}{a} (\exp^{-at} - \exp^{-2at}) + \frac{bx^2}{2a} (1 - \exp^{-at})^2 \right)}{t},$$

$$\mu = \frac{\log (r_0 \exp^{-at} + b(1 - \exp^{-at})) - \sigma^2/2}{t}.$$

<sup>4</sup>Para calcular los parámetros puede verse [5]

<sup>5</sup> $x^2$  es la volatilidad propia del activo

Al tomar en cuenta lo anterior, considere los siguientes parámetros basados en una opción sobre alguna tasa de interés

- $k = 1\%, 2\%, 3\%$  y  $4\%$ ,
- $r = 4\%$ ,
- $S_0 = 5$
- $t = 25$  años,
- $\sigma = .3$ ,
- $a = 1$ ,
- $b = 1$ .

Los resultados obtenidos son:

Strike/Fórmula	Monte-Carlo	Black-Sholes-Merton	Gulko
1 %	0.39	0.49	0.38
2 %	0.38	0.48	0.36
3 %	0.35	0.47	0.34
4 %	0.34	0.46	0.33

Observemos que existe una semejanza entre el valor simulado y el valor dado por la fórmula entrópica por lo que se puede decir que mide de manera efectiva al del proceso CIR, mientras que la fórmula de Black-Scholes-Merton sobrepondera el valor de la opción.



# Conclusiones generales

En términos generales, el objetivo de esta tesis fue alcanzado satisfactoriamente, pues como se había planteado en el objetivo se consiguió una fórmula cerrada y alterna al proceder actual en la valuación de opciones. A saber, la Fórmula (4.8):

$$Call = \frac{P}{1 - \Phi\left(-\frac{\mu}{\sigma}\right)} \left[ \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{\left(\frac{K-\mu}{\sigma}\right)^2}{2}} + (\mu - K) \left[ 1 - \Phi\left(\frac{K - \mu}{\sigma}\right) \right] \right].$$

tiene ciertas ventajas pues:

1. puede prescindir de la existencia de arbitraje, lo cual hace a este método más general que otros.
2. Toma en cuenta la eficiencia del mercado.
3. Cuenta con un bono para medir el valor presente de lo que nos costaría la opción. Así mismo, es posible verlo como un activo de cobertura ante las pérdidas que pueda ocasionar el activo riesgoso.
4. Su dependencia en dos parámetros relacionados al movimiento del activo (la media y la varianza), permite que algunos modelos como el CIR, puedan ajustarse a ella.

Sin embargo, en el caso del bono usado es necesario pensar en como debe verse y que tanto nos afecta en el precio final (tal procedimiento sale de los alcances de esta tesis). Por otro lado, después de enumerar las ventajas de la fórmula entrópica, se puede decir que tales atributos no implican que la fórmula de Black-Scholes-Merton sea mala, sino que al tener dependencia de un solo parámetro del movimiento del activo (varianza o volatilidad) la deja un poco en desventaja con la fórmula discutida.

Ahora, hablando solo en términos de las simulaciones; se puede decir que el único proceso no modelado de forma eficiente por ambas fórmulas fue el Proceso Poisson, ya que lo obtenido fue deficiente con respecto al valor simulado. Esto puede deberse a la existencia de saltos en la dinámica.

Para finalizar, se puede decir que el uso de la entropía en la valuación de opciones permite considerar aún más procesos relacionados con el entorno económico en el que se mide dicho derivado pero que necesita posiblemente una modelación matemática más compleja. Su perfeccionamiento podría permitirnos ampliarla a otros productos derivados financieros y tener una manera alterna de valorar opciones.



