



**UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO**

FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN

SIMULACIÓN POR COMPUTADORA DE
MODELOS ESTOCÁSTICOS

TESINA

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADO EN ACTUARÍA

PRESENTA:

JORGE APODACA GAONA

ASESORA: M. EN C. MAYRA LORENA DÍAZ SOSA

NOVIEMBRE, 2014

NAUCALPAN, ESTADO DE MÉXICO



Universidad Nacional
Autónoma de México



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.

Resumen

En este texto se exponen las aplicaciones de la simulación de modelos para evaluar financieramente proyectos de inversión, en los cuales el rendimiento está determinado por precios de productos que siguen un movimiento Browniano. Se presentan las bases de la teoría de cada modelo y método usado. Se hace uso del método de factorización de Cholesky para simular variables aleatorias correlacionadas y se muestra un ejemplo de simulación del precio de la canasta alimentaria con en el software estadístico R. La discusión está enfocada en los alcances de cada modelo y método.

Palabras clave: Simulación, procesos estocásticos, evaluación de proyectos, pobreza, series de tiempo.

Prefacio

Este documento muestra los campos que me generaron interés a lo largo de mis estudios en la carrera de Actuaría. Particularmente, llamó mi atención el progreso de la simulación para medir el riesgo en los últimos años porque contempla diferentes casos que de otra forma serían imposibles de evaluar. La necesidad de contar con métodos matemáticos para medir y administrar el riesgo financiero ha aumentado y el profesionista responsable de hacer frente a esta necesidad es el actuario ya que cuenta con la formación matemática y financiera para solucionar problemas afines. Es por ello que el desarrollo de textos de investigación de las aplicaciones para medir el riesgo es de importancia para el presente y futuro de la profesión actuarial. Por esa razón decidí hacer mi trabajo de titulación a manera de una guía resumida que se enfoca en ayudar a introducir al lector en algunas aplicaciones de la simulación en la evaluación de proyectos, así como mostrar sus alcances con el fin de hacer notar su dinamismo.

La materia de simulación que se imparte en la carrera de Actuaría sirvió como inspiración para emprender una investigación detallada de sus aplicaciones en las finanzas, con la intención de mostrar que es un recurso que tendrá que ser estudiado con mayor profundidad por las futuras generaciones de actuarios ya que existe la necesidad creciente de estudiar problemas por este medio.

Los modelos que se simulan en este texto están basados en soluciones de ecuaciones diferenciales de procesos estocásticos y gran parte de la dificultad de llevar a cabo una simulación es determinar y construir el modelo desde las bases teóricas con el fin de lograr mejores resultados.

Al simular procesos estocásticos se obtienen beneficios en las mediciones de flujos de efectivo y de rendimientos como es el caso de la obtención del precio de opciones por este medio y la medición de las reservas necesarias para hacer frente a futuros eventos adversos.

Introducción

El objetivo de este texto es mostrar la teoría básica para la simulación de procesos estocásticos y series de tiempo. Después de exponer los métodos necesarios para la simulación, se procederá a aplicarlos para determinar el tiempo en que un microempresario podría verse en medio de una situación alimentaria precaria de acuerdo con sus ingresos y el precio de la canasta básica rural. La hipótesis a probar con el ejemplo es que no necesariamente un humano que cuenta con un ingreso mensual mayor al de la canasta alimentaria rural se encuentra fuera del segmento de población que pasa hambre ya que la volatilidad de sus ingresos diarios afecta su consumo de alimentos diarios.

Este trabajo no se basa en una investigación exhaustiva sobre el ejemplo. La motivación del mismo es obtener un texto que sirva como guía de consulta de algunos métodos de evaluación financiera en el que se describan con sencillez aquellos que actualmente se utilizan para la simulación por computadora, así como las proposiciones asociadas y sus respectivas demostraciones con el fin de que el lector entienda el razonamiento lógico de cada concepto.

La tesina se presenta en cuatro partes: fundamentos, métodos, aplicaciones y discusión de los resultados.

En el Capítulo 1 se presenta la teoría básica sobre funciones de distribución y convergencia, necesarias para poder conceptualizar una serie de proposiciones. Se presenta también una breve discusión sobre el significado de algunas definiciones y teoremas para procurar una mejor comprensión de los métodos y herramientas de simulación que se presentarán. Los procesos estocásticos presentados son suficientes para realizar una simulación de precios y de flujos de inversión. Se demostrará el Lema de Ito y se empleará el movimiento geométrico Browniano para simular precios, por lo que se presentan sus elementos teóricos.

En el Capítulo 2 se aborda la estimación de parámetros de diferentes modelos estocásticos y la generación de números aleatorios. Asimismo, se presenta y demuestra el método de Cholesky para simular variables aleatorias correlacionadas.

En el Capítulo 3 se presenta la aplicación de los métodos mencionados y los programas en R para su implementación. El ejemplo a simular es el de la capacidad de compra de la canasta alimentaria rural de un microempresario que cuenta con ingresos diarios irregulares. El ejemplo se seleccionó debido a que cuenta con elementos tanto de operación como de precios que pueden ser simulados con un movimiento geométrico Browniano.

En el Capítulo 4 se discuten los resultados hallados y, finalmente, se presentan las conclusiones.

Índice general

Introducción	XI
1. Procesos estocásticos	1
1.1. Funciones de distribución	1
1.2. Procesos estocásticos	11
1.3. Conceptos de convergencia	15
1.4. Movimiento Browniano	19
1.5. Procesos estocásticos generalizados	21
1.5.1. Funciones generalizadas	21
1.5.2. Ruido Blanco	25
1.6. Integral estocástica de Ito	27
2. Simulación de modelos estocásticos	35
2.1. Estimación de parámetros	35
2.1.1. Método de los momentos	36
2.1.2. Método de Bayes	36
2.1.3. Máxima verosimilitud	37
2.2. Evaluación del modelo	41
2.3. Obtención de números pseudoaleatorios	42
2.4. Números aleatorios con una distribución específica	46
2.5. Generación de números aleatorios dependientes	50
2.6. Simulación con R	54
2.6.1. Aspectos generales del lenguaje R	54
2.6.2. Herramientas de simulación de R	55
3. Aplicación a los modelos de precios	59
3.1. Medidas de riesgo	59
3.2. Movimiento geométrico Browniano	62
3.3. Modelos de salto	75
3.4. Modelos de volatilidad	77
3.5. Ejemplo de aplicación de los modelos	79

3.5.1.	Descripción de la variable	81
3.5.2.	Estimación de los parámetros y verificación de los modelos	85
3.5.3.	Simulación y evaluación del riesgo de las variables	87
4.	Discusión sobre las aplicaciones	93
4.1.	Procesos estocásticos	93
4.2.	Métodos de simulación	95
4.3.	Simulación de procesos estocásticos	97
4.4.	La pobreza alimentaria y la simulación	99
4.4.1.	Causas de la pobreza alimentaria	100
4.4.2.	La medición de la pobreza en México	101
5.	Conclusiones	105
A.	Pruebas estadísticas	109
A.1.	Prueba de normalidad <i>Jarque – Bera</i>	109
A.2.	Prueba de raíz unitaria de <i>Dickey – Fuller</i> para un modelo $X_t - X_{t-1} = a_1 X_{t-1} + e_t$	109
B.	Programas en R	111
B.1.	Movimiento geométrico Browniano	111
B.2.	Simulación de los flujos de un microempresario	111
B.3.	Cálculo de los parámetros de un modelo AR(12)	115
B.4.	Código de la prueba <i>Jarque – Bera</i>	118
B.5.	Código de la prueba <i>Dickey – Fuller</i>	119

Capítulo 1

Procesos estocásticos

En este capítulo se presentarán modelos estocásticos de precios y las bases teóricas detrás de ellos. Primero se exponen conceptos propios de las funciones de distribución así como las relaciones entre las variables. Enseguida se presenta una introducción a los procesos estocásticos y se explica cómo se crea un modelo de precios a partir de las bases teóricas de éstos.

1.1. Funciones de distribución

En esta parte se presentan los conceptos relativos al estudio de las distribuciones de las variables aleatorias que se usarán más adelante. El análisis de las distribuciones de probabilidad es de gran importancia para el estudio del riesgo ya que muestra el comportamiento probable de activos en un punto específico del tiempo.

Cabe destacar que los eventos que se proponen como posibles resultados de un experimento o un fenómeno natural deben interpretarse como subconjuntos de espacios en los que existe alguna métrica con el fin de que se puedan evaluar funciones sobre de ellos¹. A continuación se presenta la definición de *evento*:

Definición 1.1. *Una colección \mathcal{F} de subconjuntos de un conjunto Ω es una σ -álgebra si cumple las siguientes condiciones:*

1. $\Omega \in \mathcal{F}$,
2. Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$,
3. Si $A_1, A_2, \dots \in \mathcal{F}$, entonces

¹En este trabajo no se discutirá sobre otros posibles espacios a los que se les pueda adaptar una teoría de probabilidad.

$$\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}.$$

A los elementos de una σ -álgebra \mathcal{F} se les suele llamar \mathcal{F} -medibles. A la intersección de todas las σ -álgebras \mathcal{F} tales que $A \subseteq \mathcal{F}$, se le llama σ -álgebra generada por A y en el caso de que A sea la colección de todos los intervalos cerrados en \mathbb{R} , entonces a la σ -álgebra generada por A se le llamará σ -álgebra de Borel y algún subconjunto de ésta es llamado *conjunto Boreliano*.

Definición 1.2. Una variable aleatoria es una función $X : \Omega \rightarrow \mathbb{R}$ tal que para cualquier conjunto Boreliano B , se cumple que el conjunto $X^{-1}(B)$ es un elemento de \mathcal{F} .

A la pareja (Ω, \mathcal{F}) se le llama *espacio medible* y X^{-1} es la función inversa de X . No se abundará en las propiedades de las variables aleatorias ni en las características de la definición.

Definición 1.3. Sea (\mathcal{F}, Ω) un espacio medible. Una función P definida en Ω es llamada función de probabilidad si cumple con:

1. $P[\Omega] = 1$,
2. $P[A] \geq 0$ si $A \in \mathcal{F}$,
3. Para $A_1, A_2, \dots \in \mathcal{F}$ tales que $A_i \cap A_j = \emptyset$ para $i, j \in \{1, 2, \dots\}$ e $i \neq j$, se tiene que

$$P\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} P[A_k].$$

Ahora se completa el espacio medible y se obtiene el espacio (\mathcal{F}, Ω, P) en el que se tratarán la mayoría de las definiciones expresadas en este texto.

La función de distribución es de gran utilidad para encontrar las probabilidades de eventos que producen las variables aleatorias y se define a continuación:

Definición 1.4. Cualquier función $F_X(\cdot)$ con dominio en la línea real y rango en el intervalo $[0, 1]$ que cumpla las siguientes propiedades:

1. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ y $\lim_{x \rightarrow \infty} F_X(x) = 1$;

2. $F_X(x)$ es una función monótona no decreciente;
3. $F_X(x)$ es continua por la derecha, esto es,

$$\lim_{0 < h, h \rightarrow 0} F_X(x + h) = F_X(x);$$

es una función de distribución.

Existen dos tipos de variables aleatorias por su naturaleza:

Definición 1.5. Se dice que una variable aleatoria X es discreta si el rango de X es contable. Se dice que una variable aleatoria X es continua si existe una función $f_X(\cdot)$ tal que $F_X(x) = \int_{-\infty}^x f_X(u)du$ para todo número real x , en donde $F_X(x)$ es función de distribución de X .

Ahora se definen las funciones de densidad discretas y continuas:

Definición 1.6. Una función $f(x)$ con dominio la línea real y rango el intervalo $[0, 1]$ es llamada función masa de probabilidad si para algún conjunto contable $x_1, x_2, \dots, x_n, \dots$:

1. $f(x_j) \geq 0$ para $j = 1, 2, \dots$;
2. $f(x) = 0$ para $x \neq x_j$ y $j = 1, 2, \dots$;
3. $\sum f(x_j) = 1$ para $x_1, x_2, \dots, x_n, \dots$

Si X es una variable aleatoria continua la función $f_X(\cdot)$ en $F_X(x) = \int_{-\infty}^x f_X(u)du$ es llamada función de densidad de probabilidad de X .

La relación entre las funciones de distribución y de densidad (o de masa de probabilidad para el caso de las distribuciones discretas) es que la función de distribución acumula las probabilidades que generan las funciones de densidad.

En ocasiones es necesario el estudio de más de una variable aleatoria que interactúa con otras en un mismo proceso al mismo tiempo, por lo que a continuación se presenta el siguiente concepto:

Definición 1.7. Sean X_1, X_2, \dots, X_k k -variables aleatorias, todas definidas en el mismo espacio de probabilidad. La función de distribución conjunta de X_1, X_2, \dots, X_k , denotada por $F_{X_1, X_2, \dots, X_k}(\cdot, \dots, \cdot)$, es definida como:

$$P[X_1 \leq x_1, X_2 \leq x_2, \dots, X_k \leq x_k] \text{ para todo } (x_1, x_2, \dots, x_k) .$$

Al vector (x_1, x_2, \dots, x_k) se le llamará variable aleatoria k -dimensional. Para variables aleatorias k -dimensionales existen funciones de densidad, las cuales serán denotadas como:

$f_{X_1, \dots, X_k}(x_1, \dots, x_k)$ en donde para el caso discreto se tiene que

$$f_{X_1, \dots, X_k}(x_1, \dots, x_k) = P[X_1 = x_1, \dots, X_k = x_k].$$

Una de las características importantes de las variables aleatorias es que en algunos casos se puede obtener un valor esperado y una variación esperada del valor de la variable:

Definición 1.8. Sea (X_1, \dots, X_k) una variable aleatoria k -dimensional con densidad $f_{X_1, \dots, X_k}(\cdot, \dots, \cdot)$, el valor esperado de una función $g(\cdot, \dots, \cdot)$ de la variable aleatoria k -dimensional, denotada como $E[g(X_1, \dots, X_k)]$ se define de la siguiente forma:

$$E[g(X_1, \dots, X_k)] = \sum g(x_1, \dots, x_k) f_{X_1, \dots, X_k}(x_1, \dots, x_k),$$

si la variable aleatoria es discreta (la suma es sobre todos los posibles valores de (X_1, \dots, X_k)); y

$$E[g(X_1, \dots, X_k)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(x_1, \dots, x_k) f_{X_1, \dots, X_k}(x_1, \dots, x_k) dx_1 \dots dx_k$$

si la variable aleatoria (X_1, \dots, X_k) es continua.

El cambio esperado al cuadrado de una variable aleatoria es denominado varianza y se define de la siguiente forma:

Definición 1.9. La varianza de una variable aleatoria se define como:

$$\text{Var}[X] = E[(X - E[X])^2].$$

Se puede demostrar que $\text{Var}[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$.

Uno de los conceptos que es muy empleado en la estadística aplicada es el de *momento*, que sirve para encontrar parámetros de algunas distribuciones. Si bien los momentos no siempre definen a una función de distribución única, pueden dar información de cómo es el comportamiento de la distribución.

Definición 1.10. A la función $E[x^k]$ se le llama k -ésimo momento de la variable aleatoria X .

Para encontrar los momentos existe un método que hace uso de una función:

Definición 1.11. La siguiente función $m(t) = E[e^{tX}]$ es llamada función generadora de momentos si la esperanza existe para cualquier valor en el intervalo $-h < t < h$ y $h > 0$.

Se llama función generadora de momentos porque $\frac{\partial^r}{\partial t^r} m(0) = E[X^r]$ y a diferencia de los momentos, define a una función de distribución de manera única. Esto quiere decir que cada función de densidad tiene una función generadora de momentos que es única en caso de que se pueda definir la esperanza en el intervalo mencionado.

Las funciones de distribución pueden ser de *cola ligera* o *cola pesada* y son de ayuda para medir riesgos grandes. Si bien un análisis gráfico de la función de densidad permite un mejor estudio del riesgo, en algunas ocasiones éste es complicado. La importancia del estudio de las colas radica en que las probabilidades de una pérdida grande se encuentran generalmente en funciones de distribución de cola pesada². A continuación la definición matemática de cola pesada y cola ligera.

Definición 1.12. Una función de distribución F_X en $(0, \infty)$ tiene una cola ligera si para toda $s > 0$,

$$m(s) < \infty$$

y una cola pesada si

$$m(s) = \infty$$

en donde $m(s)$ es la función generadora de momentos de X .

La interpretación de la definición de la *función de distribución de cola ligera* es que los momentos de la variable X son siempre acotados, mientras que la interpretación de la *función de distribución de cola pesada* es que los

²Se entenderá por *pérdida grande* al evento desfavorable que le sea difícil afrontar a una entidad.

momentos de la variable X no son siempre acotados, por lo que se esperaría que la esperanza y la varianza no estén acotadas permitiendo la existencia de eventos con valores grandes.

Hasta el momento se han descrito diferentes definiciones teóricas, pero para inferir cómo se distribuyen los resultados de un experimento aleatorio o cómo son sus valores esperados, se tienen que definir dos conceptos básicos: *muestra* y *estimador*. Si bien existe una amplia teoría acerca de estos conceptos, no se abundará en sus propiedades.

Definición 1.13. *Una muestra es un subconjunto del total de datos que están bajo estudio.*

Definición 1.14. *Un estimador es una función con dominio en los datos de una muestra y cuyo resultado se aproxima (o estima) al parámetro de la función de densidad que describe a los datos de la muestra.*

Ahora que se definieron los estimadores para calcular la esperanza y la varianza de una distribución, se introducirán aquellas distribuciones que serán de relevancia para los temas a tratar en este texto.

Definición 1.15. *Se dirá que una variable aleatoria tiene distribución normal, o se distribuye normal, si su función de densidad es descrita por:*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[\frac{-(x - \mu)^2}{2\sigma^2} \right] I_{(-\infty, \infty)}(X),$$

en donde I es la función identidad.

En ocasiones se referirá a una variable aleatoria normal X con parámetros μ y σ^2 con la notación $N(\mu, \sigma^2)$. Los parámetros μ y σ^2 resultan ser la esperanza y la varianza, respectivamente, de la variable X . Para calcular sus parámetros existen estimadores insesgados³, en concreto para una muestra de tamaño n :

$$\bar{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

³Un *estimador* $t(\cdot)$ del parámetro A es llamado insesgado si $E[t(x_1, \dots, x_n)] = A$; resulta conveniente usarlo ya que se confía en que su valor esperado no sea un valor lejano al parámetro.

y

$$\bar{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n-1} (x_i - \bar{\mu})^2.$$

Una distribución importante por su semejanza con el comportamiento final de los rendimientos y que está muy relacionada con la distribución normal es la distribución *lognormal* que nace de la siguiente transformación: $Y = \exp(X)$ en donde X es una variable aleatoria normal con parámetros μ y σ .

Definición 1.16. *Se dirá que una variable aleatoria tiene distribución lognormal si su función de densidad es descrita por:*

$$f(x) = \frac{1}{x\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(\ln(x) - \mu)^2\right] I_{(0,\infty)}(X)$$

y un espacio parametral $\sigma > 0$ y $\mu \in \mathbb{R}$.

Esta variable tiene cola pesada y sólo toma valores positivos. Una variable con distribución lognormal con parámetros μ y σ se denotará como $Lognorm(\mu, \sigma)$.

La función de distribución lognormal aparece en algunos de los ejemplos de modelos de rendimientos, por lo que se hará énfasis en el cálculo de sus parámetros para su posterior análisis. El valor esperado de una variable aleatoria y con distribución lognormal está dado por la siguiente fórmula:

$$E[y] = E[\exp(x)] = \exp\left(\mu + \frac{1}{2}\sigma^2\right);$$

y la varianza, por

$$\text{Var}[y] = E[\exp(x)^2] - E[\exp(x)]^2 = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2),$$

en donde X es una variable aleatoria normal con parámetros μ y σ . Si se busca modelar una variable aleatoria con valor esperado α y varianza β , se tienen que despejar μ y σ de las ecuaciones:

Como

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

y

$$\beta = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2),$$

entonces

$$\ln(\alpha) = \mu + \frac{1}{2}\sigma^2,$$

$$\mu = \ln(\alpha) - \frac{1}{2}\sigma^2,$$

$$\begin{aligned} \beta &= \exp\left[2\left(\ln(\alpha) - \frac{1}{2}\sigma^2\right) + 2\sigma^2\right] - \exp\left[2\left(\ln(\alpha) - \frac{1}{2}\sigma^2\right) + \sigma^2\right] \\ &= \exp[2\ln(\alpha) - \sigma^2 + 2\sigma^2] - \exp[2\ln(\alpha) - \sigma^2 + \sigma^2] \\ &= \exp[2\ln(\alpha) + \sigma^2] - \exp[2\ln(\alpha)] \\ &= \exp[2\ln(\alpha)] (\exp[\sigma^2] - 1), \end{aligned}$$

es decir,

$$\frac{\beta}{\exp[2\ln(\alpha)]} + 1 = \exp[\sigma^2]$$

y

$$\sigma = \left[\ln\left(\frac{\beta}{\exp[2\ln(\alpha)]} + 1\right) \right]^{\frac{1}{2}}.$$

Se ha mostrado que es posible obtener los parámetros de una función de distribución lognormal con esperanza α y varianza β .

En ocasiones la representación de un proceso ocurre para el caso discreto por lo que existen distribuciones que son ampliamente usadas para la modelación. Tal es el caso de la *distribución Poisson* que se puede interpretar como la distribución equivalente a la normal pero para el caso discreto.

Definición 1.17. *Se dirá que una variable aleatoria tiene distribución Poisson si su función de densidad es descrita por:*

$$f_X(x) = \frac{\exp[-\lambda]\lambda^x}{x!} I_{\{0,1,\dots\}}(X)$$

con espacio parametral $\lambda > 0$.

Ahora se describirá la función de *densidad conjunta multinormal* la cual describe el comportamiento de varias variables que se distribuyen como una normal.

Definición 1.18. La función de densidad multinormal $f_{X_1, \dots, X_k}(x_1, \dots, x_k)$ correspondiente al vector de parámetros $N(\mu, \Sigma)$ es:

$$f_{X_1, \dots, X_k}(x) = \frac{1}{(2\pi)^{1/2}} \frac{1}{\det(\Sigma)^{1/2}} \exp \left\{ -\frac{1}{2}(x - \mu)^{tr} \Sigma^{-1} (x - \mu) \right\}$$

en donde $x = (x_1, \dots, x_k)$, $\mu = (E[X_1], \dots, E[X_k])$ y la matriz Σ está definida por elementos de la forma $\Sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$.

El siguiente concepto que se mostrará es el de dependencia entre variables aleatorias. En [29] se enfatiza la importancia de contar con una medición de las dependencias entre activos debido a que el análisis de éstas ayuda a crear estrategias para la diversificación y estructura de un portafolio de inversiones. El *coeficiente de correlación de Pearson* fue una herramienta que se usó por mucho tiempo como medida de dependencia, pero se han desarrollado teorías acerca de cómo encontrar dependencias. Algunas de estas nuevas teorías hablan sobre el uso de cópulas, sin embargo, no se abundará en esa teoría. Los métodos presentados en este texto para simular variables aleatorias dependientes sólo hacen uso de los coeficientes de Pearson y Spearman.

Por medida de dependencia se entenderá una función $\tau(X, Y) \rightarrow a$ en donde X y Y son variables aleatorias y $a \in \mathbb{R}$. Se dirá que existe más dependencia entre A y B que entre C y D si $\tau(C, D) < \tau(A, B)$.

La primera medida de dependencia que se presentará es el coeficiente de correlación de Pearson:

Definición 1.19. El coeficiente de correlación de Pearson entre dos variables es la función:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

en donde $\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$ y σ_X^2, σ_Y^2 son las varianzas de X y Y , respectivamente.

Este coeficiente, desde su postulación en 1885, se ha usado como medida de dependencia. Incluso Markowitz propone al coeficiente de correlación como medida de riesgo, pero se ha visto que en algunos casos no produce buenos resultados, pues toma en cuenta valores extremos⁴, por lo que una forma de evitar el problema es usar el *coeficiente de correlación de Spearman* definido a continuación:

Definición 1.20. Sean x_1, x_2, \dots, x_n y y_1, y_2, \dots, y_n dos muestras de dos variables tales que $x_i < x_{i+1}$ donde a cada x_i le corresponde únicamente el elemento y_i . Sea entonces $y_{j_k} = y_j$ en donde $y_{j_k} < y_{j_{k+1}}$ para alguna $j \in \{1, \dots, n\}$. El coeficiente de correlación de Spearman de dos muestras está dado por:

$$\rho_s = 1 - \frac{6 \sum_{l=1}^n (l - k)_l^2}{n(n^2 - 1)}$$

en donde k es tal que $y_l = y_{l_k}$.

Ahora se definirá una medida de dependencia que puede ser usada para cualquier tipo de distribución conjunta de variables continuas.

Supóngase que se tienen dos observaciones (x', y') y (x'', y'') de un vector (X, Y) de variables aleatorias continuas. Estas observaciones son llamadas dependientes positivas si:

$$x' > x'' \text{ , entonces } y' > y'' ;$$

y dependientes negativas si

$$x' > x'' \text{ , entonces } y' < y'' .$$

Consideremos n observaciones del vector (X, Y) : $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Si denotamos como c al número de combinaciones entre estos pares que son dependientes positivos y a d como el número de dependientes negativos, entonces se puede dar la siguiente medida de dependencia:

Definición 1.21. La medida de dependencia de Kendall es:

$$\tau = \frac{c - d}{c + d}.$$

Esta medida de dependencia es la probabilidad del número de pares dependientes positivos menos los dependientes negativos.

Ahora se definirá el concepto de matriz de correlaciones que se usará para la simulación y análisis de variables dependientes:

⁴Que afectan la estimación del coeficiente sobre una muestra.

Definición 1.22. Una matriz de dimensión $n \times n$ cuyos elementos a_{ij} , $i, j \in \{1, 2, \dots, n\}$ son de la forma $a_{ij} = \rho_{ij} = \rho_{ji}$ donde ρ_{ij} es alguna medida de dependencia entre las variables i y j , es llamada matriz de correlaciones.

Este tipo de matrices serán útiles para definir parámetros multidimensionales.

1.2. Procesos estocásticos

Uno de los objetivos de la estadística es encontrar la distribución del resultado de un fenómeno aleatorio por lo que a lo largo de su historia se idearon métodos para esbozar la distribución buscada. Inicialmente, estos métodos procuraban cálculos sencillos ya que no se contaba con computadoras que los hicieran en grandes cantidades, lo cual ahora es posible hasta cierto punto.

En el quehacer cotidiano del ser humano de nuestra era se ha requerido la comprensión de fenómenos que varían a través del tiempo pues la toma de decisiones antes de que un evento suceda es de vital importancia para evitar pérdidas catastróficas en un futuro. Ya sea que se estudie la trayectoria de un ciclón o las reservas monetarias de un país, conocer el valor de estas variables a lo largo del tiempo puede resultar de interés para muchas personas. Los procesos estocásticos proporcionan formas de representar y predecir (con una cierta precisión) estos y otros fenómenos.

A continuación se darán definiciones y teoremas sobre procesos estocásticos que se han retomado de [16]. Se requiere la revisión de dicho material para profundizar en el tema.

Definición 1.23. Un proceso estocástico se define como cualquier colección de variables aleatorias X_t , $t \in T$ definidas en un espacio de probabilidad común en donde T es un subconjunto de $(-\infty, \infty)$.

El conjunto T puede ser pensado como un intervalo de tiempo. En caso de que T sea un intervalo, entonces se dirá que se trata de un proceso estocástico continuo y si T es un subconjunto de los enteros, entonces se dirá que el proceso estocástico es discreto.

Uno de los procesos estocásticos más relevantes para simular y cuantificar el riesgo es la caminata aleatoria.

Definición 1.24. Sean ξ_1, ξ_2, \dots variables aleatorias independientes que toman valores enteros, teniendo una función de densidad común f . Sea X_0 una variable aleatoria que toma valores enteros, independiente de ξ_1, ξ_2, \dots y sea la suma $X_n = X_0 + \xi_1 + \xi_2 + \dots + \xi_n$. Entonces la suma $X_n, n \geq 0$, es llamada caminata aleatoria.

Una *serie de tiempo* es un proceso estocástico X_t discreto en donde T es el tiempo y en adelante se entenderá por una *realización de una serie de tiempo* a un conjunto de valores que puedan tomar las variables del proceso X_t .

Existen dos tipos de procesos estocásticos: *estrictamente estacionarios* y *estacionarios*. La característica de las series de tiempo *estrictamente estacionarios* es la siguiente:

$$f(Z_t, Z_{t+1}, \dots, Z_{t+m}) = f(Z_{t+k}, Z_{t+k+1}, \dots, Z_{t+k+m}) \text{ para todo } t, m \text{ y } k.$$

Esta característica implica que cada variable del proceso estocástico tiene la misma distribución. En algunas ocasiones existe dificultad para hallar la función de distribución de cada variable Z_t y la distribución conjunta de algunas o todas las variables, por lo que se ha optado por buscar otra forma que no requiera de conocer toda la función de distribución de cada variable aleatoria, sino sólo sus primeros momentos. En caso de que la distribución sea conjunta, sólo se toma el coeficiente de correlación entre variables. Si bien estos elementos no definen una función de distribución, en algunos casos serán la única opción de la cual echar mano. Se dirá entonces que una serie de tiempo es *estacionaria* si las siguientes condiciones se cumplen:

1. $E(Z_t) = E(Z_{t+k})$,
2. $\gamma_k = \text{Cov}(Z_t, Z_{t+k}) = \text{Cov}(Z_{t+h}, Z_{t+h+k})$,

para cualesquiera enteros k y h . Para evitar la influencia de las unidades de medida de las autocovarianzas se debe usar un indicador como el del coeficiente de correlación pero para las series de tiempo. Éste es llamado función de autocorrelación y es definido de la siguiente forma:

$$\rho_k = \frac{\gamma_k}{\gamma_0}.$$

Si el proceso tiene una función de distribución conjunta multinormal para todas sus variables Z_t y es *estacionario*, entonces es *estrictamente estacionario* dado que la función de distribución multinormal queda definida sólo por μ y Σ .

La siguiente propiedad con que cuentan algunos de los procesos estocásticos es de gran importancia en el estudio de procesos justos en el sentido de que no importan los eventos ocurridos anteriormente en el proceso; no se modifica la probabilidad de ocurrencia de cada uno de los eventos en el futuro. Primero, se definirá lo que se entiende por eventos ocurridos:

Definición 1.25. Una filtración (en (Ω, \mathcal{F})) es una familia M_t con $t \geq 0$ de σ -álgebras tales que:

1. $M_t \subset \mathcal{F}$,
2. $0 \leq s < t \Rightarrow M_s \subset M_t$.

Ahora se definen los procesos comentados anteriormente:

Definición 1.26. Un proceso estocástico Z_t en (Ω, \mathcal{F}, P) es llamado martingala con respecto a una filtración M_t si el proceso cumple con las siguientes condiciones:

1. Z_t es M_t -medible para todo t ,
2. $E[|Z_t|] < \infty$ para todo t y
3. $E[Z_s | M_t] = Z_t$ para todo $s \geq t$.

Más adelante se retomarán estos conceptos y cobrarán importancia gracias a que los modelos de precios necesitan de las martingalas para frenar el arbitraje.

La metodología usada en este texto para construir un modelo y simularlo sigue cuatro pasos y fue presentada por Box y Jenkins en 1960 y en [13] se explica a detalle cómo funciona. Los pasos de esta metodología siguen una secuencia lógica. Sin embargo, en algunos casos pueden agregarse o quitarse pasos:

1. *Identificación.* Identificar un modelo apropiado para la serie de tiempo de acuerdo con sus características de estacionariedad y distribución de cada Z_t .

2. *Estimación.* En este paso se deben encontrar parámetros que permitan al modelo representar con la mayor precisión posible al fenómeno.

3. *Verificación.* Que el modelo cumpla con el objetivo y represente de la forma deseada al fenómeno.

4. *Uso del modelo.* Usar el modelo para la simulación, obtener resultados y analizarlos.

Más adelante, en una aplicación práctica, se retomará esta metodología. Ahora se presentarán algunos modelos de series de tiempo que servirán como antecedentes al modelo de Ito.

Definición 1.27. *El modelo $AR(p)$ (Auto regresivo “p”) es de la siguiente forma:*

$$Z_t = (1 - \alpha_1 - \alpha_2 - \dots - \alpha_p)\mu + \alpha_1 Z_{t-1} + \alpha_2 Z_{t-2} + \dots + \alpha_p Z_{t-p} + \epsilon_t$$

en donde α_t , $t \in \{1, 2, \dots\}$ son constantes; ϵ_t , una variable aleatoria independiente de ϵ_{t+m} con $m \neq t$; y μ , una constante (llamada usualmente nivel).

Este modelo está pensado para series de tiempo cuyos valores están influenciados por eventos pasados de la misma serie. Esto quiere decir que el valor $Z(t)$ de la serie de tiempo depende de valores pasados, lo cual es conveniente considerar en modelos de precios que tardan en estabilizarse después de bajas o subidas abruptas. La interpretación del modelo es que la diferencia entre Z_t y una constante μ depende de una variable aleatoria y de diferencias entre valores pasados de la serie y la constante μ :

$$Z_t - \mu = \alpha_1(Z_{t-1} - \mu) + \alpha_2(Z_{t-2} - \mu) + \dots + \alpha_p(Z_{t-p} - \mu) + \epsilon_t .$$

Para calcular los parámetros existen métodos de optimización que permiten escoger el modelo con el mínimo error. Un método usado es el de *mínimos cuadrados* dado que el modelo puede verse como una regresión en donde las variables independientes son valores pasados de la serie de tiempo.

Otro modelo usado ampliamente es el siguiente:

Definición 1.28. *El modelo $MA(q)$ (Promedios móviles “q”) es de la siguiente forma:*

$$Z_t = \mu + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \dots + \alpha_q \epsilon_{t-q} + \epsilon_t$$

en donde α_t , con $t \in \{1, 2, \dots, q\}$, son constantes; ϵ_t , una sucesión de variables aleatorias $N(0, \sigma^2)$ independientes entre ellas; y μ , una constante.

Este modelo ayudará a describir el comportamiento de un precio que depende de una variable aleatoria en el que existe gran dependencia entre valores pasados de ésta. Una característica de este modelo es que es estacionario ya que la γ_k no depende del tiempo.

Ahora se dará una generalización que contempla los dos modelos expuestos anteriormente.

Definición 1.29. *El modelo ARMA(p, q) (Autoregresivo de promedios móviles) es de la siguiente forma:*

$$Z_t - \mu - (\alpha_1(Z_{t-1} - \mu) + \alpha_2(Z_{t-2} - \mu) + \dots + \alpha_p(Z_{t-p} - \mu)) = a_1\epsilon_{t-1} + a_2\epsilon_{t-2} + \dots + a_q\epsilon_{t-q} + \epsilon_t$$

en donde α_t , con $t \in \{1, 2, \dots, p\}$, son constantes; ϵ_t , una sucesión de variables aleatorias idénticamente distribuidas e independientes entre ellas, a_t con $t \in \{1, 2, \dots, q\}$, constantes; y μ , una constante.

Una de las herramientas más importantes en el análisis de las series de tiempo es la aplicación de operadores que transforman la serie original en series diferentes que pueden tener cualidades diferentes. Por ejemplo, al restar el elemento Z_{t-1} al elemento Z_t se obtiene otra variable que puede tener mayor comportamiento estacional⁵. La transformación que se usará con mayor frecuencia en este trabajo es la de diferencias:

$$\begin{aligned}\nabla^0 Z_t &= Z_t \\ \nabla^1 Z_t &= Z_t - Z_{t-1} \\ \nabla^2 Z_t &= [(Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2})] \\ \nabla^3 Z_t &= \nabla (\nabla^2 Z_t) \\ &\vdots \\ \nabla^n Z_t &= \nabla (\nabla^{n-1} Z_t).\end{aligned}$$

Uno de los usos de las series de tiempo es hacer pronósticos, para lo cual se toma la esperanza de la variable Z_t (el pronóstico para el tiempo t). Es importante destacar que la simulación de estos modelos (y los subsecuentes) brinda un pronóstico que se encuentra dentro de un intervalo de confianza. Los modelos mostrados hasta ahora se presentan sólo como complemento e introducción a las series de tiempo y no se profundizará en ellos.

1.3. Conceptos de convergencia

Los siguientes conceptos tienen como objetivo estudiar a la convergencia de series de variables aleatorias. Las diferencias y relaciones que existen entre una y otra convergencia serán útiles para llegar a conclusiones más adelante.

⁵Existe una demostración que afirma que al aplicar las diferencias a la serie un número de veces, la serie tiende a ser estacionaria. Una demostración sobre este hecho, se puede ver en [13].

Definición 1.30. La sucesión de variables aleatorias X_1, X_2, \dots, X_n converge puntualmente a X si para cada $w \in \Omega$,

$$\lim_{n \rightarrow \infty} X_n(w) = X(w)$$

y se denotará como *ac* – $\lim_{n \rightarrow \infty} X_n = X$.

Se puede decir que esta convergencia, llamada también *convergencia con probabilidad 1*, es la que se usa usualmente en los cursos de cálculo. Este límite es estricto en el sentido de que pide la convergencia de la sucesión evaluada en cada uno de los elementos de Ω . Se pueden definir otro tipo de convergencias como la siguiente:

Definición 1.31. La sucesión de variables aleatorias X_1, X_2, \dots, X_n converge en probabilidad a X si para cada $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} P\{w \in \Omega : |X_n(w) - X(w)| > \epsilon\} \rightarrow 0$$

y se denotará como *st* – $\lim_{n \rightarrow \infty} X_n = X$.

El siguiente tipo de convergencia se usa para mostrar que el valor esperado de una sucesión X_n converge a un valor X . Se pueden usar potencias mayores que uno, pero para este texto sólo se necesitará la potencia dos.

Definición 1.32. La sucesión de variables aleatorias X_1, X_2, \dots, X_n converge en media cuadrática a X si,

$$\lim_{n \rightarrow \infty} E|X_n - X|^2 = 0$$

y se denotará como *qm* – $\lim_{n \rightarrow \infty} X_n = X$.

Este tipo de convergencia se usa para demostrar que el valor esperado de una sucesión X_n converge a X y este aspecto será de ayuda para las siguientes secciones. Se puede ver que esta convergencia implica que deben existir los segundos momentos de X_n y X . Más adelante se usará una relación entre este tipo de convergencias, a saber que la convergencia con probabilidad 1 implica la convergencia en probabilidad.

Proposición 1.1. La convergencia puntual implica la convergencia en probabilidad.

Demostración. Sea $\epsilon > 0$ fija. Para cada número natural n defínase la sucesión de eventos siguiente:

$$B_n = \bigcup_{m=n}^{\infty} (|X_m - X| > \epsilon).$$

De esta sucesión se tiene que

$$(|X_n - X| > \epsilon) \subseteq B_n \text{ implica que } P(|X_n - X| > \epsilon) \leq P(B_n)$$

por lo que

$$\begin{aligned} \lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) &\leq \lim_{n \rightarrow \infty} P(B_n) \\ &= P\left(\lim_{n \rightarrow \infty} B_n\right) \\ &= P\left(\bigcap_{n=1}^{\infty} B_n\right). \end{aligned}$$

Puesto que B_n es una sucesión decreciente,

$$P\left(\bigcap_{n=1}^{\infty} B_n\right) = P(|X_n - X| > \epsilon)$$

para toda $n \geq 1$ por lo que

$$\begin{aligned} P\left(\bigcap_{n=1}^{\infty} B_n\right) &= \lim_{n \rightarrow \infty} P(X_n \neq X) \\ &= 0. \end{aligned}$$

□

Es importante notar que el recíproco de la proposición no se cumple para todos los casos.

La siguiente definición establece la relación entre la convergencia en media cuadrática y la convergencia en probabilidad y se empleará en demostraciones posteriores.

Definición 1.33. *La sucesión de variables aleatorias X_1, X_2, \dots, X_n converge en media a X si,*

$$\lim_{n \rightarrow \infty} E |X_n - X| = 0$$

y se denotará como $m\text{-}\lim_{n \rightarrow \infty} X_n = X$.

La siguiente relación entre convergencias se usará más adelante y conectará algunas de las ideas de convergencia.

Proposición 1.2. *La convergencia en media cuadrática implica la convergencia en media.*

Demostración. Usando la desigualdad de Jensen⁶:

$$\begin{aligned} E(|X_n - X|) &\leq E(|X_n - X|)^2 \\ &\leq E(|X_n - X|^2). \end{aligned}$$

□

El recíproco de la proposición tampoco se cumple para todos los casos.

La siguiente proposición une los conceptos de convergencia en media cuadrática y de convergencia en probabilidad:

Proposición 1.3. *La convergencia en media implica la convergencia en probabilidad.*

Demostración. Para cada $\epsilon > 0$ defínase el evento $A_n = \{|X_n - X| > \epsilon\}$, entonces

$$\begin{aligned} E[|X_n - X|] &= E[|X_n - X| I_{A_n}] + E[|X_n - X| I_{A_n^c}] \\ &\geq E[|X_n - X| I_{A_n}] \\ &\geq \epsilon P(|X_n - X| > \epsilon). \end{aligned}$$

Por la desigualdad de Markov⁷ y como $E[|X_n - X|] \rightarrow 0$ entonces

$$P(|X_n - X| > \epsilon) \rightarrow 0.$$

□

⁶Para una función convexa $U(x)$ se cumple la siguiente desigualdad: $U(E(x)) \leq E(U(x))$.

⁷Para una variable aleatoria X y una constante a se cumple la siguiente desigualdad: $P(X \geq a) \leq E(X) \frac{1}{a}$.

1.4. Movimiento Browniano

Se ha observado que una partícula que se encuentra dentro de un fluido está en un constante movimiento irregular. Se ha supuesto que este movimiento es causado por el bombardeo de partículas invisibles. Este movimiento es llamado *movimiento Browniano*, en honor a uno de los primeros científicos que lo estudiaron. Un modelo que describe tal movimiento se basa en un sistema cartesiano de coordenadas cuyo origen es la posición de la partícula en el tiempo $t = 0$. Entonces las tres coordenadas de la partícula varían independientemente, cada una de ellas de acuerdo con un proceso estocástico W_t que satisface las siguientes propiedades:

Definición 1.34. *Sea W_t un proceso estocástico que cumple con las siguientes propiedades:*

1. $W_0 = 0$.
2. Para toda $t > s$ se tiene que $W_t - W_s$ es una variable aleatoria con función de distribución normal con media 0 y varianza $t - s$.
3. Para cualesquiera $t_1 < t_2 < t_3 < t_4$, las variables $W_{t_2} - W_{t_1}$ y $W_{t_4} - W_{t_3}$ son independientes.

entonces a W_t se le denomina proceso de Wiener.

La propiedad 1 se deriva de que el modelo se basa en un sistema cartesiano de coordenadas. Las propiedades 2 y 3 se basan en que el movimiento de la partícula depende de grandes cantidades de colisiones que no tienen tendencia en mover la partícula en una dirección o en la dirección contraria. Si bien la interpretación del proceso de Wiener que se ha dado es física, se muestra en este texto para dar una idea de cómo fue concebido. Una interpretación usando dinámica de precios radica en que un precio que cambia constantemente puede verse en un plano cartesiano, además los cambios de precio dependen de grandes cantidades de factores sociales, económicos, naturales, etc.

De las propiedades del proceso de Wiener, se sigue que

$$E[W_t] = E[W_t - 0] = E[W_t - W_0] = 0$$

y

$$\text{Var}[W_t] = E[W_t^2] = t,$$

y para $t > s$

$$\begin{aligned} \text{Cov}(W_s, W_t) &= \text{Cov}(W_s, W_s + W_t - W_s) \\ &= \text{Cov}(W_s, W_s + W_t - W_s) \\ &= \text{Cov}(W_s, W_s) + \text{Cov}(W_s, W_t - W_s) \\ &= s. \end{aligned}$$

Se sigue que

$$\text{Cov}(W_s, W_t) = \min\{s, t\},$$

por lo que su función de covarianza queda determinada por t y s . Una de las propiedades más importantes sobre el proceso de Wiener es que se trata de una martingala, hecho que se demostrará a continuación:

Proposición 1.4. *Un proceso de Wiener W_t es una martingala con respecto a la filtración M_s con $s \geq t$.*

Demostración. La esperanza es acotada para toda t como se mostró anteriormente, por lo que sólo queda la siguiente expresión:

$$\begin{aligned} E[W_s | M_t] &\geq E[W_s + W_t - W_t | M_t] \\ &= E[W_s - W_t | M_t] + E[W_t | M_t] \\ &= W_t \end{aligned}$$

ya que $W_s - W_t$ es independiente de M_t y W_t es \mathcal{F} -medible y pertenece a M_t . □

Uno de los aspectos más importantes con que el proceso de Wiener cuenta es que no es diferenciable ya que

$$\begin{aligned} \lim_{h \rightarrow 0} E \left[\left(\frac{W_{t+h} - W_t}{h} \right)^2 \right] &= \lim_{h \rightarrow 0} \frac{\text{Var}[W_{t+h} - W_t]}{h^2} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} = \infty. \end{aligned}$$

Ese inconveniente se resolverá en la sección siguiente.

1.5. Procesos estocásticos generalizados

En esta sección se dará una definición más general del concepto de proceso estocástico que será de utilidad para tratar temas como las ecuaciones estocásticas diferenciales.

1.5.1. Funciones generalizadas

Comenzaremos por considerar una generalización del concepto clásico de función, a saber, el concepto *función generalizada* que ha surgido en el proceso de resolución de ciertos problemas físicos (en [20] se explican más detalles) y que permite formular en el lenguaje matemático nociones idealizadas como, por ejemplo, densidad de una carga puntual, impulso instantáneo, etc.

Por medio de un ejemplo de fuerza instantánea se mostrará la necesidad y existencia de las funciones generalizadas.

Supóngase que a un cuerpo con masa $m \neq 0$ se le aplica una fuerza $F(t)$ que le comunica una velocidad $v \neq 0$ en el tiempo $t = 0$; después del tiempo $t = 0$, la fuerza termina. La segunda ley de Newton, propone que la fuerza en t queda determinada por la siguiente ecuación diferencial

$$F(t) = \frac{d(mv)}{dt}$$

por lo que para cualquier momento τ , $0 < \tau < +\infty$ en el tiempo, se tiene que:

$$\int_{-\infty}^{\tau} F(t)dt = mv. \quad (1.1)$$

Esta integral no tiene sentido: $F(t) = 0$ para todo $t \neq 0$, por lo que la integral es igual a cero y al mismo tiempo es igual a una constante mv . Se deben emplear nuevas definiciones de función y de integral con el fin de dar solución a estos problemas.

Supóngase que la cantidad de movimiento adquirido por el cuerpo es $mv = 1$ y que la fuerza que actúa sobre el cuerpo es la función $\delta(t)$ por lo que la nueva ecuación tendrá la siguiente forma:

$$\int_{-\infty}^{\tau} \delta(t)dt = 1. \quad (1.2)$$

La función $\delta(t)$ es llamada *función delta* o de *Dirac*. Supóngase que al cuerpo no se le aplica una fuerza en un instante $t = 0$, sino una fuerza

constante que se le aplica por un periodo de tiempo que va desde $-\epsilon$ hasta 0, con $\epsilon > 0$; a esta fuerza se le denotará con $\delta_\epsilon(t)$. En otras palabras, distribúyase la fuerza $\delta(t)$ a lo largo del intervalo $[0, \epsilon]$. Para hallar $\delta_\epsilon(t)$ se tiene que

$$1 = \int_{-\infty}^{\tau} \delta_\epsilon(t) dt = \int_{-\epsilon}^0 \delta_\epsilon(t) dt = \epsilon \delta_\epsilon(t)$$

con $-\epsilon \leq t \leq 0$, por lo que

$$\delta_\epsilon(t) = \begin{cases} \frac{1}{\epsilon}, & \text{si } t \in \{t : -\epsilon \leq t \leq 0\} \\ 0, & \text{si } t \notin \{t : -\epsilon \leq t \leq 0\} \end{cases} . \quad (1.3)$$

Si se quiere aproximar $\delta(t)$ por medio de $\delta_\epsilon(t)$ de la siguiente forma

$$\lim_{\epsilon \rightarrow 0} \delta_\epsilon(t) = \delta(t)$$

se obtiene

$$\delta(t) = \begin{cases} +\infty, & \text{si } t = 0 \\ 0, & \text{si } t \neq 0. \end{cases}$$

Es claro que de esta forma no se puede obtener (1.2). Pero existe un concepto que permite encontrar una definición, y es el siguiente para $\tau > 0$,

$$\int_{-\infty}^{\tau} \delta(t) dt = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\tau} \delta_\epsilon(t) dt,$$

que es pensado como el límite de la cantidad de movimiento comunicado por las fuerzas $\delta_\epsilon(t)$. Ahora se procede a extender el concepto de integral al siguiente

$$\int_{-\infty}^{\tau} \delta(t) f(t) dt = \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\tau} \delta_\epsilon(t) f(t) dt$$

en donde $f(t)$ es una función continua.

Teorema 1.1. *Sea $f(t)$ una función continua en $t = 0$, entonces*

$$\lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\tau} \delta_\epsilon(t) f(t) dt = \begin{cases} f(0) & \text{para } \tau \geq 0 \\ 0 & \text{para } \tau < 0 \end{cases} .$$

Demostración. Para $t \geq 0$ y haciendo uso de (1.3) se tiene que

$$\begin{aligned}
\left| \int_{-\infty}^{\tau} \delta_{\epsilon}(t) f(t) dt - f(0) \right| &= \left| \frac{1}{\epsilon} \int_{-\epsilon}^0 f(t) dt - f(0) \right| \\
&= \left| \frac{1}{\epsilon} \int_{-\epsilon}^0 f(t) dt - \frac{f(0)}{\epsilon} \int_{-\epsilon}^0 dt \right| \\
&\leq \frac{1}{\epsilon} \int_{-\epsilon}^0 |f(t) - f(0)| dt.
\end{aligned}$$

Como $f(x)$ es continua cuando $x = 0$, entonces para cualquier η existe un $\epsilon_{\eta} > 0$, tal que para toda t que satisface la condición $|t| < \epsilon_{\eta}$, se cumple la siguiente desigualdad

$$|f(t) - f(0)| < \eta,$$

por lo que para cualquier $\epsilon < \epsilon_{\eta}$ se cumple que

$$\left| \int_{-\infty}^{\tau} \delta_{\epsilon}(t) f(t) dt - f(0) \right| < \frac{\eta}{\epsilon} \int_{-\infty}^{\tau} dt = \eta$$

y queda demostrado el caso $t \geq 0$. El caso $t < 0$ se demuestra de la misma forma. □

No se darán todos los detalles y propiedades de las funciones generalizadas y en caso de que el lector se interese en este tema se referirá a [20].

Definamos un espacio lineal de funciones D que será el concepto principal en nuestros razonamientos. Con este fin consideraremos las funciones que vienen dadas en el conjunto de los números reales y que toman valores en los complejos. Todas las funciones finitas, siendo definidas del modo natural las operaciones de suma y multiplicación por un número, forman un espacio lineal, y las funciones finitas infinitamente derivables forman un subespacio del espacio mencionado. Ahora se introducirá en este espacio el concepto de convergencia de sucesiones.

Definición 1.35. *Una sucesión de las funciones finitas infinitamente derivables φ_n , $n=1,2,\dots$, se llama convergente hacia la función finita infinitamente derivable φ si:*

1. *Existe un segmento $[a,b]$, fuera del cual todas las funciones φ_n , $n \in \{1, 2, \dots\}$ se anulan.*

2. En dicho segmento la sucesión de funciones φ_n , $n \in \{1, 2, \dots\}$, y las sucesiones de todas sus derivadas $\varphi_n^{(k)}$, $n \in \{1, 2, \dots\}$, convergen uniformemente hacia la función φ y a sus correspondientes derivadas $\varphi^{(k)}$, $k \in \{1, 2, \dots\}$, respectivamente.

Ahora se procederá a definir el espacio D .

Definición 1.36. Un espacio de funciones finitas infinitamente derivables que convergen en el sentido de la Definición 1.35., lleva el nombre de espacio D de funciones principales.

Evidentemente, si $\varphi \in D$, toda derivada de la función φ pertenecerá también al espacio D .

Definición 1.37. Toda funcional continua lineal f , definida sobre D , recibe el nombre de función generalizada⁸.

Una funcional f en $\varphi \in D$ se denota con (f, φ) y en especial, para alguna función localmente integrable f^9 , se prestará atención a la funcional definida como:

$$(f, \varphi) = \int_{-\infty}^{\infty} f(t)\varphi(t)dt.$$

Esta funcional es una función generalizada y se le llamará en adelante la versión generalizada de f .

Al derivar f , se obtiene (f', φ) e integrando por partes:

$$\begin{aligned} (f', \varphi) &= \int_{-\infty}^{\infty} f'(t)\varphi(t)dt \\ &= - \int_{-\infty}^{\infty} f(t)\varphi'(t)dt \\ &= -(f, \varphi'). \end{aligned}$$

⁸Por funcional lineal se entenderá a toda función de la forma $R^n \rightarrow R^1$ tal que $f(X) = a_1x_1 + a_2x_2 + \dots + a_nx_n$.

⁹Una función es llamada localmente integrable si es absolutamente integrable en cualquier segmento finito.

De manera que

$$(f', \varphi) = -(f, \varphi')$$

que existe dentro del espacio D (esto es, toda función generalizada tiene derivada).

Se procede ahora a definir de manera formal la función delta.

Definición 1.38. Una funcional definida por la fórmula

$$(\delta, \varphi) = \varphi(0)$$

en donde $\varphi \in D$, se denomina función delta y se denota como $\delta(t)$. En caso de que la función sea $\varphi(t_0) \neq 0$ en donde t_0 no es el punto cero, entonces, se usa $\delta(t - t_0)$ para denotar la función delta.

A continuación se trasladará esta teoría al escenario de los procesos estocásticos.

1.5.2. Ruido Blanco

El *ruido blanco* es conocido como un *proceso estocástico Gaussiano* estacionario ξ_t , para $-\infty < t < \infty$, con $E[\xi_t] = 0$ y una función de densidad espectral constante $f(\lambda)$ en todo el eje real¹⁰. Si $E[\xi_s \xi_{s+t}] = C(t)$ es la función de autocovarianzas de ξ_t , entonces, se tiene que si

$$\int_{-\infty}^{\infty} |C(t)| dt < \infty.$$

Entonces, derivado la transformación de *Fourier*, (véase [4]), se tiene que

$$C(t) = \int_{-\infty}^{\infty} \exp(-i\lambda t) f(\lambda) d\lambda,$$

y su función de densidad espectral, $f(\lambda)$, está dada por la *transformación inversa de Fourier*:

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(i\lambda t) C(t) dt = \frac{c}{2\pi},$$

¹⁰Es llamado *ruido blanco* porque el concepto surgió en el estudio de las frecuencias de la luz del sol. Un proceso estocástico Z_t , es llamado *Gaussiano* si cada combinación lineal finita de las variables Z_t tiene una distribución normal. La *función de densidad espectral* es una función sobre un conjunto continuo t y describe (en un sentido físico) el comportamiento de una serie de ondas.

en donde c es una constante. Pero esta ecuación sólo tiene sentido si

$$C(t) = \delta(t)$$

ya que $C(t) = 0$ para toda $t \neq 0$, pues el proceso es *Gaussiano pseudo-estacionario*.

Definición 1.39. Un proceso estocástico generalizado es una función generalizada aleatoria definida de la siguiente forma: para toda $\varphi \in D$ se asigna una variable aleatoria $\Phi(\varphi)$ (esto es, $\Phi(\varphi)$ es un proceso estocástico con espacio parametral D).