# Apéndice

## Función de Costos

En este apéndice, se presenta el código de rastreo para poder encontrar la esperanza y varianza de la normal asociada a la normal truncada, esto con el fin de encontrar la esperanza y varianza para encontrar el valor de nuestra opción call con la fórmula (4.8) encontrada en el Capítulo 4; así mismo tome en cuenta la función de costo

$$c(x, y) := \left[ x + y \left( \frac{\phi\left(-\frac{x}{y}\right)}{1 - \Phi\left(-\frac{x}{y}\right)} \right) - \mu \right]^2 + \left[ y^2 \left( 1 + \frac{\frac{-x}{y}}{1 - \Phi\left(-\frac{x}{y}\right)} - \left( 1 + \frac{\frac{-x}{y}}{1 - \Phi\left(-\frac{x}{y}\right)} \right)^2 \right) - \sigma^2 \right]^2$$

De esta manera, el programa viene dado por:

```
costo<-function(x,y,mu,sigma){
  aux=x/y
  c1= abs(x + y * ( dnorm(-x/y,0,1)/(1-pnorm(-x/y,0,1)) ) - mu)
  c2= abs(y * (1 + aux*dnorm(aux,0,1)/(1-pnorm(aux,0,1))
    - (dnorm(aux,0,1)/(1-pnorm(aux,0,1)))^2) - sigma^2 )
  c=(c1^2)+(c2^2)
  return(c)
}

reparametriza<-function(n,mu,sigma,alpha,a)
{
  x=rnorm(1,mu,a)
  y=rexp(1,1/(sigma^2))

  for(i in 1:n)
  {
    if( runif(1) < alpha){
      u=rnorm(1,mu,a)
      v=rexp(1,1/(sigma^2))
    }
    else{
      u=x+runif(1,-a,a)
      v=y+runif(1,-a,a)
    }
    if(v[1]>0)
    {
      if(costo(u[1],v[1],mu,sigma)< costo(x[1],y[1],mu,sigma))
      {
```

```

        x=u
        y=v
    }
}
}
return(c(x,y))
}

```

## Movimiento Browniano

En este apéndice, se presenta el código utilizado para la simulación de un Movimiento Browniano Geométrico y el cálculo del precio de un contrato de una opción mediante el método de Monte Carlo. De esta manera, el programa viene dado por:

```

##Browniano estandar

BrowEstandar<-function(mu, sigma ,m, t){
browE<-c()
h<-t/m
tiempo<-seq(0, t, h)
for(i in 1:length(tiempo)-1){
browE[i]=rnorm(1,h*mu, sqrt(h)*sigma)
}
bestandar<-cumsum(browE)
be<-c(0, bestandar)
#x<-plot(tiempo, be, t="l")
return(be)
}

##### Browniano Geo
BG<-function(r, sigma, m, t){
browg<-c()
h<-t/m
tiempo<-seq(0, t, h)
browG=exp((((r-(sigma^2/2))*tiempo)+(sigma*BrowEstandar(0,1, m, t))))
x<-plot(tiempo, browG, t="l")
return(c(x, browG))
}

### Simulacion del precio de la opcion

PrecioBG<-function(s_0, k, r, sigma, t, m, n){
st<-c()
st_final<-c()
for(i in 1:n){
st<-s_0*BG(r, sigma, m, t)
st_final[i]=max(0, st[length(st)]-k)
}
cam<-exp(-r*t)*st_final
precio=(sum(cam))/n
return(precio)
}

```

## Proceso Poisson Compuesto

En este apéndice, se presenta el código utilizado para la simulación de un proceso Poisson Compuesto y el cálculo del precio de un contrato de una opción mediante el método de Monte Carlo. De esta manera, el programa viene dado por:

```
#Simulacion del Proceso Poisson Compuesto
PPC<-function(t, l1, l2){
  N=rpois(1, l1*t)
  unif=c()
  for(j in 1:N){
    u=runif(1, 0, t)
    unif[j]=u
  }
  eventos=sort(unif)
  saltos=rexp(N, l2)
  v=cumsum(saltos)
  tiempo=seq(1, N)
  graf=plot(eventos, v, "s", xlab="Tiempo", ylab="", col="red", lwd=2, main="
  Proceso_Poisson_Compuesto")
  return(c(graf, v))
  return(v)
}

#Simulacion del precio de una opcion

PrecioPC<-function(s_0, k, r, t, l1, l2, n){
  st<-c()
  st_final<-c()
  for(i in 1:n){
    st<-s_0*PPC(t, l1, l2)
    st_final[i]=max(0, st[length(st)]-k)
  }
  cam<-exp(-r*t)*st_final
  precio=(sum(cam))/n
  return(precio)
}
```