Un *proceso estocástico generalizado* es llamado *Gaussiano* si para las funciones arbitrarias, lineales e independientes $\varphi_1, \varphi_2, \dots, \varphi_n \in D$, la variable aleatoria $(\Phi(\varphi_1), \Phi(\varphi_2), \dots, \Phi(\varphi_n))$ se distribuye como una normal.

Ahora se evaluará al proceso de Wiener representado como una *ecuación estocástica generalizada Gaussiana*. La representación es la siguiente:

$$\Phi(\varphi) = \int_{-\infty}^{\infty} \varphi(t) W_t dt$$

con $W_t = 0$ para $t < 0$. Como $E[\Phi(\varphi)] = 0$, se tiene que

$$C(\varphi, \psi) = E[(\Phi(\varphi) - E[\Phi(\varphi)])(\Phi(\psi) - E[\Phi(\psi)])] = E[\Phi(\varphi)\Phi(\psi)]$$

es la función de covarianza del proceso. Entonces la derivada de un proceso de Wiener es un proceso con

$$E[\Phi'(\varphi)] = -E[\Phi(\varphi')] = 0$$

y

$$C'(\varphi, \psi) = E[\Phi(\varphi')\Phi(\psi')] = C(\varphi', \psi').$$

Se tiene que

$$\begin{aligned} C(\varphi, \psi) &= \int_{-\infty}^{\infty} (\Phi(\varphi)\Phi(\psi)) f(t, s) dt ds \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \varphi(t) W_t dt \int_{-\infty}^{\infty} \psi(s) W_s ds \right) f(t, s) dt ds \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_t W_s f(t, s) dt ds \varphi(t) \psi(s) dt ds \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \min(t, s) \varphi(t) \psi(s) dt ds. \end{aligned}$$

Como ψ pertenece al espacio D , entonces es integrable. Sea

$$\int_{-\infty}^{\infty} \psi'(s) ds = \check{\psi},$$

entonces

$$u(t) = t\varphi'(t)\check{\psi},$$

que pertenece a D , por lo que

$$C(\varphi', \psi') = \int_{-\infty}^{\infty} u(t) dt.$$

Pero el proceso de Wiener tiene incrementos independientes, por lo que la derivada de W_t es independiente de la derivada de W_s , lo que significa que su función de correlación es cero en $t \neq s$ y 1 cuando $t = s$. Por esta razón la integral obtenida puede verse como una integral generalizada con función $\delta(t - s)$, que es la función de correlación del *ruido blanco*:

$$C'(\varphi, \psi) = \int_{-\infty}^{\infty} \delta(t - s) u(t) dt.$$

Su media es cero y se distribuye *normal* por la definición de *proceso estocástico generalizado Gaussiano*, esto es, la derivada de un *proceso estocástico generalizado de Wiener* es el *ruido blanco*.

1.6. Integral estocástica de Ito

A continuación se definirá la integral estocástica de Ito, que contempla casos que la integral de Riemann-Stieltjes no se puede manejar.

Se empezará a definir la partición del intervalo $[t_0, T]$ como $t_0 < t_1 < \dots < t_n = T \leq T$, y con esta partición se obtendrá una serie de resultados útiles.

Gracias a la continuidad de W_t se tiene que

$$\text{máx} [|W_{t_k} - W_{t_{k-1}}|] \rightarrow 0$$

cuando $\delta_n = \text{máx}_{1 \leq k \leq n} (t_k - t_{k-1}) \rightarrow 0$.

La siguiente proposición se conoce como *variación cuadrática*:

Proposición 1.5. *La variación cuadrática de una trayectoria del movimiento Browniano sobre el intervalo $[T, t_0]$ es la longitud del intervalo casi seguramente:*

$$\sum_{k=1}^n |W_{t_k} - W_{t_{k-1}}|^2 = T - t_0$$

cuando $\delta_n \rightarrow 0$ casi seguramente.

Demostración. Para la prueba se usará la convergencia en media cuadrática y el hecho de que $E(X^4) = 3\sigma^4$ cuando X se distribuye como una $N(0, \sigma^2)$. Se cambiará a $W_{t_k} - W_{t_{k-1}}$ por ΔW_k para simplificar la notación:

$$\begin{aligned} E \left| \sum_{k=1}^n (\Delta W_k)^2 - (T - t_0) \right|^2 &= E \left(\sum_{k=1}^n (\Delta W_k)^2 - (T - t_0) \right)^2 \\ &= E \left(\sum_{k=1}^n (\Delta W_k)^2 \sum_{k=1}^n (\Delta W_k)^2 \right) \\ &\quad - 2(T - t_0) E \left(\sum_{k=1}^n (\Delta W_k)^2 \right) + (T - t_0)^2 \\ &= \sum_{k=1}^n E(\Delta W_k)^4 + \sum_{i>j, i<j}^n E(\Delta W_i)^2 E(\Delta W_j)^2 \\ &\quad - 2(T - t_0)^2 + (T - t_0)^2 \\ &= \sum_{k=1}^n 3(\Delta t_k)^2 + \sum_{i>j, i<j}^n \Delta t_i \Delta t_j - (T - t_0)^2 \\ &= \sum_{k=1}^n 2(\Delta t_k)^2 + \left(\sum_{k=1}^n \Delta t_k \right)^2 - (T - t_0)^2 \\ &= \sum_{k=1}^n 2(\Delta t_k)^2 \\ &\leq 2(T - t_0) \max_{1 \leq k \leq n} (t_k - t_{k-1}) \rightarrow 0. \end{aligned}$$

□

Haciendo uso de los resultados obtenidos, de la siguiente desigualdad:

$$\sum_{k=1}^n |W_{t_k} - W_{t_{k-1}}|^2 \leq \max(|W_{t_k} - W_{t_{k-1}}|) \sum_{k=1}^n |W_{t_k} - W_{t_{k-1}}|,$$

se deduce que

$$\sum_{k=1}^n |W_{t_k} - W_{t_{k-1}}| \rightarrow \infty,$$

cuando $\delta_n \rightarrow 0$. Por lo que la integral

$$\int_{t_0}^t G(s) \xi_s ds = \int_{t_0}^t G(s) dW_s,$$

no puede ser definida como una integral de Riemann-Stieltjes. Para $G(s) = c$, con $c \in \mathbb{R}$ se tiene que

$$\int_{t_0}^t c dW_s = c(W_t - W_{t_0}),$$

que queda definida como una integral Riemann-Stieltjes ya que la suma

$$S_n = \sum_{i=1}^n G(\tau_i)(W_{t_i} - W_{t_{i-1}})$$

con $t_0 \leq t_1 \leq \dots \leq t_n = t$, $t_{i-1} \leq \tau_i \leq t_i$, converge a ese valor. Pero la convergencia de la suma para $G(s) = W_t$ depende de la elección de los puntos τ_i como se verá a continuación.

Como $a(b-c) = \frac{1}{2}(b^2 - c^2) - \frac{1}{2}(b-c)^2 + (a-c)^2 + (b-a)(a-c)$ se llega a la siguiente igualdad:

$$\begin{aligned} S_n &= \frac{1}{2}W_t^2 - \frac{1}{2}W_{t_0}^2 - \frac{1}{2} \sum_{i=1}^n (W_{t_i} - W_{t_{i-1}})^2 \\ &\quad + \sum_{i=1}^n (W_{\tau_i} - W_{t_{i-1}})^2 + \sum_{i=1}^n (W_{t_i} - W_{\tau_i})(W_{\tau_i} - W_{t_{i-1}}). \end{aligned}$$

Ya se demostró la siguiente convergencia:

$$\sum_{i=1}^n (W_{t_i} - W_{t_{i-1}})^2 = t - t_0$$

en media cuadrática. También se tiene que

$$\sum_{i=1}^n (W_{t_i} - W_{\tau_i})(W_{\tau_i} - W_{t_{i-1}})$$

converge a cero en media cuadrática. Para ver este resultado se toma en cuenta el hecho de que $(W_{t_i} - W_{\tau_i})$, $(W_{\tau_i} - W_{t_{i-1}})$, $(W_{t_j} - W_{\tau_j})$ y $(W_{\tau_j} - W_{t_{j-1}})$ son independientes para $i \neq j$. Entonces,

$$\begin{aligned} E[(W_{t_i} - W_{\tau_i})(W_{\tau_i} - W_{t_{i-1}})] &= E[W_{t_i}W_{\tau_i}] - E[W_{t_i}W_{t_{i-1}}] \\ &\quad - E[W_{\tau_i}W_{\tau_i}] + E[W_{\tau_i}W_{t_{i-1}}] \\ &= t_i - t_i - \tau_i + \tau_i \\ &= 0 \end{aligned}$$

por lo que

$$E \left[\sum_{i=1}^n (W_{t_i} - W_{\tau_i})(W_{\tau_i} - W_{t_{i-1}}) \right]^2 = \sum_{i=1}^n E [(W_{t_i} - W_{\tau_i})^2 (W_{\tau_i} - W_{t_{i-1}})^2] \rightarrow 0$$

ya que

$$\begin{aligned} E [(W_{t_i} - W_{\tau_i})^2] &= E [W_{t_i}^2] - 2E [W_{t_i}W_{\tau_i}] + E [W_{\tau_i}^2] \\ &= t_i - 2t_i + \tau_i \\ &= \tau_i - t_i \end{aligned}$$

que tiende a cero cuando $\delta_n \rightarrow 0$.

Sólo resta la siguiente convergencia:

$$qm - \lim_{\delta_n \rightarrow 0} \sum_{i=1}^n (W_{\tau_i} - W_{t_{i-1}})^2 = \sum_{i=1}^n \tau_i - t_{i-1}.$$

La convergencia de S_n depende de los valores de τ_i y en específico se tiene que para $\tau_i = t_{i-1}$ se llega a la integral de Ito de una dimensión:

$$\int_{t_0}^t W_s dW_s = \frac{1}{2} (W_t^2 - W_{t_0}^2) - \frac{1}{2}(t - t_0).$$

Con otra elección de τ_i , por ejemplo, $\tau_i = (t_i + t_{i+1})/2$, se puede llegar a otro tipo de integrales como la de R.L. Stratanovich. Sin embargo, la integral de Ito cuenta con la característica de ser una *martingala*, por lo que se usa para modelar procesos en los que la característica de *no arbitraje* influye, por ejemplo, en el cálculo del precio de opciones. Como se vio, la integral no se puede definir como una integral de Riemann-Stieltjes a menos que G sea una función constante por lo que una forma de resolver el problema es aproximar a la función por medio de una sucesión de funciones constantes que se definirán a continuación:

Definición 1.40. Se dirá que G_n es una función paso si existe una descomposición del intervalo $t_0 \leq t_1 \leq \dots \leq t_n = t$, tal que $G_n(s) = G(t_{i-1})$ para toda $s \in [t_{i-1}, t_i)$ en donde $i \in \{1, 2, \dots, n\}$ y la siguiente desigualdad se cumple:

$$\int_{t_0}^t |G_n(s)|^2 ds < \infty.$$

Con esta nueva definición se puede introducir el concepto de la *integral de Ito* como una aproximación:

Definición 1.41. Sea $G(s)$ \mathcal{F} – medible para toda $s \in [t_0, t]$ y asumamos que se cumple con probabilidad 1 que

$$\int_{t_0}^t |G(s)|^2 ds < \infty.$$

Una función que cumpla con estas características se dirá que pertenece a $M_2[t_0, t]$. Entonces la integral de Ito se define como

$$\int_{t_0}^t G(s) dW_s = st - \lim_{n \rightarrow \infty} \int_{t_0}^t G_n dW_s$$

en donde G_n es una secuencia de funciones paso definidas en $[t_0, t]$ que convergen a G en el siguiente sentido:

$$st - \lim_{n \rightarrow \infty} \int_{t_0}^t |G(s) - G_n(s)|^2 ds = 0.$$

La necesidad de que G y G_n pertenezcan a $M_2[t_0, t]$ radica en que en caso contrario la condición de convergencia no se cumpliría. La condición de convergencia y la forma en que se definen las funciones escalón son necesarias para asegurar la existencia de la integral de Ito con la elección de $\tau_i = t_{i-1}$. Así, la integral de Ito se puede ver como el límite de integrales de Riemann-Stieltjes (para mayor detalle véase [22]). Existe una propiedad interesante que será útil en demostraciones futuras, esta propiedad es llamada la *isometría de Ito*.

Teorema 1.2. Sean $G(t)$ y W_t definidas de la misma forma que en la definición 1.41. La siguiente igualdad se cumple:

$$E \left[\int_{t_0}^t G(s) dW_s \right]^2 = E \left[\int_{t_0}^t [G(s)]^2 ds \right].$$

Demostración. De la siguiente expresión se puede ver el resultado con facilidad:

$$\begin{aligned} E \left[\left| \sum_{k=0}^{n-1} G(k) (W_{t_{k+1}} - W_{t_k}) \right|^2 \right] &= \sum_{k=0}^{n-1} \sum_{j=0}^{n-1} E [G(k)] E [W_{t_{k+1}} - W_{t_k}] \cdots \\ &\quad \cdots E [G(j)] E [W_{t_{j+1}} - W_{t_j}] \\ &= \sum_{k=0}^{n-1} E [G(k)]^2 E [W_{t_{k+1}} - W_{t_k}]^2 \\ &= \sum_{k=0}^{n-1} E [G(k)]^2 (t_{k+1} - t_k). \end{aligned}$$

Aplicando el límite a esta ecuación se llega a la siguiente igualdad:

$$E \left[\int_{t_0}^t G(s) dW_s \right]^2 = E \left[\int_{t_0}^t [G(s)]^2 ds \right].$$

□

La propiedad de linealidad implica que para toda constante a y para cualesquiera funciones $G(s)$ y $H(s)$ definidas como en 1.40, se tiene que:

$$\int_{t_0}^t a [G(s) - H(s)] dW_s = a \int_{t_0}^t G(s) dW_s - a \int_{t_0}^t H(s) dW_s.$$

Esta propiedad se deriva fácilmente de las propiedades de los límites y su prueba es trivial, como se comenta en [26] y en [22]. La integral de Ito puede

verse como un proceso estocástico que cambia en el tiempo de acuerdo con el intervalo elegido para evaluar la integral. En el caso de que la integral de Ito sea considerada como un proceso estocástico, se pueden observar características interesantes como la propiedad de martingala. Primero se verán las características básicas de todo proceso estocástico, esto es, su esperanza y función de autocovarianzas. La esperanza es cero desde que:

$$E[W_t - W_{t-1}] = 0,$$

y la función de autocovarianzas es:

$$\begin{aligned} E \left[\int_{t_0}^t G(s) dW_s \int_{t_0}^u H(s) dW_s \right] &= \int_{t_0}^{\min(t,u)} G(s) H(s) E[dW_s^2], \\ &= \int_{t_0}^{\min(t,u)} G(s) H(s) ds, \end{aligned}$$

por lo que el segundo momento es la varianza:

$$\int_{t_0}^t [G(s)]^2 ds.$$

La propiedad de martingala es de gran importancia porque permite obtener una ecuación en donde existe el juego justo, esto es, el evento del periodo siguiente a lo observado tiene las mismas probabilidades de ocurrir que en el periodo anterior, no importando lo ocurrido anteriormente.

Proposición 1.6. *La integral de Ito $\int_{t_0}^t G(s) dW_s$ es una martingala con respecto a una filtración $\{\mathcal{F}_t\}$ del movimiento Browniano.*

Demostración. Por definición, la integral es \mathcal{F}_t -medible. El primer momento y el segundo son acotados, como se vio previamente. Si la integral de Ito fuera una martingala, se tendría la siguiente igualdad:

$$E \left[\int_{t_0}^s G(s) dW_s - \int_{t_0}^t G(s) dW_s \mid \mathcal{F}_t \right] = E \left[\int_t^s G(s) dW_s \mid \mathcal{F}_t \right] = 0.$$

Es decir, el problema se reduce a demostrar la igualdad con cero de la ecuación anterior. Para una función paso evaluada en el segmento $t \leq t_k \leq t_{k+1} \leq s$, con la filtración $\mathcal{F}_t \subseteq \mathcal{F}_{t_k}$ y con $G(t_k)$ \mathcal{F}_{t_k} -medible, se tiene lo siguiente:

$$\begin{aligned} E [G(t_k) (W_{t_{k+1}} - W_{t_k}) | \mathcal{F}_t] &= E [E [G(t_k) (W_{t_{k+1}} - W_{t_k}) | \mathcal{F}_{t_k}] | \mathcal{F}_t] \\ &= E [G(t_k) E [(W_{t_{k+1}} - W_{t_k}) | \mathcal{F}_{t_k}] | \mathcal{F}_t] \\ &= E [G(t_k) E [(W_{t_{k+1}} - W_{t_k})] | \mathcal{F}_t] \\ &= E [G(t_k) 0 | \mathcal{F}_t] \\ &= 0. \end{aligned}$$

Aplicando los límites se comprueba la proposición.

□

Capítulo 2

Simulación de modelos estocásticos

Para simular un modelo estocástico primero se requiere calcular los parámetros del modelo, por lo que antes de empezar con los métodos de simulación por computadora se realizará una introducción a los métodos usados para la parametrización. Encontrar el parámetro de desviación estándar del error del modelo es imprescindible para simular una serie de tiempo y por ello se pondrá especial atención en la obtención de este parámetro. Enseguida se puede usar un método de simulación para obtener trayectorias posibles y congruentes con el comportamiento de la variables. Una vez que se calculan estas trayectorias, se guardan como datos ordenados para su posterior análisis.

2.1. Estimación de parámetros

La estimación de los parámetros de los modelos representa gran parte del proceso de modelación ya que estos parámetros determinarán los resultados del modelo. Existe un conjunto de métodos que pueden llevar a malas interpretaciones y en modelos que no cumplen con los objetivos iniciales. La mayoría de los métodos se pueden encontrar en todos los libros de estadística básica. En la práctica, aplicar estos métodos representa una inversión de tiempo y conocimiento considerable ya que la elección de los datos, los algoritmos seleccionados y la experiencia del modelador en el comportamiento de la variable determinarán los resultados del modelo. En esta sección se presentarán los métodos básicos y más usados, según [33], para encontrar parámetros; se exponen algunos otros en [3]. Los primeros métodos son los de *momentos* y de *estimación Bayesiana* que son más usados en modelos

simples como el del *movimiento geométrico Browniano*. Para modelos con más parámetros se usan métodos más robustos como el de algoritmos de optimización para encontrar un máximo único en una función de *máxima verosimilitud* de más de dos parámetros que es, en ocasiones, difícil de obtener. Después de presentar los métodos se dará un breve resumen de conceptos básicos para seleccionar los datos y probar los modelos.

2.1.1. Método de los momentos

El método consiste en usar los estimadores de los momentos para obtener los parámetros de la función de distribución de probabilidad. Fue ideado por Pearson y es el más antiguo. Al tomar una muestra X_1, X_2, \dots, X_n , se le aplican los estimadores de k momentos muestrales y éstos son usados para resolver un sistema de ecuaciones que contiene los parámetros de la distribución. El estimador del k ésimo momento es el siguiente:

$$m_k = \frac{1}{n} \sum_{i=1}^n X_i^k.$$

Como ejemplo, se hallarán los parámetros y términos de los momentos para una función de distribución $N(\theta, \sigma^2)$, para lo cual se sustrae una muestra X_1, X_2, \dots, X_n . El primer momento resulta ser una estimación de θ y, a su vez, este estimador es insesgado. Para σ^2 se tiene que

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - m_1^2 = \frac{1}{n} \sum_{i=1}^n (X_i - m_1)^2,$$

obteniendo una estimación de los parámetros de una distribución normal.

2.1.2. Método de Bayes

Este método de estimación hace uso de la regla de Bayes, por lo que recibe su nombre, y consiste en postular una función de distribución del parámetro a estimar. Esta distribución es llamada distribución previa ya que es formulada antes de que la muestra sea analizada. Después se postula una distribución sobre la variable en cuestión llamada posterior y usando la regla de Bayes se encuentra la distribución posterior. Denotemos la distribución previa como $\pi(\theta)$ y la función de distribución posterior como $f(x | \theta)$. Entonces se tiene que

$$\pi(\theta | x) = \frac{f(x | \theta) \pi(\theta)}{m(x)},$$

en donde $m(x)$ es la función de distribución marginal del modelo. Una vez que se cuenta con esta distribución, se toma el valor esperado como estimación del parámetro deseado. En lo sucesivo no se hará uso de este método aunque resulta conveniente cuando se cuenta con pocos datos. En cambio, se hará énfasis en el siguiente método ya que aunque no brinda una expresión algebraica para la estimación, sí permite realizar aproximaciones por medio de algoritmos, atributo que lo hace uno de los más usados.

2.1.3. Máxima verosimilitud

El método de *máxima verosimilitud* teóricamente se puede aplicar a todos los tipos de distribuciones. Sin embargo, el problema de encontrar un máximo de una función de varias variables puede convertirse en un problema de optimización y no en todos los casos se puede encontrar un máximo que no sea divergente o único dentro de un intervalo multidimensional de observación.

Supóngase que se tiene una muestra $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ proveniente de una realización de las variables aleatorias X_1, X_2, \dots, X_n idénticamente distribuidas y mutuamente independientes con función de distribución de probabilidades o de masa de probabilidad $f(\cdot | \theta_1, \theta_2, \dots, \theta_k)$. La *función de verosimilitud* se define como:

$$L(\theta | \mathbf{x}) = L(\theta_1, \theta_2, \dots, \theta_k | x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k).$$

Definición 2.1. *El estimador de máxima verosimilitud del parámetro θ de las variables X_1, X_2, \dots, X_n , es la función $\hat{\theta}(\mathbf{x})$ tal que para toda muestra \mathbf{x} , $L(\theta | \mathbf{x})$ alcanza el valor máximo como una función de θ .*

Para empezar con una estimación de los parámetros, se debe establecer la función de distribución de cada X_i y luego encontrar el máximo con que cuenta la función de verosimilitud, para lo cual se usa la derivada parcial con respecto a θ de la función de verosimilitud y se iguala con cero. Sin embargo, esto no asegura que el valor extremo alcanzado sea único, por lo que se necesitará de la teoría del cálculo para saber si el valor extremo encontrado es un máximo y esto restringe el universo de funciones utilizables, por ejemplo, para el caso de dos parámetros de la misma función de distribución, se necesita que cuente con segundas derivadas parciales negativas y su jacobiano sea positivo, por lo que un algoritmo para encontrar un máximo local de la

función suele ser una opción más adecuada. El lenguaje R cuenta con diferentes métodos de optimización que encuentran un máximo o un mínimo de acuerdo con una función propuesta.

Como ejemplo para emplear el método de máxima verosimilitud, considérese el siguiente modelo Z_t :

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} \cdots - \theta_q a_{t-q} + \theta_0,$$

con

$$W_t = \nabla^d Z_t,$$

y en donde $d \in \{0, 1, 2, \dots, \infty\}$. Con este modelo, se puede despejar el error de la siguiente forma:

$$a_t = W_t - \phi_1 W_{t-1} - \phi_2 W_{t-2} - \cdots - \phi_p W_{t-p} - \theta_0 + \theta_1 a_{t-1} \cdots + \theta_q a_{t-q}.$$

Considérese que el error del modelo a_t es un ruido blanco con media cero y desviación estándar σ_a^2 . Bajo este supuesto ¹ la función de distribución conjunta de $a_{d+p+1}, a_{d+p+2}, \dots, a_N$ en donde N es un número entero mayor que cero, es de la siguiente forma:

$$f(a_{d+p+1}, a_{d+p+2}, \dots, a_N) = \frac{1}{(2\pi)^{\frac{(N-d-p)}{2}} \sigma_a^{(N-d-p)}} \exp \left\{ -\frac{1}{2\sigma_a^2} \sum_{t=d+p+1}^N a_t^2 \right\}.$$

El valor que se conoce (debido a una muestra) es el del conjunto de valores

$$\tilde{W} = \{W_{d+p+1}, W_{d+p+2}, \dots, W_N\},$$

por lo que el conjunto de Φ parámetros a encontrar contiene los siguientes elementos $\phi_1, \phi_2, \dots, \phi_p, \theta_0, \theta_1, \theta_2, \dots, \theta_q$ y σ_a^2 . Así, el problema se enfoca en minimizar los errores, esto es, minimizar la siguiente función:

$$V(\Phi) = \sum_{t=d+p+1}^N |W_t - \phi_1 W_{t-1} - \phi_2 W_{t-2} \cdots - \phi_p W_{t-p} - \theta_0 + \theta_1 a_{t-1} \cdots + \theta_q a_{t-q}|,$$

¹Este supuesto puede ser diferente, pero para efectos de ejemplificación se tomará como base ya que es usado ampliamente en la literatura especializada, como en [13].

para posteriormente encontrar el parámetro que maximiza la función de verosimilitud con la estimación $\tilde{\Phi}$ de Φ . Por tratarse de la función de distribución normal, es conveniente aplicar la función logaritmo a la función de la siguiente forma:

$$\log \left\{ L \left(\sigma_a^2 \mid \tilde{\Phi} \right) \right\} = \left[-\frac{1}{2} (N - d - p) [\log (2\pi) - \log (\sigma_a^2)] \right] - \frac{1}{2\sigma_a^2} V \left(\tilde{\Phi} \right).$$

Posteriormente, se encuentra el máximo de esta función por medio de su derivada:

$$\left. \frac{\partial L}{\partial \sigma_a^2} \right|_{\sigma_a^2 = \tilde{\sigma}_a^2} = 0,$$

obteniendo un máximo en

$$\tilde{\sigma}_a^2 = \frac{V \left(\tilde{\Phi} \right)}{N - d - p}.$$

De manera que, para este modelo, el problema se reduce a minimizar $V \left(\tilde{\Phi} \right)$. El problema de minimización empieza con la elección de los valores iniciales de los parámetros para iniciar el método de optimización. La elección de estos parámetros reducirá el número de iteraciones. Los valores iniciales podrán ser elegidos mediante un método lógico y estadístico por lo que se podrá obtener un sustento científico sobre su elección. Para los modelos AR, MA y ARMA la función de autocorrelaciones proporciona un conjunto de ecuaciones que al resolverse pueden proveer de una estimación de los valores iniciales para emprender el método de optimización. Para ilustrar el método se desglosará la función de correlación primero para un modelo AR(p) estacionario que se muestra a continuación:

$$Z_t = \sum_{k=1}^p \tilde{\varphi}_k Z_{t-k},$$

en donde

$$E [Z_t] = \mu,$$

por lo que

$$\begin{aligned}\mu &= E \left[\sum_{k=1}^p \tilde{\varphi}_k Z_{t-k} + \epsilon_k \right] \\ &= \sum_{k=1}^p \tilde{\varphi}_k \mu.\end{aligned}$$

Se sigue entonces que

$$Z_t - \mu = \sum_{k=1}^p \tilde{\varphi}_k (Z_{t-k} - \mu) + \epsilon_k.$$

Por medio de este resultado se obtiene la siguiente ecuación recursiva:

$$\begin{aligned}\gamma_k &= E [(Z_t - \mu) (Z_{t-k} - \mu)] \\ &= \sum_{i=1}^p E [(\tilde{\varphi}_i (Z_t - \mu) + \epsilon_i) (Z_{t-i} - \mu)] \\ &= \sum_{i=1}^p E [\tilde{\varphi}_i (Z_t - \mu) (Z_{t-i} - \mu) + \epsilon_i (Z_{t-i} - \mu)] \\ &= \sum_{i=1}^p \tilde{\varphi}_i \gamma_{k-i}.\end{aligned}$$

Al dividir por γ_0 se obtiene el sistema de ecuaciones de Yule -Walker:

$$\begin{aligned}\rho_1 &= \tilde{\varphi}_1 + \tilde{\varphi}_2 \rho_1 + \cdots + \tilde{\varphi}_p \rho_{p-1} \\ \rho_2 &= \tilde{\varphi}_1 \rho_1 + \tilde{\varphi}_2 + \cdots + \tilde{\varphi}_p \rho_{p-2} \\ &\dots \\ \rho_p &= \tilde{\varphi}_1 \rho_{p-1} + \tilde{\varphi}_2 \rho_{p-2} + \cdots + \tilde{\varphi}_p.\end{aligned}$$

En este caso, el valor ρ es una estimación de la autocorrelación. Al encontrar la solución de este sistema de ecuaciones se pueden obtener los valores iniciales para la optimización de la función de verosimilitud de un modelo AR(p). De la misma forma se puede obtener un sistema de ecuaciones para un modelo MA(q):

$$\rho_k = \frac{-\tilde{\theta}_k + \tilde{\theta}_1 \tilde{\theta}_{k+1} + \cdots + \tilde{\theta}_{q-k} \tilde{\theta}_q}{1 + \tilde{\theta}_1^2 + \tilde{\theta}_2^2 + \cdots + \tilde{\theta}_q^2},$$

para $k = 1, \dots, q$. Con estos resultados se puede encontrar un sistema de ecuaciones para un modelo ARMA(p, q), primero resolviendo el sistema proveniente de la parte autorregresiva y posteriormente el de la parte de promedios móviles. Más adelante se presentará un ejemplo sobre el uso de este método.

2.2. Evaluación del modelo

La evaluación de los modelos de precios usados para la simulación y evaluación del riesgo no puede ser igual a la de modelos de pronósticos ya que su objetivo, precisamente, es el de valorar las diferencias en los precios y no un punto estimado en específico. Un modelo de series de tiempo se puede evaluar comparando pronósticos de la serie y valores de la misma serie, las diferencias entre éstos son llamadas *residuos*, sin embargo, en el caso de los modelos de precios que son simulados para valorar qué diferencias se pueden presentar en un precio a lo largo del tiempo, el mejor modelo será aquel cuya función de distribución sea la más parecida a la distribución de la volatilidad de la variable económica evaluada. En concreto, el conjunto de residuos se conforma por los siguientes valores:

$$\vartheta_t - \tilde{\vartheta}_t = \Lambda_t,$$

en donde ϑ_t es el valor real obtenido de una muestra para el periodo t y $\tilde{\vartheta}_t$ es el valor estimado por el modelo. La función de distribución de los residuos es de gran interés para la evaluación del modelo ya que por una parte se supone normalidad, esperanza cero y una desviación estándar baja. Siguiendo el pensamiento de la evaluación del modelo de series de tiempo, en [13] se afirma que el mejor modelo es el que cuenta con la desviación estándar más pequeña correspondiente a la función de distribución de los residuos.

Los constantes cambios en los sistemas económicos precisan considerar también valores poco probables para la variable económica, por lo que si se obtiene un modelo que tenga una función de distribución muy parecida a la de la variable en un determinado lapso de tiempo, existe riesgo de sobre estimación y no se contemplarían todos los valores factibles. Para resolver lo anterior, hasta cierto punto, en [29] proponen dividir la muestra en dos partes: una de estimación y otra para evaluar los resultados obtenidos a través del modelo, esto es, comparar la distribución estimada con la datos no usados para la estimación de los parámetros. Bajo este pensamiento se tiene que dejar en claro que los modelos vistos hasta el momento no son pronósticos de un punto, sino, herramientas para el cálculo de un intervalo

y no pueden evaluarse por medio de las diferencias contra la serie real. Para tener un modelo que utilice ciclos económicos predecibles, se puede emplear alguno de los modelos econométricos vistos.

Una de las formas de evaluar un modelo es realizar pruebas estadísticas a los datos para corroborar los supuestos del mismo. De manera que, cuando se desarrolle un ejemplo, se realizarán pruebas a los supuestos hechos sobre la serie para poder aplicar los modelos expuestos.

Las pruebas más importantes de los modelos de series de tiempo son sobre los supuestos de normalidad y estacionariedad. Una inspección visual de las variables ayuda a verificar estos supuestos, pero un método más confiable consiste en emplear pruebas de hipótesis estadísticas. En el anexo A se presentarán los estadísticos de prueba de algunas de las hipótesis más usadas.

2.3. Obtención de números pseudoaleatorios

La simulación por computadora tiene como fin analizar el comportamiento de sistemas complejos por medio de una representación reducida, pues de otra manera serían imposibles de comprender o de llevar a la práctica por su complejidad. En esta sección se presenta una introducción a los conceptos básicos de la simulación que servirá como base para presentar en el siguiente capítulo un ejemplo de aplicación.

Una de las aplicaciones de las computadoras es la representación de modelos de sistemas complejos que requieran de números aleatorios y que permitan la simulación de procesos.

Si bien las simulaciones de modelos ya existían antes de la invención de las computadoras, resultó ser más económico y fácil de llevarlas a cabo gracias al desarrollo y accesibilidad que han tenido las computadoras y los lenguajes de programación a lo largo de los últimos años (véase [21]). En las siguientes secciones se verán algunas aplicaciones y repercusiones que ha tenido la simulación por computadora en el mundo moderno.

En lo sucesivo se entenderá por un *sistema* al proceso que es objeto de estudio, al cual se le atribuirán supuestos. Por *modelo* se entenderá a la versión matemática reducida del comportamiento de un fenómeno; por *simulación*, al proceso en el que se emplea una computadora para evaluar un modelo numéricamente y en el que los datos son generados para estimar las características del mismo.

Para generar números aleatorios para los modelos se requiere un algoritmo que comience con un escalar llamado *semilla*. Este tipo de algoritmos suelen generar secuencias fijas de números reales, por lo cual se les denomina generadores de números *pseudoaleatorios*.

Para generar números pseudoaleatorios existen diferentes métodos pero uno de ellos sobresale dado que cuenta con un gran periodo (concepto que se explicará en breve) y es el que utilizan los programas usados en este texto, por lo que no se abundará en otros métodos².

El *Generador Congruencial Lineal* (GCL) consiste en la siguiente fórmula recursiva:

$$Z_i = (aZ_{i-1} + c)(\text{mod}(m)),$$

en donde $\text{mod}(m)$ es el módulo m , a y c son dos constantes arbitrarias y Z_0 es la semilla (que es un número inicial arbitrario).

Para obtener un número y que tenga función de distribución uniforme en el intervalo $(0, 1)$ se divide a Z_i entre m .

Se usa este método ya que la división de cualquier número entre un número m es siempre única.

Para la realización de una simulación conviene obtener la mayor cantidad posible de números aleatorios. Sin embargo, el GCL tiene ciertas limitaciones que dependen de la elección de sus parámetros. Se darán los principios básicos sobre la elección de los mismos.

Los números generados por el GCL se repiten cada cierto número de veces y esta propiedad puede verse fácilmente ya que el generador sólo puede dar a lo más un subconjunto finito de números racionales $\left\{0, \frac{1}{m}, \frac{2}{m}, \dots, \frac{(m-1)}{m}\right\}$, por lo que si se repite el proceso infinitamente entonces en algún momento se tendría que escoger otra vez alguno de los números ya usados como semilla, volviendo a repetir la secuencia de números. A la cantidad de números generados antes de que se repitan se le llama *periodo* del generador. Como sólo puede tomar los números $\left\{0, \frac{1}{m}, \frac{2}{m}, \dots, \frac{(m-1)}{m}\right\}$, entonces el periodo debe ser a lo más m . En caso de que el periodo sea m entonces se dice que el generador tiene *periodo completo* o *full period*. Para obtener un generador con periodo completo existe el siguiente teorema.

²En [24] se pueden ver los alcances de otros métodos.

Teorema 2.1. *El LCG tiene periodo completo si y sólo si las siguientes tres condiciones se cumplen.*

1. *El único entero positivo que divide a m y c es el número 1.*
2. *Si q es un número primo que divide a m , entonces q divide a $a-1$.*
3. *Si 4 divide a m , entonces 4 divide a $a-1$.*

Este teorema puede ayudar a construir un generador apropiado, pero existe otro problema importante que es la capacidad de las computadoras. Cuando las computadoras guardan un número usan cierta cantidad de bits llamada *longitud de palabra* por lo que la representación del número m no puede exceder dicha longitud. El número máximo que se puede almacenar en una computadora con *longitud de palabra* de l bits es 2^{l-1} pues un bit es usado para el signo. El número l suele ser muy grande pues la capacidad de las computadoras ha aumentado, permitiendo que existan generadores con periodos grandes.

El caso en el que $c > 0$, se conoce como GCL mixto; si $c = 0$, entonces se dice que es un GCL multiplicativo. En caso de usar el GCL mixto se puede obtener periodo completo, pero con el GCL multiplicativo esto no es posible pues la condición 1 no puede ser satisfecha ya que m divide a m y a cero al mismo tiempo. Por otra parte, el GCL multiplicativo es más fácil de usar ya que evita perder recursos en la investigación de un parámetro c apropiado.

Hasta ahora se ha mostrado el algoritmo básico de los GCL, pero existen diferentes generadores de números pseudoaleatorios. También existen generadores de números aleatorios que no siguen un algoritmo computacional y son imposibles de predecir. Algunos de estos generadores están basados en la recopilación de datos aleatorios desde la atmósfera o el movimiento del cursor hecho por el usuario. El uso de estos generadores es bueno desde el punto de vista de que no tienen periodo y no son predecibles, lo que resulta útil en el uso de la encriptación. Para el software R existen librerías que permiten descargar secuencias de números aleatorios desde internet³, por ejemplo, del servidor de www.random.org que en lo particular obtiene los números con datos de la atmósfera. Sin embargo, no se hará uso de estos números aleatorios y en cambio se usarán generadores de números pseudoaleatorios propios de R y que han demostrado tener un periodo grande, sobrepasando a los generadores convencionales del tipo GCL. El uso de uno u otro generador se deja al juicio del lector y para los objetivos de este trabajo, el generador que usa R

³Software estadístico del que se hará uso en el presente trabajo.

será suficiente ya que sobrepasa la cantidad de números necesarios para obtener los resultados requeridos. A continuación se mostrarán las características básicas del generador que usa R con el fin de mostrar su potencial.

La siguiente fórmula recursiva es en la que se basa el generador conocido como *Mersenne Twister* (MT). El algoritmo genera una secuencia de números enteros entre 0 y $2^w - 1$ en donde w es la *longitud de palabra* del procesador. Dividiendo estos números entre $2^w - 1$ se obtiene un número entre 0 y 1.

La fórmula recursiva es la siguiente:

$$X_{k+n} := X_{k+m} \oplus (X_k^u | X_{k+1}^l)A \text{ con } k = 0, 1, \dots,$$

en donde n es un entero llamado *grado de recurrencia*; r , un entero tal que $0 \leq r \leq w - 1$; m , un entero tal que $1 \leq m \leq n$; y A , una matriz constante de $w \times w$ con elementos en el conjunto $\{0, 1\}$. En el lado derecho de la ecuación, X_k^u representa los $w - r$ bits superiores de X_k y X_{k+1}^l los otros r bits no superiores de X_{k+1} . $(X_k^u | X_{k+1}^l)$ es la concatenación de X_k^u con X_{k+1}^l . \oplus es un operador aditivo bit a bit módulo 2 (XOR). Primero se dan las semillas iniciales X_1, X_2, \dots, X_{n-1} en donde X_i ($i = 1, \dots, n - 1$) es un vector de tamaño w que tiene sus elementos en el conjunto $\{0, 1\}$ y que es la representación en bits de algún entero. Al obtener $(X_k^u | X_{k+1}^l)$ se procede a multiplicar por A , después al vector resultante se le adhiere X_{k+m} por medio de \oplus y se obtiene X_{k+n} .

El generador usa las $n - 1$ semillas para empezar a generar números. En 1998, Makoto Matsumoto y Takuji Nishimura presentaron una versión que mejora la uniformidad de los datos y alcanza un periodo de $2^{19937} - 1$. Este generador requiere de 624 semillas. En este trabajo se usará el MT19937 para generar números aleatorios ya que supera a los demás algoritmos en el tamaño del periodo y en la calidad de la uniformidad de los datos (véase [24]).

Algunos de los programas y lenguajes de programación más usados para la simulación se presentan a continuación y aunque cada uno tiene características diferentes, la utilidad o conveniencia en el uso de uno u otro es a criterio de cada usuario.

1. Arena es un software de simulación de propósito general⁴ que es producido por Systems Modeling Corporation. Funciona usando diagramas de

⁴Con propósito general se quiere decir que el software es competente para llevar a cabo la simulación de un sistema de cualquier campo.

flujo conectando procesos y asignando a cada proceso atributos. Se puede iniciar una simulación que pasa datos de un proceso a otro. Cuenta con herramientas gráficas que permiten un mayor entendimiento y descripción de cada proceso y del sistema.

2. Extend es un software de simulación de propósito general que comparte las mismas características que Arena conectando procesos para lograr construir el sistema bajo estudio.

3. GPSS es un lenguaje de simulación de propósito general que permite analizar un sistema paso a paso por cada uno de sus procesos. Es ampliamente conocido por ser uno de los primeros lenguajes enfocados en simulación.

No necesariamente se necesita usar software especializado en simulación, pues con una computadora que pueda generar números aleatorios y hacer operaciones elementales se puede hacer una simulación. En este trabajo las simulaciones se harán con el software estadístico R de distribución *libre* con el fin de programar funciones estadísticas y mostrar el funcionamiento de los procesos. En el anexo B se muestran los códigos de R para cada simulación y gráfica de este trabajo.

2.4. Números aleatorios con una distribución específica

Para empezar una simulación con los supuestos sobre la distribución de las variables, se necesita crear números aleatorios que sigan dicha distribución. Para ello existen métodos que tienen características diferentes y que cada usuario escoge, de acuerdo con su criterio y experiencia. El método de *aceptación y rechazo* es el más usado por los programas computacionales y es el que sirve para obtener cualquier distribución que se pueda graficar en un plano. Sin embargo, por sus características ocupa gran cantidad de recursos computacionales comparado con los otros métodos. Afortunadamente, el avance en la velocidad de las computadoras que se ha dado en años recientes permite que el tiempo ocupado en procesamiento sea menor al crear los números aleatorios que tienen una distribución dada.

A continuación se introducirá un método para simular variables aleatorias con funciones de distribución discreta. Para este tipo de funciones es suficiente simular una variable aleatoria uniforme y asignar rangos de diferentes tamaños como lo muestra el siguiente teorema.

2.4. NÚMEROS ALEATORIOS CON UNA DISTRIBUCIÓN ESPECÍFICA 47

Teorema 2.2. *Considere que p es una función de probabilidad para una función de distribución discreta. Sean $x_1 < x_2 < \dots$ todos los valores para los que $p(x_i) > 0$, y sea U_1 una variable aleatoria que se distribuye $U(0, 1)$. Sea*

$$Y = \text{mín} \left\{ x_j : \sum_{k=1}^j p(x_k) \geq U_1 \right\}.$$

Entonces Y es una variable aleatoria discreta cuya función de probabilidad es p .

Demostración.

$$\begin{aligned} P(Y = x_i) &= P \left\{ \sum_{k=1}^{i-1} p(x_k) < U_1 \leq \sum_{k=1}^i p(x_k) \right\} \\ &= \sum_{k=1}^i p(x_k) - \sum_{k=1}^{i-1} p(x_k) \\ &= p(x_i). \end{aligned}$$

□

Para las variables aleatorias continuas existen otros métodos que son diferentes a los de las variables discretas.

El método de la *transformación inversa* está basado en el siguiente teorema:

Teorema 2.3. *Si se genera un número l que se distribuye como una variable uniforme en el intervalo $(0, 1)$, entonces un número x con función de distribución F se obtiene con la función inversa F^{-1} (si es que existe) de la siguiente manera:*

$$x = F^{-1}(l).$$

Demostración. Sea y que se distribuye como una $U(0, 1)$, i.e, que $P(l \leq \xi) = \xi$ para cualquier ξ tal que $0 \leq \xi \leq 1$. Entonces

$$\begin{aligned} P(F^{-1}(l) \leq x) &= P(l \leq F(x)) \\ &= F(x). \end{aligned}$$

□

Este método es muy eficiente con respecto a la cantidad de cálculos que se necesitan para obtener los datos necesarios. Sin embargo, para aplicarlo se requiere contar con la función F^{-1} que en muchas ocasiones es imposible de encontrar.

Ahora se dará un ejemplo para mostrar el desarrollo del algoritmo. Simulemos una variable aleatoria con distribución exponencial con parámetro λ cuya función de distribución es la siguiente:

$$F(x) = \int_0^x \lambda \exp(-\lambda x) dx = 1 - \exp(-\lambda x).$$

Si se simula una variable Y con distribución $U(0, 1)$, entonces se obtiene una variable aleatoria con distribución exponencial con el método descrito:

$$x = -\frac{1}{\lambda} \ln(1 - y).$$

Otro método que se usa para todo tipo de funciones de distribución es el de *aceptación y rechazo* que fue ideado por el físico Von Neumann en 1951. Consiste en aceptar números aleatorios que cumplen con ciertas condiciones y rechazar a los que no las cumplen; las condiciones son elaboradas de tal forma que los números aceptados tengan una función de distribución dada (véase [21]).

Teorema 2.4. *Un algoritmo para generar una variable aleatoria X con densidad f es el siguiente:*

1. *Generar una variable aleatoria Y con función de distribución $G(\cdot)$.*
2. *Generar $U \sim U(0, 1)$, independiente de Y .*
3. *Si $U \leq \frac{f(Y)}{cg(Y)}$ (en donde $g(\cdot)$ es la función de densidad de Y), entonces aceptar $X = Y$, de lo contrario volver a 1 e intentar de nuevo.*

Demostración. 1. Como U es $U(0, 1)$, entonces

$$\begin{aligned} P\left(U \leq \frac{f(y)}{cg(y)} \mid Y = y\right) &= P\left(U \leq \frac{f(y)}{cg(y)}\right) \\ &= \frac{f(y)}{cg(y)}. \end{aligned}$$

2. Por el teorema de la probabilidad total (véase [25]), se tiene que:

$$\begin{aligned}
 P\left(U \leq \frac{f(Y)}{cg(Y)}\right) &= \int_{-\infty}^{\infty} P\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y = y\right) g(y) dy \\
 &= \int_{-\infty}^{\infty} P\left(U \leq \frac{f(y)}{cg(y)} \mid Y = y\right) g(y) dy \\
 &= \int_{-\infty}^{\infty} \frac{f(y)}{cg(y)} g(y) dy \\
 &= \frac{1}{c} \int_{-\infty}^{\infty} \frac{f(y)}{g(y)} g(y) dy \\
 &= \frac{1}{c} \int_{-\infty}^{\infty} f(y) dy \\
 &= \frac{1}{c}.
 \end{aligned}$$

3. También se tiene que

$$\begin{aligned}
 P\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y \leq y\right) &= \frac{P\left(U \leq \frac{f(Y)}{cg(Y)}, Y \leq y\right)}{G(y)} \\
 &= \frac{1}{G(y)} \int_{-\infty}^y P\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y = w\right) g(w) dw \\
 &= \frac{1}{G(y)} \int_{-\infty}^y P\left(U \leq \frac{f(w)}{cg(w)} \mid Y = w\right) g(w) dw \\
 &= \frac{1}{G(y)} \int_{-\infty}^y \frac{f(w)}{cg(w)} g(w) dw \\
 &= \frac{F(y)}{cG(y)}.
 \end{aligned}$$

Usando 1, 2 y 3 se obtiene:

$$\begin{aligned}
 P\left(Y \leq y \mid U \leq \frac{f(Y)}{cg(Y)}\right) &= \frac{P\left(U \leq \frac{f(Y)}{cg(Y)} \mid Y \leq y\right) P(Y \leq y)}{P\left(U \leq \frac{f(Y)}{cg(Y)}\right)} \\
 &= \frac{F(y)}{cG(y)} G(y) c \\
 &= F(y).
 \end{aligned}$$

□

2.5. Generación de números aleatorios dependientes

El tema de generación de números aleatorios dependientes es de interés para los capítulos siguientes ya que se usarán variables aleatorias dependientes. Se presenta un método llamado *factorización de Cholesky* y que sirve para simular variables aleatorias con función de distribución conjunta multinormal. El algoritmo consiste en obtener la matriz de covarianzas de un conjunto de variables aleatorias y después encontrar la matriz inferior de la descomposición de Cholesky para posteriormente hacer una pequeña transformación (véase [28]).

Teorema 2.5. *Sea $X = (X_1, \dots, X_p)$ un vector de variables aleatorias con función de distribución conjunta multinormal con media $\mu = 0$ y matriz de covarianzas $\Sigma = I$ en donde I es la matriz unidad (esto quiere decir que las variables X_i , con $i = 1, 2, \dots, p$ son independientes unas de otras). Sea A una matriz de covarianzas de dimensión $p \times p$ que puede descomponerse en sus factores de Cholesky y sea C la matriz inferior de la descomposición de Cholesky de A . Entonces el vector aleatorio $Y = CX$ tiene distribución p -variada multinormal con media $\mu = 0$ y matriz de covarianzas A .*

Demostración. Sea X_1, \dots, X_p una secuencia de p variables aleatorias continuas con función de densidad conjunta $f_{X_1, \dots, X_p}(x_1, \dots, x_p)$; y $y_i = g_i(x_1, \dots, x_p)$, un conjunto de funciones continuas inyectivas. Se puede despejar a cada x_i de las funciones $y_i = g_i(x_1, \dots, x_p)$ de tal forma que $x_i = h_i(y_1, y_2, \dots, y_n)$. Sea $|J|$ el determinante de la siguiente matriz:

$$\begin{pmatrix} \frac{\partial h_1}{\partial y_1} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial y_1} & \cdots & \frac{\partial h_n}{\partial y_n} \end{pmatrix}.$$

$|J|$ es llamado *acobiano* de f y se sabe del cálculo (véase [23]) que:

$$f_{Y_1, \dots, Y_p}(y_1, \dots, y_p) = f_{X_1, \dots, X_p}(h_1(y_1, \dots, y_p), \dots, h_p(y_1, \dots, y_p)) |J|.$$

Ahora se encontrará $|J|$ para una transformación $y = Cx$:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

2.5. GENERACIÓN DE NÚMEROS ALEATORIOS DEPENDIENTES 51

que es equivalente a:

$$\begin{aligned} y_1 &= a_{11}x_1 \\ y_2 &= a_{21}x_1 + a_{22}x_2 \\ &\vdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n. \end{aligned}$$

Despejando cada x_i recursivamente se obtiene lo siguiente:

$$\begin{aligned} x_1 &= \frac{1}{a_{11}}y_1 \\ x_2 &= \frac{1}{a_{22}}y_2 - \frac{a_{21}}{a_{11}a_{22}}y_1 \\ &\vdots \\ x_n &= \frac{1}{a_{nn}}y_n + t(y_1, \cdots, y_{n-1}), \end{aligned}$$

en donde $t(y_1, \cdots, y_{n-1})$ es una función obtenida al despejar $x_n - \frac{1}{a_{nn}}y_n$.

Con estas ecuaciones se puede obtener el jacobiano de la transformación:

$$|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} = \begin{vmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ \alpha_{21} & \frac{1}{a_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \frac{1}{a_{nn}} \end{vmatrix} = \frac{1}{a_{11}} \frac{1}{a_{22}} \cdots \frac{1}{a_{nn}},$$

en donde α_{ij} es una constante ya que todas las derivadas parciales son de orden uno.

Como $\det(C) = a_{11}a_{22} \cdots a_{nn}$ se concluye que $|J| = \frac{1}{\det(C)}$.

Usando el hecho de que $\det(A) = \det(C)\det(C)$, al ser f la función de distribución multinormal, X un vector aleatorio y $y = Cx$:

$$\begin{aligned}
f(y) &= f(C^{-1}y)|C|^{-1} \\
&= \frac{1}{(2\pi)^{1/2}} \frac{1}{\det(I)^{1/2}} \exp \left\{ -\frac{1}{2}(C^{-1}y)^{tr} I^{-1}(C^{-1}y) \right\} \frac{1}{\det(C)} \\
&= \frac{1}{(2\pi)^{1/2}} \exp \left\{ -\frac{1}{2}y^{tr}(C^{-1})^{tr} I(C^{-1}y) \right\} \frac{1}{\det(C)} \\
&= \frac{1}{(2\pi)^{1/2}} \frac{1}{\det(C)} \exp \left\{ -\frac{1}{2}y^{tr}(C^{-1})^{tr}(C^{-1}y) \right\} \\
&= \frac{1}{(2\pi)^{1/2}} \frac{1}{\det(A)^{1/2}} \exp \left\{ -\frac{1}{2}y^{tr}(C^{tr})^{-1}C^{-1}y \right\} \\
&= \frac{1}{(2\pi)^{1/2}} \frac{1}{\det(A)^{1/2}} \exp \left\{ -\frac{1}{2}y^{tr}(CC^{tr})^{-1}y \right\} \\
&= \frac{1}{(2\pi)^{1/2}} \frac{1}{\det(A)^{1/2}} \exp \left\{ -\frac{1}{2}y^{tr}A^{-1}y \right\},
\end{aligned}$$

por lo que Y es una variable aleatoria multinormal con media cero y matriz de covarianzas A .

□

Este método ayudará a concebir simulaciones de varias variables que son dependientes y así obtener resultados más apegados a la realidad. En la Figura 2.1 se pueden observar dos variables aleatorias con distribución $N(0, 1)$ graficadas conjuntamente pero con un coeficiente de correlación cero, por lo que en ella no se aprecia una dependencia entre las variables. En cambio en la Figura 2.2 se aprecia que usando el método de Cholesky y un coeficiente de correlación de 0.7 se forman dos variables con función de distribución normal y que existe relación entre las variables. Ahora se presenta código en lenguaje R para generar estas gráficas:

```

#Input: A = Matriz de covarianzas
#Output: resultado = Gráfica requerida
#La función chol() calcula la factorización de Cholesky
for(i in 1:1300)
{
    a<-t(chol(A))%*%rnorm(2)
    Z[i]=a[1]
    G[i]=a[2]
}
resultado<-plot(Z,G)

```

2.5. GENERACIÓN DE NÚMEROS ALEATORIOS DEPENDIENTES 53

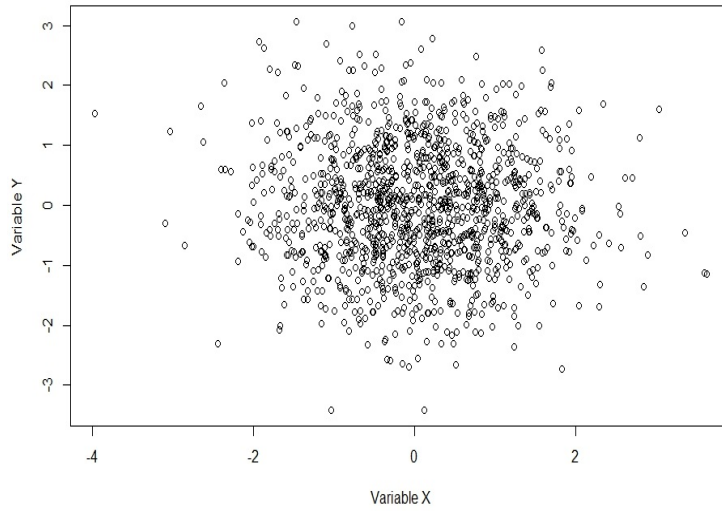


Figura 2.1: Simulación de dos variables independientes. Fuente: Elaboración propia.

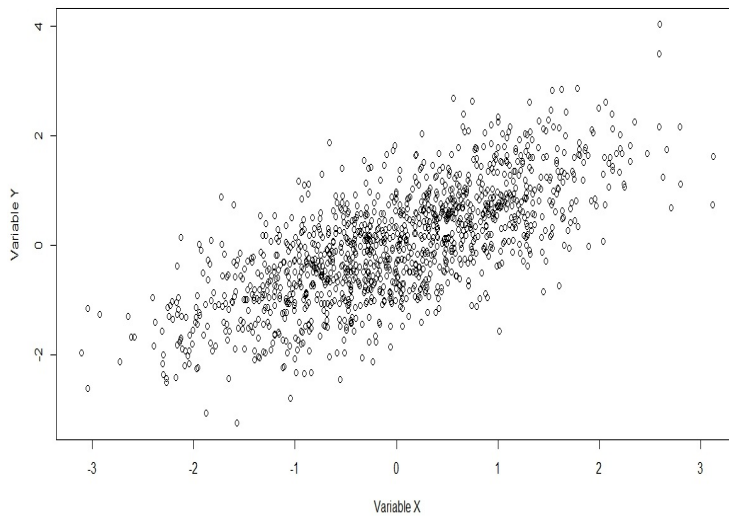


Figura 2.2: Simulación de dos variables dependientes con el uso del método de Cholesky. Fuente: Elaboración propia.