## Proceso Cox-Ingersoll-Ross

En este apéndice, se presenta el código utilizado para la simulación de un proceso Cox-Ingersoll-Ross y el cálculo del precio de un contrato de una opción mediante el método de Monte Carlo. De esta manera, el programa viene dado por:

```
CIR<-function(a, b, sigma, r_0, tiempo, n){
  particion=tiempo/n
  aux=c()
  aux[1]=r_0
  for(i in 1:n){
    aux[i+1]=aux[i]+a*(b-aux[i])*particion+sigma*(sqrt(aux[i]))*sqrt(
    particion)*rnorm(1, 0, 1)
  }
  x=seq(0, tiempo, particion)
  graf=plot(x=x, y=aux, type="l", main="Cox-Ingersoll-Ross", col="red", xlab="
  Tiempo", ylab="")
  return(c(aux, graf))
}
```

```
}  
  
#Precion de una Opcion Call  
PrecioCir<-function(s_0,k,r,t,a,b,sigma,r_0,n){  
  st<-c()  
  st_final<-c()  
  for(i in 1:n){  
    st<-s_0*cir(a,b,sigma,r_0,tiempo,n)  
    st_final[i]=max(0,st[length(st)]-k)  
  }  
  cam<-exp(-r*t)*st_final  
  precio=(sum(cam))/n  
  return(precio)  
}
```

# Bibliografía

- [1] ARROW, K.J (1971), *The Value and Demand for Information*, In Decisions and Organizations, edited by C.B. McGuire and R. Radner. Amsterdam: North-Holland, 131-139.
- [2] BAZARAA, M.,JARVIS, J. y SHERALI, H. (2010), *Linear Programming and Network Flows*, cuarta edición, Wiley, New Jersey.
- [3] BILLINGSLEY, P (1999), *Convergence of probability measures*, Wiley, New York.
- [4] COVER, T. y THOMAS, J. (2006), *Elements of Information Theory*, segunda edición, Wiley, New Jersey.
- [5] COX, J.C,INGERSOL, J.E, y ROSS, S.A (1985), *A Theory of the Term Structure of Interest Rates*, Econometrica 53.
- [6] GULKO, L. (1996), *The Entropy Theory of Stock Option Pricing*, Yale School of Management, Mayo.
- [7] GULKO, L. (1997), *Dart Boards and Asset Prices: Introducing the Entropy Pricing Theory*, Capítulo 14 en *Advances in Econometrics*, T.B. Fomby and R.C Hill, eds. JAI Press, Greenwich, CT.
- [8] HULL, J. (2012), *Options, Futures and Other Derivatives*, cuarta edición, Prentice Hall, Boston.
- [9] IZE, J.(2002), *Cálculo de variaciones*, Serie FENOMECC, UM. Editorial Impretei.
- [10] JOHNSON, N.,KOTZ, S. y BALAKRISHNAN, N. (1994), *Continuous Univariate Distributions Vol. 1* , segunda edición, Wiley, New York.
- [11] LUENBERGER, D. (1969), *Optimization by vector space methods*, primera edición, Wiley, Stanford, California.
- [12] MERTON, R. C.(1973), *Theory of Rational Option Pricing*. The Bell Journal of Economics and Management Science, 141-183.
- [13] RUBINSTEIN, M.(1994), *Implied Binomial Trees*. Journal of finance 49, 771-818.
- [14] ROSS, S.A (1976), *The Arbitrage Theory of Capital Asset Pricing*. Journal of Economic Theory 13, 341-360.
- [15] ROSS, S.A (2010), *A first course in probability*, octava edición, Pearson, México.