Una limitante de este método es que para poder realizar la descomposición de Cholesky, la matriz de correlaciones debe ser *positiva definida*, pues en caso contrario, las raíces cuadradas tendrán valores negativos como argumento.

2.6. Simulación con R

El lenguaje de programación R es una herramienta que ayuda a obtener soluciones a problemas estadísticos. A través del tiempo ha llegado a generar confianza en sus métodos y es actualizado constantemente por científicos de todo el mundo por medio de nuevas librerías.

2.6.1. Aspectos generales del lenguaje R

Una de las ventajas de usar R como lenguaje de programación para la evaluación de proyectos es que cuenta con funciones estadísticas pre-programadas que ayudan a obtener inferencias sobre los posibles valores que puede tomar el rendimiento del negocio.

Como se precisa en [31]:

1. R es un lenguaje de programación y un entorno para la estadística computacional, manipulación de datos y técnicas de graficación. Es un proyecto de GNU⁵ similar al lenguaje de programación estadística S y a su entorno que fueron desarrollados en los laboratorios Bell por John Chambers y otros colaboradores. Sus desarrolladores señalan que R puede ser pensado como una implementación diferente a S pero que gran cantidad de código de S puede ser utilizado en R sin muchos cambios.

2. Una de las ventajas de R es que se pueden obtener gráficas y documentos con calidad de publicación y que pueden contener nomenclatura matemática gracias a que puede funcionar conjuntamente con el entorno LaTeX.

3. El entorno de R fue pensado para obtener una plataforma planeada y coherente de tal forma que fuera fácil y no tuviera funciones específicas e inflexibles.

⁵GNU (Gnu's Not Unix) es un sistema de software que contiene herramientas con denominación libre, esto quiere decir, entre otras cualidades, que su código puede ser manipulado por el usuario.

4. Las versiones actuales de R son el resultado de varios contribuyentes de todo el mundo. R fue escrito por primera vez por Robert Gentleman y Ross Ihaka del Departamento de Estadística de la Universidad de Auckland y en lo sucesivo se les fue integrando un grupo de colaboradores a nivel mundial y que resguarda el código fuente de R para las versiones oficiales del R Project.

5. Gracias a que R forma parte de GNU, se tiene que su código es abierto y puede ser manipulado para crear nuevas funciones o crear nuevas librerías por lo que se han escrito gran cantidad de librerías con aplicaciones en diferentes ramas de la ciencia. Para el trabajo que se realizará en este documento no se necesitará de alguna librería extra a las que contiene la distribución estándar de R.

6. La calidad de los procesos internos de R pasa una serie de pruebas que realizan sus desarrolladores y estos procesos, como el resguardo del código fuente de las versiones oficiales, han obtenido certificaciones (véase [31]) y se encuentran a código abierto para la evaluación de las personas interesadas.

A continuación se presentarán brevemente algunas de las funciones que se usarán posteriormente en este documento.

2.6.2. Herramientas de simulación de R

R puede crear números pseudoaleatorios de acuerdo con diferentes métodos que el usuario puede elegir por medio de la función `RNGkind`. Sin embargo, recordemos que R usa por defecto el generador MT, mismo que se usará por tener suficiente periodo y calidad de los datos. La semilla, el valor inicial, fue obtenida aleatoriamente desde R, que la calculó automáticamente de acuerdo con la hora y el día en que se realizó el proceso. Para poder implementar el uso de una semilla específica para la simulación en R se usará la función `set.seed`.

R cuenta con un conjunto de funciones que se pueden usar para simular variables aleatorias con una distribución específica⁶. En general, se usa el prefijo `r` antes del nombre de la distribución para que R simule los números aleatorios. Para obtener otro tipo de valores de la distribución, se usan diferentes prefijos, por ejemplo, para obtener percentiles o probabilidades de la distribución, se usan los prefijos `q` y `p` respectivamente. Como ejemplo de

⁶Para mayor detalle de los métodos usados por R se recomienda revisar [30].

la sintaxis, se muestra el código para simular una función de distribución normal:

Sintaxis:

```
rnorm(n, mean = a, sd = b)
```

Argumentos:

1. **n**, número de simulaciones a realizar.
2. **a**, media de la función de distribución.
3. **b**, desviación estándar de la distribución.

Como ejemplo numérico se presenta la simulación de dos números aleatorios con función de distribución $N(10, 4)$:

```
rnorm(2, mean = 10, sd = 2)
```

Cuyos resultados arrojados son:

```
[1] 9.677161    8.489975.
```

Otro tipo de distribuciones se pueden simular bajo una sintaxis parecida, en donde sólo cambian la cantidad y tipos de parámetros que tiene cada función de distribución. Para obtener la sintaxis de las distribuciones se puede obtener ayuda de los manuales que publica el R-project o directamente de la ayuda en la consola de R.

Para simular series de tiempo, se requieren simular recursivamente varias variables aleatorias por medio de bucles. La sintaxis de los bucles en R es parecida a la del lenguaje C y otros lenguajes derivados y en especial, para este documento, se pondrá atención en la interacción de los bucles `for` y los arreglos multidimensionales.

Para ejecutar un bucle `for` en R se tiene la siguiente sintaxis:

```
for(i in a:b)
  {... },
```

en donde *a* es el entero desde el cual el contador *i* empieza a cambiar y *b* es el número entero en el que el bucle se detendrá. Por ejemplo, para simular las primeras treinta observaciones de una caminata aleatoria con valor inicial cero y cuyos valores sean guardados en un arreglo o vector, la sintáxis será:

```
for (i in 1:30)
  {
    z<-rnorm(1,mean=0,sd=1)
    xt[i+1]=xt[i]+z
  }.
```

Por medio de arreglos, los resultados de la simulación pueden ser guardados para su posterior análisis estadístico.

Capítulo 3

Aplicación a los modelos de precios

En este capítulo se exponen algunas aplicaciones de la teoría expuesta hasta el momento. Se presenta una introducción a las medidas de riesgo con el fin de mostrar la necesidad del uso de un modelo estocástico para el cálculo de la solvencia.

3.1. Medidas de riesgo

A continuación, se brinda una definición de riesgo con el fin de medirlo, complementando el análisis de las distribuciones que se obtendrán al hacer una simulación. Se muestran diferentes medidas de riesgo y se usan las características de las medidas de riesgo para exponer los alcances de cada una.

La definición de riesgo se ha discutido en los círculos actuariales durante años pero este texto no abundará al respecto y se adoptará la definición que se proporciona en [1].

Definición 3.1. *Se llama riesgo al retorno futuro de una inversión (considerado como una variable aleatoria).*

El problema de la medición del riesgo ha dado lugar a la invención de diversos tipos de medidas con características diferentes. Aunque se ha tratado de crear algún conjunto de propiedades que debieran tener las medidas de riesgo, no se ha llegado a un consenso al respecto. Existe la idea de “medida coherente de riesgo” que es una denominación que se le da a algunas medidas que cumplen con ciertas propiedades y se definirá con el fin de mostrar los

alcances de las medidas que se presentarán. La definición de medida coherente de riesgo fue propuesta por un grupo de investigadores internacionales en el año de 1997 con el fin de que los resultados de las mediciones fueran lógicos (véase [1]). No obstante, se añade que la elección de una medida de riesgo debe tomar en cuenta más factores económicos, por lo que la elección de alguna de ellas es responsabilidad del administrador del riesgo.

Definición 3.2. *Una medida de riesgo es una función $P : X \rightarrow \mathbb{R}^+$, en donde X es un conjunto de variables aleatorias (en este caso los retornos de un conjunto de inversiones dado).*

De acuerdo con [1], existen algunas propiedades deseables en una medición del riesgo con el fin de ser coherente con lo que se mide y el resultado obtenido. Se ha llegado a decir que estas propiedades son esenciales más que deseables pero sólo nos limitaremos a presentarlas y no se discutirán (véase [29]).

Definición 3.3. *Una medida coherente de riesgo es una medida de riesgo $P(\cdot)$ que satisface las siguientes condiciones:*

1. *Propiedad HP (Homogeneidad positiva): $P(\lambda x) = \lambda P(x)$ para toda variable aleatoria x y para todo real $\lambda > 0$.*
2. *Propiedad S (Subaditividad): $P(x + y) \leq P(x) + P(y)$ para todas las variables aleatorias x e y .*
3. *Propiedad M (Monotonía): Si $x < y$, entonces $P(x) < P(y)$ para todas las variables aleatorias x e y .*
4. *Propiedad T (Invariancia traslacional): $P(x + \alpha r_0) = P(x) - \alpha$ para toda variable aleatoria x , para todo real α y para toda tasa de rendimiento r_0 libre de riesgo.*

La propiedad T indica que si un monto α es invertido a una tasa libre de riesgo y es añadido a un portafolio, entonces el riesgo de éste decrece α unidades.

La propiedad S se basa en el principio de que “una fusión no genera riesgo extra”. Sin embargo no necesariamente esto ocurre en la realidad, pues, por ejemplo, al fusionarse dos bancos, uno podría tener tanta deuda que sobrepasara la máxima deuda que pudieran asumir los dos juntos.

Si se considera que un monto más grande tiene menos liquidez, entonces el riesgo de este monto grande debería ser más alto que si se tiene ese monto esparcido en unidades pequeñas, generando así la siguiente desigualdad:

$P(\lambda x) \geq \lambda P(x)$. Si se considera la propiedad de subaditividad, se tiene que $P(\lambda x) \leq \lambda P(x)$. De esta manera se obtiene la igualdad de la propiedad HP: $P(\lambda x) = \lambda P(x)$.

En el caso de la propiedad M la idea implícita es que a mayor retorno, mayor el riesgo.

Las primeras medidas de riesgo tenían la finalidad de ser fáciles de calcular más que razonables. Markowitz propuso a la varianza en el caso de una sólo inversión y al coeficiente de correlación de Pearson en el caso de un conjunto de inversiones (véase [29]). La funcionalidad de esta medida propuesta depende de la función de distribución de los retornos, pues en el caso de que ésta no sea simétrica puede que se obtengan mediciones incorrectas ya que la relación entre variables puede no ser lineal y no se puede usar para ciertas distribuciones por lo que su uso es restringido.

De acuerdo con [19], en el año de 1994 se presentó una medida cuyo uso se generalizó rápidamente en todo el mundo por su facilidad de interpretación. A continuación su definición:

Definición 3.4. *El VaR^1 es la función definida a continuación:*

$$\text{VaR}_r(X) = \inf \{x : F_X(x) \geq r\}$$

en donde X es una variable aleatoria, $0 \leq r \leq 1$ y $F_X(x)$ es la función de distribución de X .

Se puede observar que el VaR no dice mucho sobre la posible pérdida que cae afuera del intervalo de confianza. Además la elección del intervalo de confianza es subjetiva y depende del administrador del riesgo.

La razón más usual por la que el VaR es criticado es que, si bien, cumple con las propiedades T , HP y M , falla en cumplir con la subaditividad para algunos casos².

Una vez que en 1997 se presentaron las características que debían tener las medidas de riesgo para llegar a ser coherentes, se empezó a desechar al VaR como medida de riesgo por no cumplir con la propiedad S . El mundo necesitaba una medida que sustituyera al VaR y encontró refugio en la siguiente medida:

¹Valor en riesgo por sus siglas en inglés.

²Para el caso de distribuciones conjuntas elípticas la subaditividad sí se cumple.

Definición 3.5. El CVaR es la función definida a continuación:

$$\text{CVaR}_r = E[X : X > \text{VaR}_r]$$

en donde X es la variable aleatoria que representa a los rendimientos de un activo.

CVaR son las siglas en inglés de Conditional Value at Risk, que en español quiere decir valor en riesgo condicional. Se define como el valor esperado de las posibles pérdidas que exceden al VaR. El CVaR considera pérdidas catastróficas que el VaR no toma en cuenta y además es una medida coherente de riesgo [27]. Las comparaciones entre el VaR y el CVaR no se hicieron esperar desde que se presentaron las medidas coherentes de riesgo. Un trabajo al respecto y que extiende los conceptos de la funcionalidad del CVaR para la optimización de portafolios de inversión es [27].

A continuación se definen diferentes tipos de riesgo que se pueden encontrar en una compañía según se explica en [15].

1. El *riesgo de mercado* se define como la pérdida que puede sufrir un inversionista debido a la diferencia en los precios que se registran en el mercado.

2. El *riesgo de crédito* es la pérdida potencial producto del incumplimiento de la contraparte en una operación que incluye un compromiso de pago.

3. El *riesgo de liquidez* se refiere a las pérdidas que puede sufrir una institución al requerir una mayor cantidad de recursos para financiar sus actividades a un costo posiblemente inaceptable. También se le llama así a la imposibilidad de vender un activo en el mercado.

4. El *riesgo operativo* se refiere a las pérdidas posibles que son debidas a una falla en la operación. Este es un concepto muy amplio pero no se hará énfasis en este tipo de riesgo.

5. Por *riesgo sistemático* se entenderá la pérdida causada por una crisis generalizada del sistema financiero.

3.2. Movimiento geométrico Browniano

En el Capítulo 1 se definió la integral de Ito. Ahora se procederá a definir la ecuación diferencial de Ito, de la cual se obtiene un modelo de precios.

Supongamos que el fenómeno que se quiere describir es el del valor de un activo que cambia en el tiempo instantánea y aleatoriamente, y que además los cambios en el tiempo son producto de diferentes factores aleatorios que afectan constantemente el valor. Se pensaría que el modelo que describe el fenómeno es $\frac{dx}{dt} = a(x, t)$ en donde $a(x, t)$ es un coeficiente de x que puede depender del tiempo; sin embargo, este modelo no proporciona movimientos aleatorios estocásticos continuos, por lo que un modelo más adecuado puede ser de la siguiente forma:

$$\frac{dx}{dt} = a(x, t) + b(x, t)\xi_t$$

en donde $a(x, t)$ es llamada la *parte determinista*, $b(x, t)$ es un coeficiente de x que puede depender del tiempo (llamado *parte estocástica*) y ξ_t es la razón de cambio de un proceso estocástico continuo.

La interpretación de este modelo es que la razón de cambio de un proceso estocástico es igual a una función determinista más una razón de cambio de un proceso estocástico conocido multiplicado por una función determinista.

Para modelar ξ_t se puede usar el *ruido blanco*, por lo que $\xi_t = \frac{dW_t}{dt}$, de donde se obtiene la *ecuación diferencial de Ito*.

Definición 3.6. Una ecuación estocástica diferencial de Ito (EED) es:

$$dX_t = f(X_t, t)dt + G(X_t, t)dW_t$$

o en su otra forma

$$X_t = c + \int_{t_0}^t f(X_s, s)ds + \int_{t_0}^t G(X_s, s)dW_t$$

con valor inicial $X_{t_0} = X_0$.

Uno de los aspectos más importantes sobre esta ecuación es que existe un lema que ayuda a encontrar una solución a una transformación de una EED. Gracias a este lema se han realizado avances importantes en la investigación en finanzas como, por ejemplo, el cálculo del precio de opciones europeas Put a través del modelo Black-Scholes-Merton. La prueba tradicional de la existencia y unicidad de la solución se puede ver en [26] y en [22].

Lema 3.1. Sea $u = u(t, x)$ una función escalar continua definida en $[t_0, T] \subset \mathbb{R}$ con derivadas parciales continuas u_t , u_x y u_{xx} . Si X_t es un proceso definido en $[t_0, T]$ con ecuación diferencial estocástica

$$dX_t = f dt + G dW$$

en donde f , G y W_t son funciones escalares, entonces $Y_t = u(t, X_t)$ define en $[t_0, T]$ la siguiente ecuación diferencial estocástica:

$$dY_t = \left(u_t(t, X_t) + u_x(t, X_t)f(t) + \frac{1}{2}u_{xx}(t, X_t)G(t)^2 \right) dt + u_x(t, X_t)G(t)dW_t$$

en donde

$$u_x = \frac{\partial u(t, x)}{\partial x}$$

y

$$u_{xx} = \frac{\partial^2 u(t, x)}{\partial x \partial x}.$$

Demostración. El proceso inicial X_t tiene la siguiente forma:

$$X_t = X_0 + f(t - t_0) + G(W_t - W_{t_0}) \text{ con } t_0 \leq t \leq T,$$

por lo que el proceso Y_t tiene la siguiente forma:

$$Y_t = u(t, X_0 + f(t - t_0) + G(W_t - W_{t_0})).$$

Supóngase que $t_0 \leq t_1 \leq \dots \leq t_n = t \leq T$. Entonces,

$$Y_t - Y_0 = \sum_{k=0}^{n-1} (u(t_k, X_{t_k}) - u(t_{k-1}, X_{t_{k-1}})).$$

Aplicando la expansión de Taylor se obtiene:

$$\begin{aligned} u(t_k, X_{t_k}) - u(t_{k-1}, X_{t_{k-1}}) &= u_t(t_{k-1} + d_k(t_k - t_{k-1}), X_{t_{k-1}})(t_k - t_{k-1}) \\ &\quad + u_x(t_{k-1}, X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}}) \\ &\quad + \frac{1}{2}u_{xx}(t_{k-1}, X_{t_{k-1}} + \bar{d}_k(X_{t_k} - X_{t_{k-1}}))(X_{t_k} - X_{t_{k-1}})^2 \end{aligned}$$

en donde $0 < d_k$ y $\bar{d}_k < 1$. En virtud de la continuidad de X_t , u_t y u_{xx} , existen variables α_n y β_n que convergen a cero con probabilidad 1 cuando

$$\delta_n = \max_{1 \leq k \leq n} (t_k - t_{k-1}) \rightarrow 0$$

y que satisfacen las siguientes desigualdades:

$$\max_{1 \leq k \leq n} |u_t(t_{k-1} + d_k(t_k - t_{k-1}), X_{t_{k-1}}) - u_t(t_{k-1}, X_{t_{k-1}})| \leq \alpha_n$$

y

$$\max_{1 \leq k \leq n} |u_{xx}(t_{k-1}, X_{t_{k-1}} + \bar{d}_k(X_{t_k} - X_{t_{k-1}})) - u_{xx}(t_{k-1}, X_{t_{k-1}})| \leq \beta_n.$$

Sustituyendo a d_k y \bar{d}_k por cero, queda demostrar la siguiente ecuación:

$$\begin{aligned} st - \lim_{\delta_n \rightarrow 0} \sum_{k=1}^n [& u_t(t_{k-1}, X_{t_{k-1}})(t_k - t_{k-1}) \\ & + u_x(t_{k-1}, X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}}) \\ & + \frac{1}{2} u_{xx}(t_{k-1}, X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}})^2] \\ & = \int_{t_0}^t (u_t(s, X_s) + u_x(s, X_s)f \\ & + \frac{1}{2} u_{xx}(s, X_s)G^2) ds + \int_{t_0}^t u_x(s, X_s)GdW_s. \end{aligned}$$

En virtud de los supuestos de continuidad se tiene que:

$$ac - \lim_{\delta \rightarrow 0} \sum_{k=1}^n u_t(t_{k-1}, X_{t_{k-1}})(t_k - t_{k-1}) = \int_{t_0}^t u_t(s, X_s) ds$$

y

$$\begin{aligned} st - \lim_{\delta \rightarrow 0} \sum_{k=1}^n u_x(t_{k-1}, X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}}) \\ = \int_{t_0}^t u_x(s, X_s) f ds + \int_{t_0}^t u_x(s, X_s) G dW_s. \end{aligned}$$

Entonces la atención se centra en la suma

$$\sum_{k=1}^n u_{xx}(t_{k-1}, X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}})^2$$

que es igual con

$$\begin{aligned} & f^2 \sum_{k=1}^n u_{xx}(t_{k-1}, X_{t_{k-1}})(t_k - t_{k-1})^2 \\ & + 2fG \sum_{k=1}^n u_{xx}(t_{k-1}, X_{t_{k-1}})(t_k - t_{k-1})(W_{t_k} - W_{t_{k-1}}) \\ & + G^2 \sum_{k=1}^n u_{xx}(t_{k-1}, X_{t_{k-1}})(W_{t_k} - W_{t_{k-1}})^2. \end{aligned}$$

Primero se demostrará la convergencia del siguiente término:

$$st - \lim_{\delta \rightarrow 0} \sum_{k=1}^n u_{xx}(t_{k-1}, X_{t_{k-1}})(W_{t_k} - W_{t_{k-1}})^2 = \int_{t_0}^t u_{xx}(s, X_s) ds.$$

Como se sabe que

$$ac - \lim_{\delta \rightarrow 0} \sum_{k=1}^n u_t(t_{k-1}, X_{t_{k-1}})(t_k - t_{k-1}) = \int_{t_0}^t u_t(s, X_s) ds$$

se tiene que demostrar que

$$st - \lim_{\delta_n \rightarrow 0} S_n = 0$$

en donde

$$S_n = \sum_{k=1}^n u_t(t_{k-1}, X_{t_{k-1}})((W_{t_k} - W_{t_{k-1}})^2 - (t_k - t_{k-1})),$$

por lo que se procede a usar la convergencia en media cuadrática.

$$E[(S_n)^2] = \sum_{k=1}^n E[u_t(t_{k-1}, X_{t_{k-1}})^2] 2(t_k - t_{k-1})^2$$

ya que $E[(S_n)] = 0$, pues $E[(W_{t_k} - W_{t_{k-1}})^2 - (t_k - t_{k-1})] = 0$.

Entonces

$$\begin{aligned} & \sum_{k=1}^n E [u_t(t_{k-1}, X_{t_{k-1}})^2] 2(t_k - t_{k-1})^2 \\ & \leq 2 \max_{1 \leq k \leq n} \{ E [u_{xx}(t_{k-1}, X_{t_{k-1}})^2] \} \delta_n \sum_{k=1}^n (t_k - t_{k-1}) \end{aligned}$$

que tiende a cero cuando $\delta_n \rightarrow 0$

Como se ha expuesto al inicio del presente trabajo, la convergencia en media cuadrática implica la convergencia en media; y esta a su vez, la convergencia en probabilidad. Esto es, la convergencia en media cuadrática implica la convergencia en probabilidad, de donde se sigue el resultado deseado. Sólo resta demostrar que

$$\begin{aligned} & f^2 \sum_{k=1}^n u_{xx}(t_{k-1}, X_{t_{k-1}})(t_k - t_{k-1})^2 \\ & + 2fG \sum_{k=1}^n u_{xx}(t_{k-1}, X_{t_{k-1}})(t_k - t_{k-1})(W_{t_k} - W_{t_{k-1}}) \end{aligned}$$

tiende a cero.

La demostración de este último hecho requiere de un tratamiento específico entre valores grandes y pequeños que tome u_{xx} mediante un método de truncamiento. En este trabajo se presentará un resumen de la demostración como se presenta en [26]. Para ver una demostración completa de esta parte se sugiere revisar [28].

Como ejemplo se tiene que:

$$\begin{aligned} & E \left[\left(\sum_{k=1}^n u_{xx}(t_{k-1}, X_{t_{k-1}})(t_k - t_{k-1})(W_{t_k} - W_{t_{k-1}}) \right)^2 \right] \\ & = \sum_{k=1}^n E [u_{xx}(t_{k-1}, X_{t_{k-1}})^2] (t_k - t_{k-1})^3 \\ & \leq \max_{1 \leq k \leq n} \{ E [u_{xx}(t_{k-1}, X_{t_{k-1}})^2] \} \sum_{k=1}^n (t_k - t_{k-1})^2 \end{aligned}$$

que tiende a cero cuando $\delta_n \rightarrow 0$.

□

Cabe destacar que el lema anterior es una versión corta del teorema presentado en [18] por Ito.

El modelo que se verá a continuación necesitará del Lema 3.1 para mejorar los resultados de una simulación.

Definición 3.7. *El Movimiento Geométrico Browniano (MGB) es un modelo de precios que tiene la siguiente forma:*

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

en donde S_t es el precio de un activo, μ es la fuerza de interés esperada, σ es la volatilidad de S_t ; y W_t , un proceso generalizado de Wiener.

La primera parte de este modelo se basa en la ecuación diferencial determinista del rendimiento:

$$\frac{dS_t}{dt} = \mu S_t,$$

en donde μ es la fuerza de interés. Una solución de esta ecuación determinista dado un valor inicial S_0 es:

$$S_t = S_0 \exp(\mu t).$$

La filosofía del modelo MGB es la misma del modelo determinista pero con la adición del término aleatorio $\sigma S_t \frac{dW_t}{dt}$.

La solución de la ecuación diferencial estocástica del MGB se deduce del Lema 3.1 y el modelo obtenido será de utilidad para la simulación de precios que no alcanzan valores negativos ya que la distribución de la serie de tiempo generada por este modelo es una variable aleatoria lognormal.

La solución de la ecuación diferencial estocástica del MGB está dada por la siguiente función:

$$Y_t = u(t, X_t) = S_0 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma X_t \right\}$$

con $X_t = W_t$ como la forma trivial de una ecuación diferencial estocástica con $f = 0$ y $G = 1$.

Para demostrarlo se usa el Lema 3.1:

$$\begin{aligned} dY_t &= \left(\mu - \frac{\sigma^2}{2} \right) Y_t dt + \frac{\sigma^2}{2} Y_t dt + \sigma Y_t dW_t \\ &= \mu Y_t dt + \sigma Y_t dW_t. \end{aligned}$$

Consecuentemente,

$$S_t := S_0 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma X_t \right\}$$

es una solución de la ecuación diferencial estocástica del MGB.

Proposición 3.1. *Para el MGB se cumple lo siguiente:*

1. $E[S_t] = S_0 \exp(\mu t)$.
2. $Var[S_t] = S_0^2 \exp(2\mu t) (\exp(\sigma^2 t) - 1)$.
3. $Cov(S_t, S_r) = S_0^2 \exp(\mu(t+r)) (\exp(\sigma^2 r) - 1)$.

Demostración. Dado que W_t se distribuye como una normal entonces $\exp(\sigma W_t)$ se distribuye lognormal por lo que:

$$\begin{aligned} E[S_t] &= E \left[S_0 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right\} \right] \\ &= S_0 \exp \left(\mu - \frac{\sigma^2}{2} \right) t E[\exp(\sigma W_t)] \\ &= S_0 \exp \left(\mu - \frac{\sigma^2}{2} \right) t \exp \left(\frac{\sigma^2}{2} t \right) \\ &= S_0 \exp(\mu t). \end{aligned}$$

Para la varianza se tiene que

$$\begin{aligned}
\text{Var} [S_t] &= E [S_t^2] - E [S_t]^2 \\
&= S_0^2 \exp \left(2 \left(\mu - \frac{\sigma^2}{2} \right) t \right) E [\exp (2\sigma W_t)] - S_0^2 \exp (2\mu t) \\
&= \left(2 \left(\mu - \frac{\sigma^2}{2} \right) t \right) E [\exp (2\sigma W_t)] - S_0^2 \exp (2\mu t) \\
&= S_0^2 \left[\exp \left(2 \left(\mu - \frac{\sigma^2}{2} \right) t \right) \exp \left(\frac{(2\sigma)^2}{2} t \right) - \exp (2\mu t) \right] \\
&= S_0^2 [\exp (2\mu t - \sigma^2 t) \exp (2\sigma^2 t) - \exp (2\mu t)] \\
&= S_0^2 [\exp (2\mu t + \sigma^2 t) - \exp (2\mu t)] \\
&= S_0^2 [\exp (2\mu t)(\exp (\sigma^2 t) - 1)].
\end{aligned}$$

Para la covarianza se tiene que:

$$\begin{aligned}
\text{Cov}(S_t, S_r) &= E \left[S_0^2 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) (t+r) + \sigma(W_t + W_r) \right\} \right] - E [S_t] E [S_r] \\
&= S_0^2 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) (t+r) \right\} E [\exp (\sigma(W_t + W_r))] - S_0^2 \exp (\mu(t+r)) \\
&= S_0^2 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) (t+r) \right\} E [\exp (2\sigma W_r)] E [\exp (\sigma(W_t - W_r))] \\
&\quad - S_0^2 \exp (\mu(t+r)) \\
&= S_0^2 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) (t+r) \right\} \exp (2\sigma^2 r) \exp \left(\frac{1}{2} \sigma^2 (t-r) \right) \\
&\quad - S_0^2 \exp (\mu(t+r)) \\
&= S_0^2 \exp \{ \mu(t+r) + r\sigma^2 \} - S_0^2 \exp (\mu(t+r)) \\
&= S_0^2 \exp \{ \mu(t+r) \} (\exp (\sigma^2 r) - 1).
\end{aligned}$$

□

La esperanza resulta ser la ecuación diferencial del valor presente. La distribución lognormal, obtenida de la solución de la ecuación estocástica diferencial MGB, está definida para cada t por su esperanza y su varianza, lo que permitirá obtener intervalos de confianza para la serie de tiempo simulada.

Para simular un MGB se requiere simular el ruido blanco, para lo cual existe el siguiente método:

$\Delta W = \epsilon\sqrt{\Delta t}$ en donde $\epsilon \sim N(0, 1)$, derivado de que $\Delta W^2 = \Delta t$.

Así, el MGB se puede simular llevándolo a su forma discreta, valorando el modelo en n intervalos equidistantes de tamaño $\frac{1}{n}$:

$$S_{t+1} - S_t = \mu S_t \frac{1}{n} + \sigma S_t \epsilon_t \sqrt{\frac{1}{n}}.$$

Despejando se sigue que

$$S_{t+1} = S_t \left(1 + \mu \frac{1}{n} + \sigma \epsilon_t \sqrt{\frac{1}{n}} \right),$$

con los valores iniciales S_{t_0} y $t_0 = 0$. Al implementar este modelo se obtiene un rendimiento con distribución normal, lo que conlleva a la posible obtención de rendimientos negativos, los cuales pueden no ser convenientes en algunos casos. Para descartar este tipo de variables se puede usar la transformación $\ln(S_t)$ y usar el lema 3.1 para obtener una solución. Primero se reconocerán los elementos del lema: $f = \mu S$, $G = \sigma S$, $\frac{\partial u}{\partial S} = \frac{1}{S}$ y $\frac{\partial^2 u}{\partial S^2} = -\frac{1}{S^2}$. Sustituyendo estos valores se obtiene la siguiente ecuación:

$$d\ln(S_t) = \left(\mu - \frac{1}{2}\sigma^2 \right) dt + \sigma dz.$$

La forma discreta de esta fórmula es la siguiente:

$$\ln(S_{t+1}) - \ln(S_t) = \left(\mu - \frac{1}{2}\sigma^2 \right) \frac{1}{n} + \sigma \epsilon_t \sqrt{\frac{1}{n}}.$$

Despejando a S_t , se obtiene la siguiente variable:

$$S_{t+1} = \exp \left\{ \left(\mu - \frac{1}{2}\sigma^2 \right) \frac{1}{n} + \sigma \epsilon_t \sqrt{\frac{1}{n}} \right\} S_t,$$

que es la solución de la ecuación diferencial estocástica del MGB vista previamente. Esta variable se distribuye como una lognormal ya que

$$\left(\mu - \frac{1}{2}\sigma^2 \right) \frac{1}{n} + \epsilon_t \sqrt{\frac{1}{n}},$$

se distribuye como una variable aleatoria normal.

En la Figura 3.1 se aprecia la simulación de un precio. En este caso se simularon 1000 periodos y se establecieron como parámetros un valor esperado de 0.03 de rendimiento al final del periodo y una volatilidad de 0.01 (en el anexo B.2 se encuentra el código para generarla).

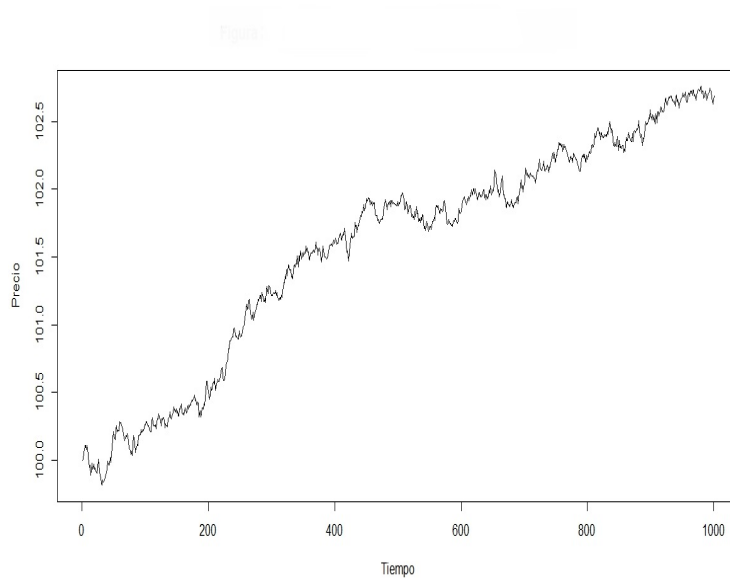


Figura 3.1: Simulación de un MGB. Fuente: Elaboración propia.

A lo largo de la simulación se pueden apreciar cambios de tendencia y el efecto del movimiento Browniano, pero el resultado es cercano al valor esperado del rendimiento.

La dinámica de las tasas de interés es diferente a la de los precios ya que las tasas de interés deben de cumplir con características propias por lo que a continuación se mostrarán características que sugieren el uso de modelos diferentes al movimiento geométrico Browniano. Sea $P(t, s)$ el precio de un bono en el tiempo t y con plazo s , y con $t \leq s$ con un valor unitario al vencimiento:

$$P(s, s) = 1.$$

La tasa de retorno $R(t, T)$ que se conseguirá desde el tiempo t hasta el tiempo $s = t + T$ se puede representar de la siguiente forma de acuerdo con la fórmula del valor presente:

$$R(t, T) = -\frac{1}{T} \log(P(t, t + T)),$$

con $T > 0$. Una función que establece la tasa de interés en términos de T es llamada estructura del plazo en t . La problemática se presenta cuando se requiere saber la tasa futura dependiendo del precio del bono; sin embargo, la tasa futura $F(t, T)$ debe cumplir que

$$R(t, T) = \frac{1}{T} \int_t^{t+T} F(t, \tau) d\tau,$$

y esta ecuación tiene la siguiente solución diferencial:

$$F(t, s) = \frac{\partial}{\partial s} ((s - t) R(t, s - t)).$$

La tasa al momento t debe ser entonces:

$$r(t) = R(t, 0) = \lim_{\tau \rightarrow 0} R(t, \tau).$$

En 1977 Vasicek asume que el precio del bono $P(t, s)$ al tiempo t debe ser valorado de acuerdo con la tasa de interés al momento t en $\{r(t), t \leq \tau \leq s\}$ (véase [32]). Este supuesto es justificado con la proposición de que la tasa al momento t sigue una ecuación diferencial estocástica de Ito; el precio durante el plazo $s - t$ está determinado por la tasa al momento t . Bajo estos supuestos, la ecuación de los rendimientos queda de la siguiente forma:

$$R(t, T) = -\frac{1}{T} \log (P(t, t + T, r(t))).$$

Así, Vasicek encuentra que se debe hacer énfasis en los modelos de tasas de interés para modelar precios de bonos. Bajo esta lógica propone usar el modelo físico de Ornstein-Uhlenbeck. El modelo es el siguiente:

$$dr_t = k(\theta - r_t) dt + \sigma dW_t,$$

en donde $\theta > 0$ determina la tasa de interés a largo plazo y $k > 0$ la fuerza con que r_t regresa a la tasa a largo plazo, razón por la cual se le da el nombre de modelo de tasa de interés con factor de media revertida. La ventaja de este modelo, de acuerdo con [32], es que es un modelo estacionario por lo que no diverge al evaluarse por largos periodos de tiempo. Vasicek no describe más ventajas sobre este modelo. Usando la fórmula de Ito respecto a una función $f(r_t) = \exp\{kt\} r_t$, se puede obtener una solución a la ecuación estocástica diferencial. Sustituyendo en el resultado del Lema 3.1 se sigue que:

$$d(\exp\{kt\} r_t) = \exp\{kt\} k\theta dt + \sigma \exp\{kt\} dW_t.$$

Integrando de ambos lados:

$$\exp\{kt\} r_t = r_0 + \int_0^t \exp\{ks\} k\theta ds + \int_0^t \sigma \exp\{ks\} dW_s,$$

esto es,

$$r_t = r_0 \exp \{-kt\} + \theta (1 - \exp \{-kt\}) + \sigma \int_0^t \exp \{k(t-s)\} dW_s.$$

Como la integral es evaluada desde $t = 0$, se tiene que:

$$E[r_t] = r_0 \exp \{-kt\} + \theta (1 - \exp \{-kt\}).$$

Para la covarianza se tiene el siguiente resultado para $s \leq t$:

$$\begin{aligned} \text{Cov}[r_s, r_t] &= E[r_s - E[r_s]] E[r_t - E[r_t]] \\ &= E \left[\left(\sigma \int_0^s \exp \{-k(t-p)\} dW_p \right) \left(\sigma \int_0^t \exp \{-k(t-q)\} dW_q \right) \right] \\ &= \sigma^2 \exp \{-k(s+t)\} E \left[\int_0^s \exp \{k(t-p)\} dW_p \int_0^t \exp \{k(t-q)\} dW_q \right]. \end{aligned}$$

Usando la isometría de Ito y excluyendo al límite de la notación se obtienen los siguientes resultados:

$$\begin{aligned} \text{Cov}[r_s, r_t] &= \sigma^2 \exp \{-k(s+t)\} E \left[\sum_{p=0}^s \exp \{k(t-p)\} \Delta W_p \sum_{q=0}^t \exp \{k(t-q)\} \Delta W_q \right] \\ &= \sigma^2 \exp \{-k(s+t)\} \sum_{p=0}^s \exp \{k(t-p)\} E[\Delta W_p] \sum_{q=0}^t \exp \{k(t-q)\} E[\Delta W_q] \\ &= \sigma^2 \exp \{-k(s+t)\} \sum_{p=0}^s \exp \{k(t-p)\} E[\Delta W_p] \exp \{k(t-p)\} E[\Delta W_p] \\ &= \sigma^2 \exp \{-k(s+t)\} \sum_{p=0}^s \exp \{2k(t-p)\} E[\Delta W_p]^2 \\ &= \sigma^2 \exp \{-k(s+t)\} E \left[\int_0^s \exp \{2k(t-p)\} dp \right] \\ &= \frac{\sigma^2}{2k} \exp \{-k(s+t)\} (\exp \{2ks\} - 1). \end{aligned}$$

La varianza se derivará fácilmente de esta fórmula con $s = t$:

$$\begin{aligned} \text{Var}[r_t] &= \frac{\sigma^2}{2k} \exp \{-2kt\} (\exp \{2kt\} - 1) \\ &= \frac{\sigma^2}{2k} (1 - \exp \{-2kt\}). \end{aligned}$$

Si se considera que $t \rightarrow \infty$, se puede observar que el modelo lleva a la tasa de interés a ser estacionaria en el largo plazo.

3.3. Modelos de salto

En la vida real se han presentado ocasiones en las que los cambios de tendencia son dramáticos y estos cambios han sido estudiados en diferentes precios de activos. Para algunos modelos se ha concluido que la distribución de los precios en un tiempo t contiene colas largas y probabilidades acumuladas diferentes a las que el movimiento geométrico Browniano considera, por lo que se han creado modelos que cambian los parámetros en el tiempo y consideran cambios de tendencia denominados “saltos” y así se consideran escenarios más apegados a la realidad. El primer modelo que se considerará combina el movimiento geométrico Browniano con un proceso Poisson compuesto, después se presentarán modelos de volatilidad estocástica.

En un proceso Poisson los tiempos en que ocurren los saltos se denotarán como $\tau_1 < \tau_2 < \tau_3, \dots$

Para contar el número de saltos hasta el tiempo t , se usará la variable J_t . Supongamos que se divide el tiempo de observación en n intervalos iguales de tiempo, esto es, de tamaño $\frac{t}{n}$. El resultado de un experimento es 1 si ocurre un salto dentro del intervalo de tiempo y 0 si no ocurre. El experimento puede ser descrito como una variable Bernoulli de la siguiente forma:

$$\begin{aligned} P\left(J_t - J_{t-\frac{t}{n}} = 1\right) &= \lambda \frac{t}{n} \\ P\left(J_t - J_{t-\frac{t}{n}} = 0\right) &= 1 - \lambda \frac{t}{n} \end{aligned}$$

para alguna λ tal que $0 < \lambda \frac{t}{n} < 1$. λ es llamada *intensidad del proceso*. De estas definiciones se obtiene lo siguiente para un número k :

$$P(J_t - J_0 = k) = \binom{n}{k} \left(\lambda \frac{t}{n}\right)^k \left(1 - \lambda \frac{t}{n}\right)^{n-k}.$$

Cuando n tiende a infinito, el modelo tiende a una distribución Poisson:

$$\frac{(\lambda t)^k}{k!} \exp(-\lambda t).$$

Definición 3.8. El proceso estocástico $\{J_t, t > 0\}$ es llamado proceso de Poisson si se cumplen las siguientes condiciones:

1. $J_0 = 0$.
2. $J_t - J_s$ es un entero tal que $0 \leq t < s < \infty$ y

$$P(J_t - J_s = k) = \frac{(\lambda(t-s))^k}{k!} \exp(-\lambda(t-s))$$

para $k = 0, 1, 2, \dots$.

3. Los incrementos $J_{t_2} - J_{t_1}$ y $J_{t_4} - J_{t_3}$ son independientes para todo $0 \leq t_1 < t_2 < t_3 < t_4$.

Supongamos que el precio S_t salta en $t = \tau$ y denotemos con $S_{\tau+}$ al instante justo después de que ocurra el salto y con $S_{\tau-}$ al instante anterior. Bajo estas definiciones se obtiene:

$$\Delta S = S_{\tau+} - S_{\tau-}.$$

Si se interpreta el cambio como una proporción del precio anterior, el modelo puede verse de la siguiente forma:

$$S_{\tau+} = qS_{\tau-},$$

con $q > 0$. Suponer que q es constante en todos los saltos no sería congruente, por lo que se tendrán que generar variables aleatorias independientes e idénticamente distribuidas q_{τ_i} ($i = 1, 2, \dots$), llegando a la siguiente ecuación:

$$dS_t = (q_t - 1)(S_{t-}) dJ_t,$$

en donde J_t es un proceso Poisson. Al sumarle a un proceso de movimiento geométrico Browniano un proceso de salto Poisson, se obtiene la siguiente ecuación:

$$dS_t = (S_{t-}) [\mu dt + \sigma dW_t + (q_t - 1) dJ_t].$$

Si se aplica la solución de la ecuación diferencial del movimiento geométrico Browniano para el intervalo (τ_j, τ_{j+1}) , se obtiene:

$$S_t = S_0 + \int_0^t S_s (\mu ds + \sigma dW_s) + \sum_{j=1}^{J_t} S_{\tau_j^-} (q_{\tau_j} - 1).$$

Usando el cambio de variable $Y_t = \log(S_t)$ se puede obtener $Z_t = \log(qS_{\tau-}) - \log(S_{\tau-})$ y una solución de la siguiente forma:

$$Y_t = Y_0 + \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t + \sum_{j=1}^{J_t} Z_{\tau_j}.$$

3.4. Modelos de volatilidad

El movimiento geométrico Browniano tiene una distribución conocida en el tiempo t lo que significa una ventaja para el cálculo del riesgo. Sin embargo, en la práctica los parámetros del modelo no son constantes y esto ocasiona que la distribución de los activos no sea la misma que la que propone el modelo. El principal problema al suponer una volatilidad constante es que el riesgo depende en gran medida de este parámetro. A continuación se presentan modelos de volatilidad estocástica, sugeridos en tiempos recientes y que intentan resolver el problema. Los primeros estudios sobre modelos estocásticos de volatilidad se presentaron por Hull y White en 1987 (véase [33] para más referencias) en donde proponen modelar la volatilidad con un movimiento geométrico Browniano:

$$dS_t = \mu S_t dt + \sigma_t S_t dW_t$$

y

$$d\sigma_t^2 = \beta \sigma_t^2 dt + v \sigma_t^2 dW'_t,$$

en donde W_t y W'_t se encuentran correlacionadas con coeficiente de correlación ρ . A partir de este modelo se obtuvieron otros como el de volatilidad de media revertida que es de la siguiente forma:

$$d\sigma_t = k(\theta - \sigma_t) dt + v dW'_t,$$

donde los parámetros θ y k son la volatilidad esperada a largo plazo y la fuerza de reversión del modelo respectivamente.

Simular la volatilidad como un proceso estocástico permite obtener volatilidades dinámicas. Sin embargo, los cambios sistemáticos de las variables económicas determinan diferentes comportamientos predecibles a lo largo del periodo de observación, por lo que existe la necesidad de contar con un modelo que siga cierta lógica en el tiempo y un modelo de series de tiempo puede ayudar a representar las autocorrelaciones de la volatilidad. Los modelos de series de tiempo que se mostrarán para la volatilidad son el de promedios

móviles y el Garch que contemplan la heteroscedasticidad, esto es, la varianza de la serie tiene cambios sistemáticos en el tiempo.

El modelo de *promedios móviles* se basa en el estimador del segundo momento para la varianza, pero le asigna un peso a los últimos valores presentados en la serie de tiempo. La forma en que se asignan los pesos es llamada *Half-Life* y es usada usualmente para modelar el decaimiento radiactivo. Supóngase que la estimación del segundo momento usa la muestra de las últimas n observaciones:

$$\sigma = \left\{ \frac{1}{n} \sum_{i=1}^n (r_i - \mu)^2 \right\}^{\frac{1}{2}},$$

en donde r_i es el rendimiento del activo. μ es el primer momento de los rendimientos, pero en lo sucesivo se considerará cero como se recomienda en [33] y en [28] ya que al dividir entre n , la estimación tiende a cero. Sustituyendo el peso uniforme $\frac{1}{n}$ por un peso que se degrada a medida que la observación del rendimiento es más antigua, $(\lambda^{i-1}(1-\lambda))$, se obtiene el siguiente modelo:

$$\begin{aligned} \sigma_t^2 &= \sum_{i=1}^n (1-\lambda) \lambda^{i-1} r_{t-i}^2 \\ &= (1-\lambda) \lambda^{(1-1)} r_{t-1}^2 + \sum_{i=2}^n (1-\lambda) \lambda^{i-1} r_{t-i}^2 \\ &= (1-\lambda) r_{t-1}^2 + \lambda \sum_{i=2}^n (1-\lambda) \lambda^{i-2} r_{t-i}^2 \\ &= (1-\lambda) r_{t-1}^2 + \lambda \sigma_{t-1}^2. \end{aligned}$$

Esto es, se tiene un modelo de promedios móviles para la volatilidad. El parámetro λ se puede encontrar observando la gráfica de autocorrelaciones de la serie y con ayuda de un algoritmo de optimización que ajuste el parámetro a un rango de error deseado. El siguiente modelo es llamado comúnmente *Garch* por sus siglas en inglés (*General Autoregressive Conditionally Heteroscedastic*). Se trata de modelar la volatilidad con un ARMA usando la experiencia de los valores pasados, pero también de los errores vistos como ruido blanco. La versión ARMA(1,1) es la siguiente:

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2,$$

en donde ω , α y β son constantes mayores a cero y ϵ es ruido blanco. Los parámetros necesitan de una condición adicional a las de los ARMA con el

fin de apegarse más a la realidad. Esta condición es que $\alpha + \beta < 1$, con lo cual el proceso es estacionario.

3.5. Ejemplo de aplicación de los modelos

Las aplicaciones de la teoría expuesta abarcan los precios de activos accionarios, tasas de interés o dividendos, entre otros. Una aplicación interesante es la de los precios de los alimentos pues afectan principalmente a la población de escasos recursos, lo cual se traduce en problemas de desnutrición y en otros problemas de carácter social. El número de personas que cuentan con los recursos suficientes para comprar los alimentos necesarios para una dieta nutritiva básica, es una de las formas de medir la pobreza. La pobreza afecta a millones de personas en el mundo y en su forma más grave, la miseria por falta de alimentación, afecta diariamente a personas que no pueden obtener recursos suficientes por medio de un trabajo. La volatilidad de los precios de los alimentos y la inflación afecta de forma determinante a millones de personas que viven cerca o por debajo de la línea de pobreza (véase [10]) y el seguimiento de estos indicadores podrá determinar cambios en la población afectada y en la evolución de políticas implementadas para combatir la pobreza. Bajo este pensamiento, se decidió usar la variable económica del precio de la canasta básica para mostrar cómo se aplican los modelos a variables reales y que la teoría de los modelos de precios no sólo está restringida a los modelos de los precios bursátiles.

La hipótesis a probar con el ejemplo es que no necesariamente un humano que cuenta con un ingreso mensual mayor al de la canasta alimentaria rural (véase [8]) se encuentra fuera del segmento de población que pasa hambre ya que la volatilidad de sus ingresos diarios afecta su consumo de alimentos diarios. Para modelar esto, se simulará el ingreso diario de un microempresario, por ejemplo, el de un vendedor ambulante de artesanías. En caso de que se requiera comprobar la existencia de este segmento de la población y la importancia de su estudio, se pueden consultar [5] y [11].

Así, el objetivo de la simulación será encontrar el tiempo en el que el microempresario tarda en llegar a pasar un día con hambre, para lo cual se simulará el precio diario de la canasta básica alimentaria rural de acuerdo con un proceso geométrico Browniano, en el que sus parámetros serán simulados con un modelo de series de tiempo ya que se quiere contemplar la ciclicidad mensual del precio de la canasta básica y su tendencia de crecimiento, pues se ha observado que la inflación de los precios en México contiene ciclicidad

periódica como se precisa en [12]. Para simular los ingresos del microempresario, se usará una distribución Poisson con una media de 45 pesos de ingreso diario que, al multiplicarse por treinta, claramente se traducirá en un ingreso mensual mayor al costo mensual de la canasta básica rural. El supuesto sobre el ingreso diario del microempresario es arbitrario ya que se desconoce la existencia de algún estudio sobre la distribución del ingreso diario de un microempresario ya que al ser tan volátiles, éstos cambian significativamente por día y por tipo de microempresario como se precisa en [5]. En el caso de México, no se ha realizado un muestreo sobre los ingresos diarios de un tipo de microempresario en específico y en caso de existir, la estimación tendría un error grande gracias a la volatilidad de los ingresos de este sector. En [5] se realizó un estudio al respecto a nivel mundial, en el que se pone como ejemplo el testimonio de una microempresaria y sus ingresos diarios, pero no se tomó una muestra diaria representativa sobre sus ingresos día a día. Sin embargo, en este caso se puede observar que la simulación ayudará a mostrar cómo es el día a día de las personas que viven con un ingreso similar que, de otra forma, no se podría observar a falta de una muestra representativa.

Según el párrafo anterior, será necesario obtener el ingreso diario de un micro empresario y éste se modelará como una variable aleatoria ya que un ingreso diario constante no reflejaría la realidad de sus ingresos ya que estos cambian día a día como se mencionó anteriormente. En el caso del precio diario de la canasta alimentaria es necesario considerarlo como una variable aleatoria ya que cuenta con cambios diarios en los precios de sus elementos como se comentó anteriormente, es por eso que se requiere un modelo que arroje precios diarios aleatorios pero que considere la volatilidad y el crecimiento que el precio de la canasta básica espera tener de acuerdo al comportamiento histórico de la variable. Bajo este pensamiento, el modelo de movimiento geométrico Browniano sería adecuado ya que sus parámetros contemplan la volatilidad y el crecimiento esperados. Un modelo que contenga otras variables económicas, necesitaría la prueba de diferentes supuestos de correlación entre las variables y la implementación de modelos económicos que no forman parte de este trabajo. El modelo de movimiento geométrico Browniano tiene la ventaja de que su volatilidad permite contemplar todos los escenarios que se requieran y un modelo que contemple otras variables macroeconómicas necesitaría de un estudio extra sobre el futuro de estas y se tendría que contemplar el error implícito en estos supuestos. Por otra parte, con el uso del movimiento geométrico Browniano, se pueden considerar sus parámetros como variables aleatorias y en específico, como procesos estocásticos que pueden ser estimados bajo los métodos vistos anteriormente. Esto permitirá contemplar parámetros del modelo acorde al comportamiento

cíclico de los precios y esto proporcionará una ventaja con respecto a otros modelos que sólo consideran parámetros calculados de acuerdo al promedio de los últimos meses. En caso de que se quisiera considerar un modelo no lineal se necesitaría el uso de métodos que requieren de gran número de datos históricos como es el caso de las redes neuronales de más de una capa.

El tiempo de evaluación del proyecto será de diez meses en los cuales se podrá apreciar el riesgo de que el microempresario no cuente con recursos para alimentarse en algún momento. En caso de que el microempresario cuente con excedentes respecto a la canasta alimentaria, éstos se supondrán gastados inmediatamente y no corresponderán a un ahorro ya se asume que la persona cuenta con otros gastos. Por supuesto, esto puede cambiarse si se reproduce el ejercicio.

La aplicación comprenderá los cuatro pasos de Box y Jenkins (véase el Capítulo 2): en el primero, se describen las variables y se muestran los datos a utilizar así como la elección del modelo; en el segundo, se calculan los parámetros por medio del método de máxima verosimilitud; en el tercero, se verifica el modelo para lo cual se realizan pruebas de hipótesis sobre los supuestos; y en el cuarto, se simulará el modelo para medir el riesgo inherente con la volatilidad de los precios.

3.5.1. Descripción de la variable

Como se ha mencionado, la variable será el costo de la canasta alimentaria rural necesaria para que un microempresario pueda contar con una alimentación adecuada para vivir sanamente. La variable es calculada por el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) y publicada en [6]. Para su cálculo se usan variables como el costo de los alimentos, las costumbres de alimentación y recomendaciones médicas para que la canasta contenga los alimentos necesarios para que el consumidor pueda llevar su vida sanamente (para ver la metodología a detalle véase [8]). El CONEVAL calcula el precio de dos tipos de canasta alimentaria que son la rural y la urbana; la diferencia radica en los precios y costumbres diferentes en las áreas rurales con respecto a las urbanas. En este trabajo sólo se tomará en cuenta la canasta alimentaria rural. En el caso de la variable que representa el ingreso del microempresario, como se justificó anteriormente, se simulará una variable aleatoria Poisson con media 45.

La canasta alimentaria es construida con base en la ingesta energética y de nutrientes de la población mexicana y la información sobre esta ingesta proviene de tres fuentes. La primera fuente es el *conjunto de tablas de*

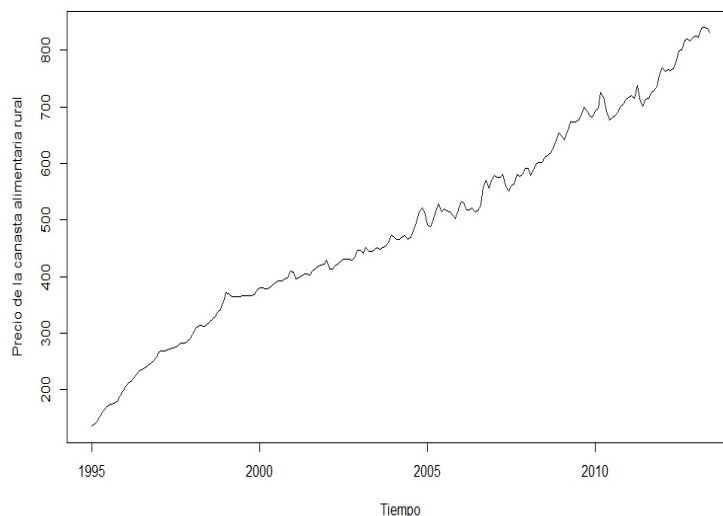


Figura 3.2: Precio de la canasta alimentaria rural. Fuente: Elaboración propia con datos del CONEVAL (véase [6]).

aportaciones nutricionales de los alimentos (véase [8]). La segunda fuente son documentos que contienen recomendaciones sobre los requerimientos energéticos emitidas por la Food and Agriculture Organization of the United Nations (FAO) y otras instituciones. La tercera fuente es la Encuesta Nacional de Ingresos y Gastos en los Hogares (ENIGH) del año 2006 en la que se encuentran las costumbres alimentarias del país. En general, se selecciona un estrato poblacional que cuente con las características de nutrición recomendadas y una vez que se encuentra este estrato, se observan las cantidades y tipos de alimentos consumidos por el mismo. Una vez construida la canasta alimentaria rural, se necesitará saber cuántas personas -que vivan en comunidades rurales- pueden acceder a ésta y se procederá a cuantificar el costo mes a mes de todos los alimentos incluidos en ésta. La evolución de la variable desde enero de 1995 hasta junio del año 2013 se muestra en la Figura 3.2.

Con una exploración visual se pueden apreciar tendencias temporales y cambios en la volatilidad, pero para ver con mayor detalle la variable en la Figura 3.3 se graficó el crecimiento porcentual.

Como la variable contiene cambios en las tendencias de crecimientos en diferentes lapsos de tiempo, no sería apropiado usar toda la serie para obtener

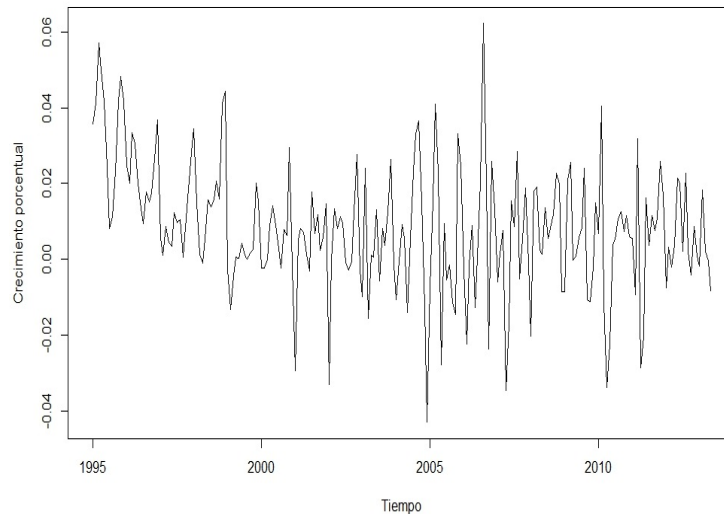


Figura 3.3: Crecimiento porcentual del precio de la canasta alimentaria rural. Fuente: Elaboración propia con datos del CONEVAL (véase [6]).

un modelo sobre el futuro inmediato (de uno a cuatro meses) del comportamiento de la variable, por lo que para el cálculo de los parámetros sólo se usará el periodo comprendido entre enero del año 2009 y el mes de junio de 2013. La elección de este intervalo se justifica en tener datos recientes y que no han sido influenciados por eventos de gran impacto en la economía como la crisis económica del año 2008. En la Figura 3.4 se muestra el coeficiente de autocorrelación de los crecimientos porcentuales de la serie histórica y servirá para seleccionar el modelo a utilizar.

Se puede apreciar que la variable no se encuentra significativamente autocorrelacionada con meses anteriores y sólo los atrasos uno, siete, ocho, nueve, once y doce muestran una autocorrelación alta según las bandas del 20% de autocorrelación, por lo que un modelo coherente debería usar a lo más los periodos señalados. El mes doce es altamente autocorrelacionado ya que muestra la tendencia temporal del mes de observación. En la Figura 3.5 se muestra la autocorrelación de la volatilidad o crecimiento por periodo del crecimiento porcentual.

La volatilidad cuenta con autocorrelaciones bajas, pero se puede apreciar un patrón de autocorrelación en los atrasos uno, dos y doce.

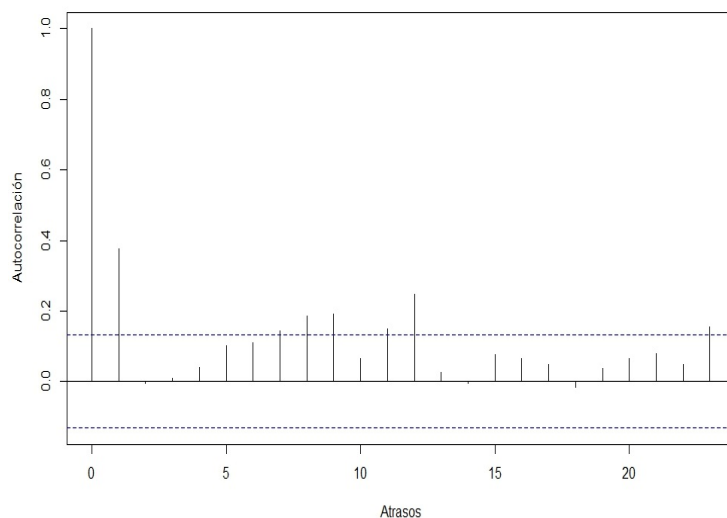


Figura 3.4: Autocorrelación del crecimiento porcentual del precio de la canasta alimentaria rural. Fuente: Elaboración propia con datos del CONEVAL (véase [6]).

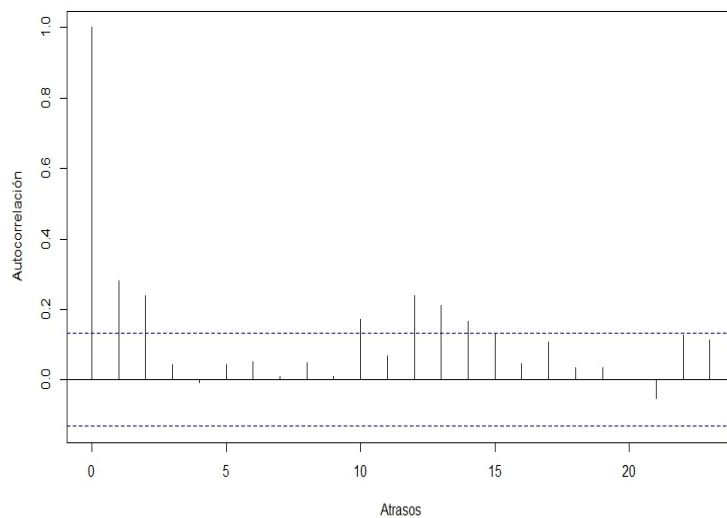


Figura 3.5: Autocorrelación de la volatilidad del precio de la canasta alimentaria rural. Fuente: Elaboración propia con datos del CONEVAL (véase [6]).

3.5.2. Estimación de los parámetros y verificación de los modelos

Para calcular los parámetros se considerarán los métodos de los momentos y de máxima verosimilitud. El método de los momentos será usado para el cálculo tradicional de los parámetros del movimiento geométrico Browniano y para verificar estos parámetros se usarán intervalos de confianza suponiendo normalidad del comportamiento de los parámetros. Por último se realizarán pruebas de estacionariedad y normalidad de los residuos de los modelos autoregresivos usados para calcular la volatilidad y el crecimiento porcentual en el siguiente periodo. El crecimiento porcentual del precio es calculado con la siguiente fórmula:

$$x_i = \frac{(V_i - V_{i-1})}{V_i},$$

en donde V_i es un valor observado del precio y n es el tamaño de la muestra. La volatilidad se calcula con la siguiente fórmula:

$$v_i = x_i - x_{i-1}.$$

Para una muestra que considera datos desde el mes de enero del año 2009 hasta el mes de junio del 2013, la prueba *Jarque-Bera* de normalidad de la variable de crecimiento porcentual, arroja un valor p de 0.024 por lo que se tomará como cierto el supuesto de normalidad. La variable de volatilidad cuenta con un valor p de 2.2×10^{-16} , por lo que también se puede suponer normalidad (el código en R para realizar estas pruebas se puede ver en el anexo B.4). Para calcular los parámetros del modelo con el método de los momentos se usan los siguientes estimadores:

$$\bar{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

y

$$\bar{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n-1} (x_i - \bar{\mu})^2.$$

Un intervalo de confianza se calcula suponiendo normalidad y calculando el promedio y la desviación estándar de la serie. Con estos datos se obtiene el percentil 0.9. Este cálculo arroja el intervalo mostrado en el Cuadro 3.1.

Uno de los supuestos más fuertes sobre los modelos es que la serie a evaluar es estacionario, por lo que se procede a realizar la prueba *Dickey-Fuller* de estacionariedad sobre la tasa de crecimiento porcentual y sobre la

Cuadro 3.1: Estimación de los intervalos de confianza de los parámetros.

Parámetro	Estimación	Intervalo de confianza al 90 %
Crecimiento porcentual	0.004584063	± 0.01856
Volatilidad	0.014478814	± 0.01715

volatilidad. Los resultados se muestran en el Cuadro 3.2. El código en R para realizar estas pruebas se puede ver en el anexo B.6.

Cuadro 3.2: Pruebas de estacionariedad.

Variable	Estadístico Dickey Fuller	Valor p
Crecimiento porcentual	-5.2347	0.01
Volatilidad	-5.187	0.01

Con los valores obtenidos en el Cuadro 3.2, se puede concluir que las series son estacionarias por lo que se podrán usar modelos AR para su simulación.

La siguiente estimación consta de dos partes: en la primera se estiman los parámetros de un modelo AR(12)³ para el crecimiento porcentual y un AR(2) para la volatilidad; en la segunda, se realiza un pronóstico del siguiente periodo para los parámetros y así se obtiene una estimación de los parámetros del movimiento geométrico Browniano. Los parámetros de los modelos AR(12) y AR(2) se presentan en el anexo B.3 y se calcularon usando la metodología expuesta en el capítulo anterior en el tema que trata sobre la estimación con el método de máxima verosimilitud. Los datos utilizados para las estimaciones comprenden los precios desde febrero de 1995 hasta junio del 2013. El código del programa en R para realizar este cálculo se encuentra en el anexo B.4.

Cuadro 3.3: Pruebas de normalidad y estacionariedad.

Prueba	Crecimiento porcentual	Volatilidad
Jarque-Bera	0.01	0.01
Shapiro-Wilk	0.007369	5.005×10^{-06}
Dickey-Fuller	7.752×10^{-5}	0.01

En el Cuadro 3.3 se resumen los valores p de las pruebas de normalidad y estacionariedad de los residuos de los modelos. Con los modelos obtenidos,

³La elección del número de los retrasos se realizó de acuerdo con las autocorrelaciones de estas variables.

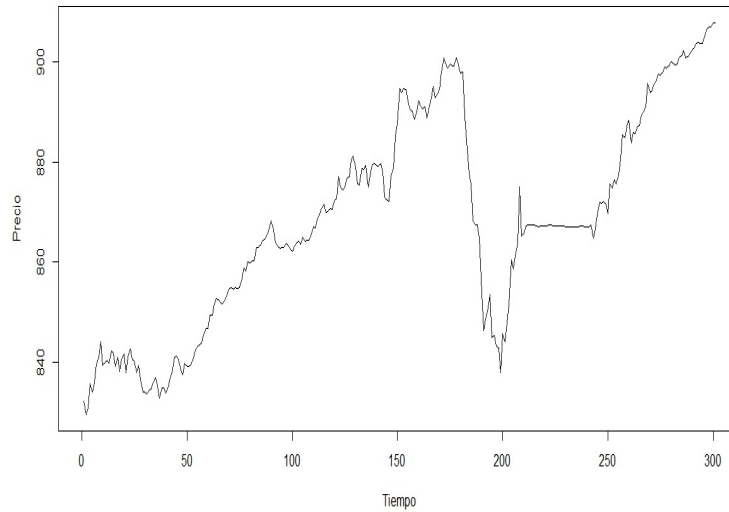


Figura 3.6: Trayectoria simulada del precio de la canasta alimentaria. Fuente: Elaboración propia.

se pronosticó cada parámetro para el mes de julio del 2013, obteniendo los valores que se muestran en el Cuadro 3.4.

Cuadro 3.4: Pronóstico de parámetros en el mes de julio.

Parámetro	Pronóstico
Crecimiento porcentual	-0.00101
Volatilidad	0.00366

3.5.3. Simulación y evaluación del riesgo de las variables

Para realizar el experimento se simularon 10,000 trayectorias posibles de la variable del precio de acuerdo con los parámetros de los modelos, para lo cual se ejecutó el programa que se puede ver en el anexo B.3.

En la Figura 3.6 se puede apreciar una trayectoria simulada durante el experimento y se observa que la volatilidad de la variable cambia cada treinta días.

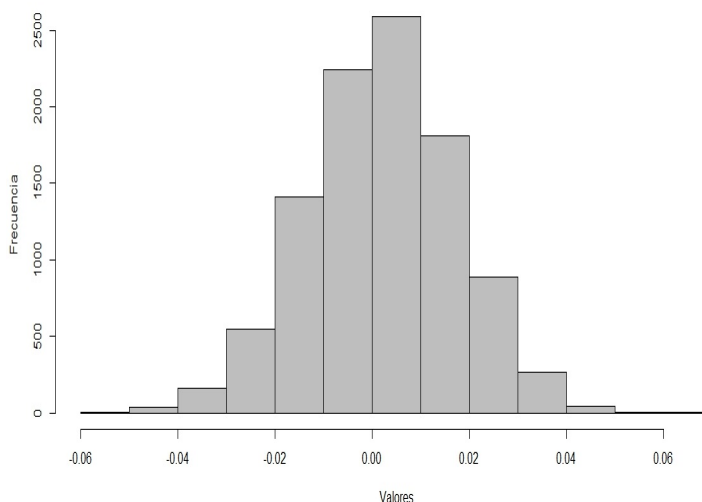


Figura 3.7: Crecimiento porcentual del precio simulado al final de diez meses. Fuente: Elaboración propia.

En la Figura 3.7 se puede apreciar la frecuencia con que aparecen valores simulados del crecimiento porcentual del precio de la canasta alimentaria rural al final del último mes simulado. Se observa que la variable se encuentra con media positiva (0.003897). Sin embargo, la simulación arrojó casos en los que el crecimiento resultó negativo.

En la Figura 3.8 se puede observar el histograma de los valores que tomó la volatilidad al final del último mes simulado. Se puede apreciar que aparecen simulaciones con posibles valores estresados, por lo que el riesgo que mide el modelo toma en cuenta estos posibles casos.

En la Figura 3.9 se puede apreciar la frecuencia con la que se simuló al final del periodo de evaluación el precio de la canasta alimentaria rural. El valor de la media es \$825.9 que es un valor que se puede tomar como pronóstico.

Finalmente, en la Figura 3.10, se puede apreciar el histograma del tiempo que tarda un microempresario en tener un día en el que su ingreso no alcanza para comprar la canasta alimentaria rural (*evento desfavorable*). Se puede apreciar un alto número de casos en los que ningún día del lapso de observación, el microempresario cayó en esta eventualidad (casos que se encuentran en el evento -2). Sin embargo, en varias ocasiones, el microempresario sí se encontró en el evento desfavorable. Se pueden usar los cuantiles

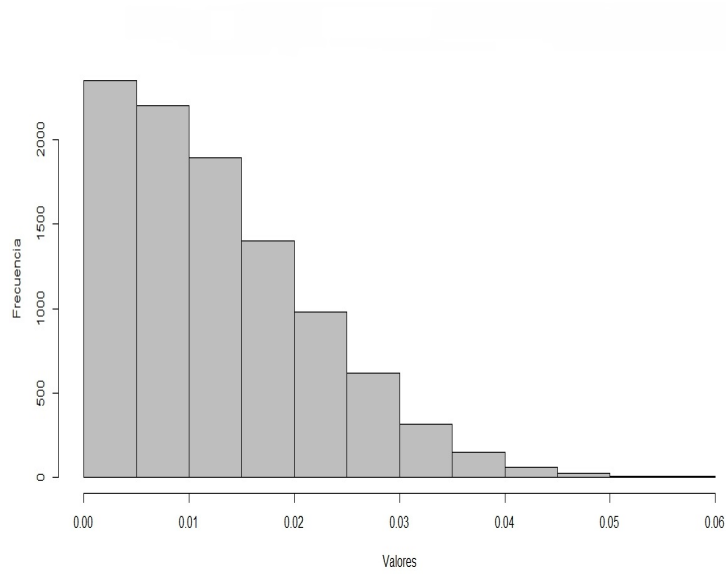


Figura 3.8: Volatilidad del precio simulado al final de diez meses. Fuente: Elaboración propia.

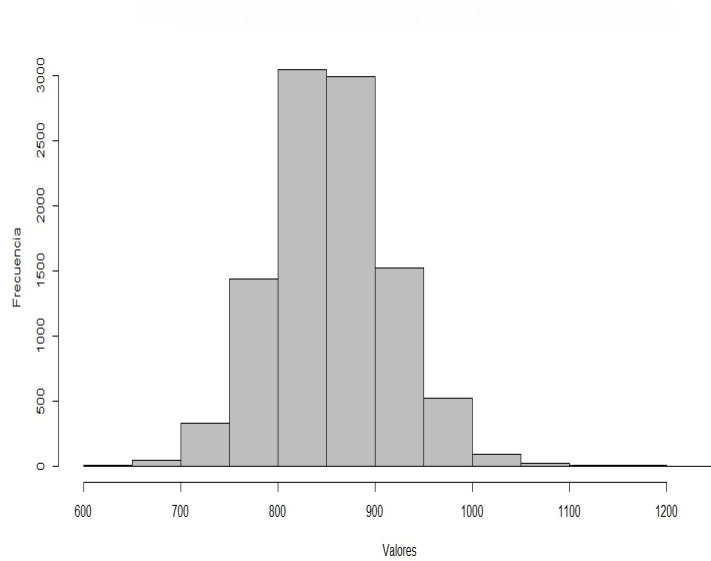


Figura 3.9: Precio simulado al final de diez meses. Fuente: Elaboración propia.

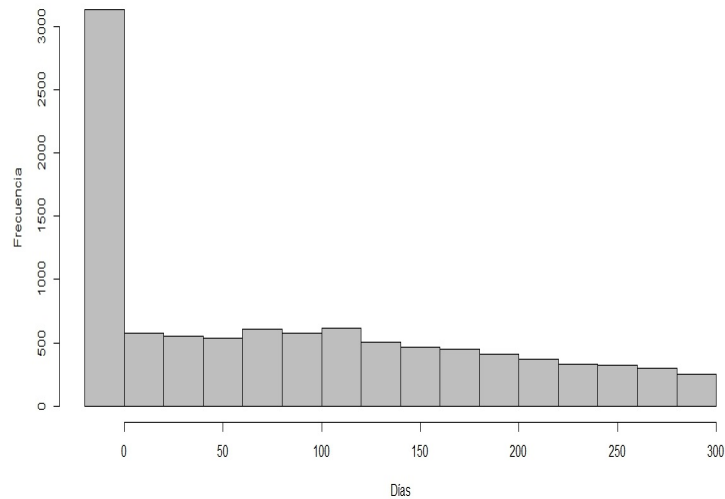


Figura 3.10: Tiempo restante para la falta de alimentación. Fuente: Elaboración propia.

de la distribución para medir con mayor detalle cuántos eventos son adversos. Se observa que el evento desfavorable se presenta en el 69 % de los casos. También se puede observar que el tiempo medio en lo que esto sucede es de 87.37 días.

Ahora que se tienen los resultados de la simulación será conveniente encontrar el VaR para una confianza del 95 % de las variables obtenidas para cuyo cálculo se pone a disposición el respectivo código en R en el anexo B.2. Gracias a la simulación, se pueden calcular diferentes riesgos fácilmente. Para ejemplificar este argumento, se calculó el VaR del precio de la canasta básica rural para el último día observado y también el VaR del mínimo día en que el evento desfavorable ocurre. Así se podrá obtener mayor detalle respecto a los diferentes riesgos que enfrenta el microempresario. El VaR del precio de la canasta alimentaria después de diez meses es de \$851, y el VaR del evento desfavorable es de 262 días. El VaR del evento desfavorable se puede interpretar como el tiempo de vida mínimo del ingreso del microempresario generando un indicador extra para la medición de la *elasticidad de la pobreza*.

Así, se puede concluir que el VaR no dice mucho sobre el comportamiento del riesgo por lo que lo mejor será obtener varias medidas de cuantiles de diferentes variables (variables estratégicas como son cuánto y cuándo) y

mediciones de valores esperados (como los obtenidos en este ejemplo). Con la simulación se puede obtener una *medida multidimensional del riesgo*, que en este caso puede observarse que es más apropiada.

Capítulo 4

Discusión sobre las aplicaciones

La discusión se presentará en cuatro partes: la primera se concentra en la teoría expuesta sobre los procesos estocásticos; la segunda parte hablará sobre los métodos de simulación; la tercera sobre la simulación de los procesos estocásticos; y la cuarta, sobre la pobreza alimentaria.

4.1. Procesos estocásticos

La continuidad del proceso de Wiener permite que éste se pueda diferenciar e integrar si se considera como un proceso gaussiano generalizado. Pasar al sistema matemático de funciones generalizadas podría tener sólo sentido a medida de que se requiera modelar un fenómeno real. En este caso, si no se contara con la contradicción de la correlación instantánea unitaria, no se podría definir un proceso de Wiener como generalizado.

En el campo de los procesos estocásticos se puede hallar un conjunto de teorías diferentes pero que hacen uso del ruido blanco para modelar el error instantáneo de una variable, tal es el caso de los modelos autorregresivos y el proceso de Wiener. Sin embargo, no existe evidencia de que la distribución de los errores sea normal para todas las aplicaciones, y al no cumplirse este supuesto, gran parte de la teoría de ecuaciones diferenciales estocásticas o modelos autorregresivos pierde utilidad. En tal caso, tendría que desarrollarse una teoría diferente para cada tipo de distribución. Pero se pueden encontrar diferentes ventajas en el uso de la normalidad e incluso justificación para su uso, ya que es conveniente cuando se cuenta con poca información sobre el fenómeno a modelar. En el caso específico del cálculo de las opciones, se tiene una distribución que cuenta con las mismas probabilidades de crecer o decrecer.

Una ventaja del uso del proceso de Wiener es que se pueden encontrar modelos de precios y ecuaciones diferenciales que cuenten con solución. Como es el caso del movimiento geométrico Browniano cuya distribución queda definida para toda t , por lo que la simulación de éste es redundante ya que fácilmente se puede calcular el VaR para cualquier intervalo de confianza y con esta idea se llega a que la simulación sólo tiene sentido si se desea aplicar a un proceso cuya distribución se desconoce o si la distribución de los parámetros a simular es diferente de la normal, ya que concebir un modelo para una distribución en específico puede ser más costoso. Si bien hasta este momento se ha limitado el uso de la simulación a escenarios en los que se desconoce la distribución del proceso, ésta ofrece ventajas también en cuanto a resultados se refiere. Por ejemplo, es sencilla la obtención de variables - como las vistas en el ejemplo del capítulo anterior- si se compara con tratar de encontrar la distribución del proceso en algún tiempo t , pues el costo en tiempo de investigación puede elevarse y en ocasiones se requiere incluso de supuestos inamovibles como es el caso de la teoría actuarial tradicional, en la que se suponen valores esperados provenientes de una tabla de mortalidad año a año en lugar de realizar una simulación de la siniestralidad bajo estudio.

En el caso de los modelos de series de tiempo para realizar pronósticos, se necesita de una serie estacionaria, que su variación no contenga una función de distribución con cola pesada de una muestra representativa de datos y que cuente con una distribución normal. No es seguro que esto suceda en la realidad y la simulación puede ayudar a encontrar intervalos de confianza para los pronósticos y a valorar el riesgo de que el modelo no sea del todo predictivo pues incluso se puede simular el error del modelo por medio de la distribución de los residuos y así obtener un intervalo de confianza que no se encuentra restringido a una fórmula que necesite del supuesto de normalidad.

Es interesante observar el desarrollo histórico del cálculo estocástico, que en el siglo XX derivó en la obtención de modelos de precios. En los inicios del cálculo estocástico se debieron obtener las propiedades básicas del proceso de Wiener y posteriormente desarrollar teoremas y proposiciones respecto a la integrabilidad de las ecuaciones. Pero el teorema que mayor importancia contiene en este campo es el del Lema de Ito ya que gracias a él se pueden obtener modelos que usan el movimiento geométrico Browniano, lo que permitió obtener diversos modelos de precios en las últimas décadas. La concepción de la integral de Ito resultó ser sólo una justificación de la existencia de estos modelos, por lo que puede ser hecha a un lado en cursos básicos. Sin embargo, las propiedades de la integral pueden ser usadas para demostrar futuros teoremas que la usen como base, por lo que en un curso superior que

contemple la dinámica de precios, debería ser un tema obligado. Esta teoría hace uso de varios supuestos que, como se ha mencionado, no siempre se cumplen en la realidad, sobre todo si la aplicación no se basa en el supuesto de no arbitraje usado en el cálculo de precios de opciones. El parámetro de la volatilidad juega un papel central en estos modelos ya que al suponer que la distribución de éste en t es normal o lognormal determina al proceso estocástico junto con el crecimiento porcentual. Afortunadamente la simulación de estos parámetros puede arrojar un riesgo factible respecto de los valores que un precio o tasa pueden tomar en un tiempo t , por lo que el cálculo de una opción o de un riesgo por medio de la solución de una ecuación diferencial basada en el Lema de Ito queda reducida, es inadecuada y no valora el riesgo proveniente de la incertidumbre creada con el cálculo de los parámetros.

La teoría básica expuesta sobre funciones de distribución, los diferentes tipos de convergencias y las definiciones básicas sobre variables aleatorias, son un conjunto de herramientas que son utilizadas en la mayoría de las construcciones de los teoremas referentes a los procesos estocásticos ya que el objetivo más importante de su estudio es el de encontrar la función de distribución para toda t , por lo que estas definiciones y sus propiedades deben ser presentadas en todo curso básico de procesos estocásticos.

4.2. Métodos de simulación

De acuerdo con lo observado en el Capítulo 2, la simulación por computadora depende de tres grandes campos de investigación: la generación de números pseudoaleatorios por medio de algoritmos matemáticos, la obtención de números pseudoaleatorios con funciones de distribución diferentes a una función de distribución uniforme y la creación de un algoritmo que pueda modelar el fenómeno que se requiera. Bajo este pensamiento, se discutirán algunos aspectos.

En el siglo XX la generación de números pseudoaleatorios derivó en diferentes modelos que cuentan con características diferentes. Los primeros modelos lineales congruenciales son fáciles de aplicar pero su periodo es corto. Estos generadores de números aleatorios son fáciles de replicar, por lo que son excelentes en caso de que se requiera reproducir un experimento. Tampoco requieren de un gran número de iteraciones computacionales por lo que no ocupan gran cantidad de memoria. Así, los Generadores Lineales Congruenciales (GLC) son adecuados para la investigación básica, pero al requerir una uniformidad mayor de los números generados y una cantidad infinita de ellos, los generadores naturales son de gran ayuda. No obstante,

la única forma de reproducir el experimento es guardando todos los números obtenidos en la memoria de la computadora, lo que puede generar grandes requerimientos de hardware. En los trabajos de las últimas décadas, el generador Mersenne Twister (MT) ha superado el periodo de cualquier intento de optimización de los GLC, y además ha mejorado la uniformidad de los datos, como se muestra en [24]. Los generadores MT siguen usando la teoría de módulos, por lo que el estudio de este campo podrá dar resultados en el futuro sobre los números pseudoaleatorios. El algoritmo se basa en fórmulas predictivas por lo que se sugiere utilizar pruebas de uniformidad para complementar los resultados de la simulación y poder demostrar que el generador utilizado es mejor con respecto a otro generador.

Ya que se cuenta con los números uniformes pseudoaleatorios, éstos deben convertirse a números con una distribución dada, para lo cual existen diferentes técnicas, las cuales contemplan algunos tipos de distribuciones o requieren de gran capacidad computacional ya que se basan en aceptación y rechazo. Afortunadamente, la mayoría del software de simulación, como es el caso de R, cuenta con las transformaciones predefinidas, de tal forma que la simulación de estas distribuciones es fácil y no se debe dedicar tiempo en su programación. Una buena forma de evaluar los resultados es graficar un histograma y observar el parecido con que cuenta la distribución de acuerdo con los supuestos hechos. También comprobar los primeros momentos permitirá determinar si la simulación se aproxima a los supuestos. Así, la elección de la función de distribución a simular es un paso fundamental en la modelación, para lo cual se emplean las pruebas de hipótesis de normalidad.

En ocasiones se requerirá conocer el error de un modelo y el posible valor de un parámetro y en algunos casos no se tendrán datos sobre estos parámetros. La observación y la elección del modelo ayudarán a concluir qué distribución se desea adoptar. La simulación permitirá comparar diferentes funciones de distribución y elegir la que se considere más conveniente. Por ejemplo, en el capítulo anterior, se supuso una distribución Poisson para simular los ingresos de un microempresario. Se puede realizar el experimento para algún otro microempresario cuyos ingresos se distribuyan de forma diferente y evaluar diferentes tipos de ingresos.

No obstante, no siempre se puede aplicar una simulación o no siempre es la herramienta ideal para determinar el comportamiento de una función de distribución, ya que si se cuenta con pocos datos sobre el fenómeno, será imposible saber qué función de distribución describe algún supuesto del modelo. Por ejemplo, en el capítulo anterior, se supuso que la función de distribución

del microempresario era una función de distribución Poisson. Esto no quiere decir que todos los microempresarios cuentan con esa función de distribución.

4.3. Simulación de procesos estocásticos

La ecuación estocástica de Ito permite obtener una ecuación diferencial de los precios igual a la del interés compuesto pero con la cualidad de tener un factor aleatorio que permite obtener una solución de la ecuación diferencial dentro de un intervalo de confianza, con lo que se obtiene una solución con comportamiento lognormal. Esto permite encontrar el valor presente de precios aleatorios con sólo considerar el crecimiento porcentual y la volatilidad del modelo. El modelo también sirve para tasas de interés por lo que en el estudio de los mercados de derivados es de gran importancia ya que permite saber el valor presente de la distribución en t del precio o de la tasa. También permite obtener el precio de una opción además de un análisis sobre la conveniencia de comprar o vender la opción a medida de que el tiempo se aproxima al tiempo en que se hace válida la opción.

Antes de los modelos de Black y Sholes para el cálculo de opciones, el mejor método para encontrar el precio de una opción era el de la simulación del movimiento geométrico Browniano y después la simulación se fue enfocando a las opciones que contienen correlación entre activos como es el caso de las opciones del tipo *rainbow* para las que no se cuenta con una ecuación que calcule el precio de la opción (véase [28]).

La utilidad de simular la volatilidad ha resultado en estudios como el de Hull (véase [33]), en el que incluso se propone simular la volatilidad con movimiento geométrico Browniano. Los modelos de Black y Sholes no pueden considerar esta ventaja (véase [28]). También se puede simular la volatilidad bajo diferentes supuestos de tendencias y modelos, como en el ejemplo del capítulo anterior, en el que se maneja un modelo con ciclicidades temporales conocidas previamente por el implementador.

Otras transformaciones del movimiento geométrico Browniano pueden ser utilizadas para obtener resultados más parecidos a la realidad, como es el caso del modelo de salto, en el que se pueden obtener valores estresados del precio. Pero se necesitaría observar una serie que contenga estos saltos frecuentemente para realizar una suposición acertada del tamaño y la frecuencia de estos saltos. Por ejemplo, en el caso de la inflación, se tendría que esperar a que existiera un periodo de hiperinflación el cual es cada vez menos frecuente pero sería imprescindible considerar estos saltos en la evaluación

de un préstamo a largo plazo como un préstamo hipotecario (en el que el pago crece conforme a la inflación). Así, aunque el modelo de salto pudiera no justificarse estadísticamente, puede utilizarse como práctica para obtener valores estresados. Una investigación que use la simulación por computadora no siempre resultará en conclusiones justificadas científicamente ya que un evaluador financiero puede simular casos que no se sabe si sucederán en el futuro.

Un caso contrario al modelo de salto es el modelo de Ornstein-Uhlenbeck (veáse [32]) el cual se deriva de una aplicación en el campo de la física y que converge a medida que el tiempo tiende a infinito. El supuesto es que la tendencia a largo plazo no cambia ya que nuevas tecnologías son creadas periódicamente dentro de una empresa, por lo que el precio de una acción se recupera después de un tiempo. A este fenómeno se le conoce como reversión a la media y se ha implementado en diversos modelos como se precisa en [28]. Este supuesto puede parecer razonable, pero se necesitarían datos a largo plazo que lo sustentaran, por lo que esto lleva a la conclusión de que la elección del modelo depende también de las necesidades de estudio en el corto o largo plazo.

En el ejemplo del Capítulo 3 se mostró la simulación de escenarios de los ingresos de un microempresario y cómo estos determinan si se encuentra en marginación alimenticia independientemente de que sus recursos mensuales acumulados se encuentren por arriba del precio de la canasta alimentaria rural. Se decidió usar la simulación de los parámetros de acuerdo con un modelo autoregresivo. Esto produjo parámetros más adaptados a lo que pudiera suceder en el futuro. Además, al usar la simulación se pudo establecer cuándo el microempresario no podría alimentarse adecuadamente y, en promedio, en qué tiempo sucedería esta eventualidad. Se mostró que, bajo el modelo supuesto de ingresos, existen seres humanos que, aunque cuentan con ingresos mensuales para no considerarse dentro del rango de personas en marginación alimentaria, pueden pasar días con hambre. La población en riesgo de tener hambruna debería estudiarse bajo este entendido.

Otro aspecto importante que se puede observar en el ejemplo es que la simulación brinda una base para crear una política pública de subvenciones para que las personas en la misma situación que el microempresario puedan sortear los días de hambre que padecen o, en el caso de intereses privados, se puede crear un seguro o instrumento futuro para que el cliente obtenga raciones de alimentos o préstamos suficientes para que no pase hambre. Este último argumento se expone en [5] en donde se observó que las personas con

ingresos irregulares y que cuentan con bajos ingresos, hacen uso de préstamos para ayudarse a solventar los días de hambre o de insolvencia financiera. Es por ello que un evaluador financiero puede usar un modelo similar para encontrar el préstamo propicio para que el microempresario no caiga en sobre endeudamiento y pueda sortear sus días de insolvencia, así como evaluar la capacidad de pago y el tiempo promedio en que el cliente entraría en cartera en riesgo hasta que se recupere de los días con bajos ingresos.

Se puede usar un modelo similar para la industria microfinanciera o para la creación de un fondo de subvenciones. Por ejemplo, en el caso de que una microfinanciera quisiera implementar modelos para determinar la solvencia de cada cliente y su frecuencia de pago adecuada para que éste no caiga en cartera en riesgo, la entidad microfinanciera tendría que determinar una distribución para cada cliente, o bien, para un grupo de clientes con micronegocios similares para poder implementar el modelo. Esto requeriría capacitar a su fuerza de ventas, así como obtener datos verídicos de los flujos del cliente, lo que representa una inversión significativa.

La simulación puede implementarse en diferentes campos del conocimiento en donde se requiera evaluar el riesgo de diferentes variables que afectan el resultado de una inversión. Para ello es preciso contar con datos suficientes para realizar estimaciones, estudios estadísticos complementarios para fundamentar los supuestos, conocimiento del evaluador sobre el fenómeno y capacidad de cómputo suficiente para realizar los cálculos.

4.4. La pobreza alimentaria y la simulación

La pobreza alimentaria es uno de los mayores problemas en el mundo. Este fenómeno existe en diferentes formas según la situación geográfica o cultural de las zonas en las que se presenta. Si bien la Food Agriculture Organization (FAO) ha definido diferentes tipos de pobreza como la pobreza visual, sonora y cultural, la pobreza alimentaria es una de las más importantes ya que supone un sufrimiento físico que puede ocasionar daños permanentes e incluso la muerte en un corto plazo. Aunado a esto, la pobreza alimentaria afecta en mayor medida a personas con cuerpos poco resistentes a la baja ingesta de alimentos, esto es, sus efectos físicos son más evidentes en personas de edad avanzada y niños. Es por ello que la Organización de las Naciones Unidas (ONU) propuso erradicar la pobreza extrema y el hambre como uno de los objetivos de desarrollo del milenio para lo cual se requiere que los países en el mundo tengan avances en el tema.

4.4.1. Causas de la pobreza alimentaria

Como se precisa en [9], la carencia alimentaria en México es el estado económico que le impide a un individuo adquirir los alimentos necesarios para poder tener una alimentación sana y, además presenta tres o más carencias sociales. A este conjunto de alimentos se le llama canasta alimentaria. Si la persona evaluada no cuenta con el ingreso suficiente para comprar la canasta alimentaria, se dice que la persona se encuentra por debajo de la línea de bienestar mínimo. Al observar esta definición se puede apreciar que existen dos variables que determinarán el estado de pobreza y estas variables son el precio de la canasta alimentaria y el ingreso de las personas. Como se señala en [14], estas importantes variables dependen de las condiciones económicas de la zona geográfica en la que se encuentra el individuo. Así, el entendimiento, monitoreo y predicción de las variables macroeconómicas ayudarían a crear políticas y prácticas para la disminución de la pobreza.

Un ejemplo de cómo las variables macroeconómicas influyen en el precio de la canasta básica es el del limón: su precio depende de las exportaciones, importaciones, la oferta, la demanda, la especulación del precio y el costo de los combustibles que serán usados para impulsar los transportes que lo llevarán al lugar de venta, la lejanía entre el lugar de cosecha y el lugar de venta, la inflación, entre muchos otros factores.

En cuanto a los ingresos como variable, algunos de los indicadores que lo determinan son las remesas recibidas por los familiares, el número de migrantes, el salario mínimo y el empleo. Al considerar estos factores en un modelo predictivo, se necesitaría una muestra considerable para determinar las correlaciones conjuntas entre cualquier combinación de ellas además de un esfuerzo considerable para lograr la reducción del error. Un modelo basado en redes neuronales podría considerar todas estas variables. Sin embargo, podría existir alguna restricción para obtener todos los datos necesarios, como se especifica en [14]. En este caso la segmentación de grupos poblacionales por medio de conglomerados resultaría útil para atender a la población con estrategias diferenciadas y así obtener políticas más adecuadas. Sin embargo, para aplicar este método y evaluar a toda la población de México, se necesitarían los datos provenientes de un censo y para cuando concluyera el levantamiento del mismo los datos dejarían de ser representativos al no ser los actuales.

En [2] se presentan modelos alternativos que pretenden funcionar como alertas tempranas de acuerdo con los cambios de las variables macroeconómicas. Estos modelos se basan en la simulación de las variables de modelos

económicos y brindan escenarios para diferentes países en caso de que se presente alguna crisis económica.

En este sentido cabe destacar la importancia de la simulación para poder crear un fondo nacional contra el aumento de la pobreza en nuestro país. En [14] se explica que la población mexicana es más sensible a crisis económicas mundiales ya que la población que se encuentra en peligro de tener ingresos por debajo de la línea de pobreza es en proporción mayor que en otros países. También señala que después de la crisis mundial del 2009 en México se redujeron las exportaciones de las empresas maquiladoras hacia los Estados Unidos de América, lo que afectó en mayor medida a la población urbana que a la población rural. Este resultado muestra que la segmentación de las personas que se encuentran en pobreza es vital para implementar mejores respuestas a eventos desfavorables.

4.4.2. La medición de la pobreza en México

En México el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) realiza la medición oficial de la pobreza por lo que busca medirla multidimensionalmente, esto es, mide diferentes indicadores de ésta. Entre ellos, la pobreza alimentaria. Uno de los problemas de esta medición es que depende de encuestas realizadas por el Instituto Nacional de Estadística y Geografía (INEGI) que no son necesariamente representativas por estado (véase [17]). Y por otra parte, como se comentó en la sección previa, falta una investigación más profunda en lo que respecta a la categorización de las personas que la padecen ya que de esta forma se podrían aplicar políticas públicas más eficientes. Por ejemplo, de quienes sufren pobreza alimentaria, no se ha calculado la proporción de personas que perciben ingresos del sector agropecuario o por remesas. La simulación de los ingresos y de los precios de las canastas para ciertos tipos de población con pobreza alimentaria, permitiría predecir qué tipo de necesidades tendrán y cómo será afectada su vida según se presenten cambios macroeconómicos desfavorables.

Para erradicar el hambre en México en el año de 2013 se emprendió la Cruzada Nacional Contra el Hambre (CNCH) que es liderada por la Secretaría de Desarrollo Social. La Cruzada es una iniciativa que pretende unir esfuerzos del sector privado, los diferentes órganos del gobierno y la población civil interesada en ayudar. De acuerdo con [9], la población objetivo es de 7.01 millones de mexicanos (personas que no pueden pagar la canasta alimentaria y además presentan dos o más carencias sociales del CONEVAL). Es una población de casi tres veces la de Panamá. A pesar de que esta iniciativa representa un avance significativo en el apoyo a la población vulnerable,

esta cifra es presumiblemente mayor si se considera a aquellos que sufren de hambre uno o más días dentro de un mes dependiendo de las variaciones diarias de los precios de los alimentos (sin tomar en cuenta los precios de la canasta básica no alimentaria).

Para poder mitigar la carencia alimentaria se necesita de un conjunto de medidas que a priori y a posteriori resuelvan el problema. Las medidas a priori deberían tener como objetivo lograr la sana alimentación de las poblaciones vulnerables. Un ejemplo de este tipo de medidas son los comedores comunitarios que puso en marcha la CNCH. Sin embargo, éstos son costosos a largo plazo en caso de que atiendan a toda la población objetivo.

Una inversión adicional y que cuenta con mayores ventajas es la capacitación en la mejora de la producción de alimentos en las zonas vulnerables. Sin embargo, como se comentó anteriormente, no todas las personas cuentan con las mismas necesidades y la capacitación no tendría resultados satisfactorios si no se prevén puestos de trabajo suficientes para contratar a las personas capacitadas. Por lo que se puede concluir que cualquier política será temporal si no se cuenta con crecimiento económico alto en las zonas más remotas del país, además de la generación de empleos con prestaciones suficientes.

El problema es mayor en el caso de poblaciones rurales alejadas si la región no cuenta con recursos naturales suficientes para su explotación, por lo que otra estrategia es capacitar e invertir en negocios comunitarios productivos.

En este trabajo se aplicaron modelos autoregresivos con base en la volatilidad de los precios de la canasta alimentaria rural. La relevancia del trabajo radica en que permite observar el comportamiento dinámico tanto de los ingresos como del costo de la canasta alimentaria así como evaluar diferentes tipos de ingresos y canastas, escenarios de crisis económica y los efectos de una hiperinflación. Por ejemplo, con la simulación de los ingresos promedio por remesas se puede evaluar cuál es el efecto de una disminución de éstas en un sector específico de la población.

En el caso del ejemplo del microempresario, se podría calcular el ingreso promedio en alguna zona geográfica específica y posteriormente evaluar cómo afectaría en su nivel de pobreza una baja en el turismo de la zona o un posible aumento en el cobro de impuestos.

Cabe destacar que la metodología presentada utiliza modelos básicos autoregresivos. Sin embargo, brinda un pronóstico de las tendencias del precio

de la canasta básica y elimina la necesidad de hacer diferentes supuestos sobre el sistema económico del país como necesitan los modelos presentados en [2].

Si se realiza una mayor inversión en el análisis de los diferentes tipos de personas que cuentan con pobreza alimentaria, esta inversión podría recuperarse rápidamente gracias a la optimización de recursos empleados en las políticas para combatir la pobreza, así como a la anticipación de los eventos negativos en la economía de zonas particulares. Para lograr esto, el análisis debe considerar la caracterización de grupos de individuos. Un censo podría ser el único medio para obtener los datos necesarios a nivel localidad, por lo que la planeación del cuestionario del siguiente censo debería tener un mayor énfasis en el tema de los ingresos. Una propuesta para la obtención y análisis de datos de pobreza es realizar una encuesta en cada gobierno municipal para determinar las razones que causan la pobreza alimentaria en cada municipio.

Capítulo 5

Conclusiones

A pesar de su extendido uso en la actualidad, la volatilidad es una medida de riesgo que puede ser malinterpretada ya que una baja volatilidad no implica que la distribución de los crecimientos de los incrementos no tenga una cola larga, esto es, se puede observar una volatilidad que tiende a bajar pero un VaR que tiende a aumentar, lo que recuerda la discusión sobre las medidas de riesgo y qué deben de cumplir. Si las distribuciones simuladas no son normales, deben usarse diferentes medidas de riesgo para verificar el riesgo. En este caso una buena medida del riesgo debería ser multidimensional, contemplando diferentes aspectos del riesgo ya que utilizar un sólo dato unidimensional como medida estándar de riesgo resultaría en un error al valorar a todos los activos de un portafolio de inversión. Para medir los diferentes riesgos presentes en un activo, la simulación resulta una herramienta valiosa, ya que en una sólo simulación, además de la pérdida posible se puede observar la distribución del tiempo en que ésta ocurre.

Si bien la simulación parece ayudar en diferentes áreas y acelerar el estudio de un fenómeno, ésta depende del cálculo de parámetros de los modelos, los cuales pueden ser erróneos o no necesariamente apegados a la realidad. Por esta razón se puede complementar la simulación con pronósticos, como se realizó en el Capítulo 3, en donde fue posible hallar un intervalo de confianza aceptable del modelo final y alimentar tomando en consideración incluso la incertidumbre que existe en el valor de los parámetros. Pero el modelo debe ser parsimonioso, debe ser reducido con el fin de que pocas variables determinen gran parte del comportamiento del fenómeno ya que en otro caso el modelo puede ser sobre estimado.

Para que una simulación resulte eficiente, se deben hacer pruebas sobre los supuestos del modelo empleado, lo cual demanda contar con datos suficientes

para realizarlas. Esto es, un estudio estadístico previo sobre las variables implicadas es imperativo antes de elegir un modelo. Sin embargo, en caso de que no se cuente con los datos adecuados, la simulación ayudará a ver qué pasaría si un supuesto arbitrario se cumpliera, por lo que el realizador de la simulación debe ser una persona con conocimientos no sólo de carácter matemático, sino también sobre el área del conocimiento en que la simulación se aplique. Esto con el fin de realizar un análisis lo más cercano a la realidad posible.

Con la simulación de los procesos estocásticos se pueden obtener diversos resultados sobre la función de distribución del proceso ya que, al simular los parámetros y al poner a prueba gran cantidad de supuestos sobre el modelo, es posible obtener una mejor apreciación del riesgo multidimensional. En cambio, con el uso de las soluciones de las ecuaciones diferenciales estocásticas sólo se pueden modificar un número restringido de supuestos.

La veracidad de la estimación de los parámetros de los modelos debe revisarse en cada simulación. Para ello se deben obtener intervalos de confianza para decidir si un parámetro debe o no ser tomado como variable aleatoria. Por esa razón, antes de iniciar una simulación se debe realizar un estudio estadístico sobre el fenómeno y sobre los supuestos se debería tratar de simular la distribución del parámetro en caso de que el intervalo de confianza de éste resulte muy amplio, pues de no hacerlo el modelo se alejaría de la realidad. Así, la simulación brinda una visión más realista del comportamiento de los modelos estadísticos como en el caso de las series de tiempo, ya que al realizar un pronóstico sin la ayuda de simulación, por ejemplo, con un modelo $AR(p)$, se pueden dejar de considerar valores en el intervalo de confianza de la predicción.

Con estas observaciones, se puede concluir que el uso de la simulación crecerá de acuerdo con la capacidad de cálculo con que cuentan las computadoras actuales, lo que supone un futuro alentador. También se tiene un avance en la simulación de modelos estadísticos pequeños en el sentido de que el número de cálculos utilizados para ello es comparable con el de los cálculos necesarios para simular el ejemplo expuesto en el Capítulo 3.

La autocovarianza en algunos modelos de precios, como el movimiento geométrico Browniano, resulta ser determinista. En caso de que se simulen los parámetros, la autocovarianza es convertida en una variable aleatoria dependiente del parámetro de volatilidad. Este fenómeno altera la función de distribución en cualquier instante y produce una mejor aproximación a la realidad ya que el modelo contiene la ciclicidad del parámetro y no sólo

una autocovarianza conocida. Para usarse, se tendrían que realizar pruebas estadísticas para comprobar este supuesto.

La relevancia social del trabajo radica en que muestra cómo la simulación permite observar el comportamiento diario tanto de los ingresos como del costo de la canasta alimentaria, permitiendo evaluar diferentes tipos de ingresos y diferentes composiciones de canastas alimentarias. El modelo propuesto permite considerar escenarios de crisis económica y sus efectos en el incremento de la pobreza alimentaria. Esto es necesario ya que los recursos usados para la implementación de políticas públicas es limitado y una mejor planeación de su uso puede optimizar los recursos destinados a las políticas sociales.

Una limitante del estudio es que los modelos utilizados para los pronósticos del costo de la canasta alimentaria no consideran la relación con otras variables macroeconómicas que podrían cambiar las tendencias de los precios de los alimentos en el corto o mediano plazos. Por esta razón un nicho para investigaciones futuras es estudiar las relaciones que existen entre el costo de las canastas alimentarias y la evolución de variables económicas como el crecimiento en la producción de los elementos de la canasta, o bien la simulación de la dinámica de los diferentes tipos de ingresos que existen en los hogares como son las remesas, los salarios e ingresos diarios de micronegocios.

Anexo A

Pruebas estadísticas

A.1. Prueba de normalidad *Jarque – Bera*

H_0 : La muestra se distribuye normal.

H_1 : La muestra no se distribuye normal.

Estadístico de prueba:

$$JB = \frac{N}{6}SK + \frac{N}{24}KU,$$

en donde

$$SK = (sk_1, sk_2, \dots, sk_p)' (sk_1, sk_2, \dots, sk_p),$$

y

$$KU = (ku_1 - 3, \dots, ku_p - 3)' (ku_1 - 3, \dots, ku_p - 3),$$

en donde ku_i es una estimación de la curtosis y sk_i es la oblicuidad del elemento $i = 1, 2, \dots, N$ de la muestra (véase [13]).

Región de rechazo: El estadístico de prueba se distribuye χ^2 cuadrada con $2p$ grados de libertad. La hipótesis nula se rechaza con un grado α de significancia cuando $JB \leq \chi^2_{2p, \alpha}$.

A.2. Prueba de raíz unitaria de *Dickey–Fuller* para un modelo $X_t - X_{t-1} = a_1 X_{t-1} + e_t$

H_0 : La serie es no estacionaria ($a_1 \neq 0$).

H_1 : La serie es estacionaria ($a_1 = 0$).

Estadístico de prueba: El estadístico de prueba para la a_1 es obtenido mediante simulaciones de Montecarlo (véase [13]).

Región de rechazo: Rechazar H_0 si ($p > 0.05$).

Anexo B

Programas en R

B.1. Movimiento geométrico Browniano

```
#Input: precio[1] = Precio inicial
#Output: resultado = Gráfica requerida

precio<-c(1)
precio[1]=100

for(i in 2:1001)
  {
      precio[i]=precio[i-1]*(1+(1/1000)*0.03
      +0.01*sqrt(1/1000)*rnorm(1))
  }

#Se grafican los precios obtenidos

resultado<-plot(ts(precio))
```

B.2. Simulación de los flujos de un microempresario

```
#Input: a = Parámetros del modelo AR(12)
# del crecimiento porcentual del precio
# b = Parámetros del modelo AR(2) de la volatilidad
# y = Últimos 12 valores del crecimiento porcentual del precio
# u = Últimos 2 valores de la volatilidad
```

```
#Output: c = Crecimiento
# t = Volatilidad
# w = Precio
# v = Tiempo de quiebra
# varprecio = VaR .95 del precio
# vardesf = VaR .95 del evento desfavorable

# El vector "a" define los parámetros del modelo AR(12)
# del crecimiento porcentual del precio,
# calculados con el programa del Anexo B.4.

a<-c(0.347099274,
      -0.310330345,
      0.072665898,
      -0.127229413,
      0.091422279,
      -0.001864348,
      0.109426224,
      0.066795555,
      0.173358307,
      -0.028602726,
      0.139602503,
      0.17008097
)

# El vector b define los parámetros del modelo AR(2)
# de la volatilidad del precio

b<-c(0.3495271 ,
      0.3537163
)

y<-c( 0.02004113 ,
      0.00210373 ,
      0.02279259 ,
      0.00223580 ,
      -0.00426658 ,
      0.00871662 ,
```

B.2. SIMULACIÓN DE LOS FLUJOS DE UN MICROEMPRESARIO113

```
        0.00074034 ,
        -0.00186766 ,
        0.01840782 ,
        0.00200437 ,
        -0.00057153 ,
        -0.00843490
    )
u<-c( 0.0025759 ,
      0.0078634
    )
)
```

#Se declaran las variables e índices para los bucles

```
v<-c(1)
beta<-c(50)
c<-c(1)
z<-c(1)
l<-c(1)
t<-c(1)
w<-c(1)
e<-c(1)
o<-c(1)
up<-c(1)
h<-c(1)
zer<-c(1)
```

```
dif<-c(1)
```

```
z[1]=832.29
e[1]=z[1]/30
```

```
for(j in 1:10000) #Se realizan 10,000 simulaciones
{
  v[j]=-2 #Se asigna el valor -2
  l=1
  for(i in 1:10)
  {
```



```

y[12+i]=0
for(zer in 1:12)
{
  #Se realiza la estimación del crecimiento
  #del precio para el mes

  y[12+i]= y[12+i]+y[12+i-zer]*a[zer]

}

y[12+i]=y[12+i]+rnorm(n=1,mean=0,sd=0.014478814)

#Se realiza la estimación de la volatilidad
u[2+i]=abs(u[2+i-1]*b[1]+u[2+i-2]*b[2]+
rnorm(n=1,mean=0,sd=0.013381892))

for(r in 1:30)
{

  #Se realiza la simulación diaria del precio
  z[l+1]=z[l]*(1+y[12+i]*(1/30)+
  abs(u[2+i])*sqrt(1/30)*rnorm(n=1,mean=0,sd=1))
  e[l+1]=z[l+1]/30

  #Simulación del ingreso diario

  beta[l+1]=rpois(1,lambda=45)

  #Ganancia diaria

  dif[l+1]=beta[l+1]-e[l+1]

  #Se determina si se tiene una ganancia negativa
  o<-ifelse(dif[l+1]<0,1,2)
  up<-ifelse(v[j]>-2,1,10)

  #En caso de que la ganancia sea negativa,
  # se asigna el valor 11
  h[j]=o+up

  #Se detiene la cuenta

```

B.3. CÁLCULO DE LOS PARÁMETROS DE UN MODELO AR(12) 115

```
        #para una ganancia negativa
        ifelse(h[j]==11,v[j]<-c(1),3)

        l=l+1
      }
    }
  #Se guardan los output en un array
  c[j]=y[10+12]
  t[j]=u[10+2]
  w[j]=z[301]

}

#Se grafican los resultados

par(mfrow=c(3,3))
plot(ts(y),main="Crecimiento")
plot(ts(u),main="Volatilidad")
plot(ts(z),main="Precio")
hist(c,main="Crecimiento",col="green", freq=50)
hist(t,main="Volatilidad",col="blue",freq=50)
hist(w,main="Precio",col="red",freq=50)
hist(v,main="Espera",col="red",freq=50)

#Calculo del VaR
varprecio<-quantile(z,0.95)
vardesf<-quantile(v,0.95)
```

B.3. Cálculo de los parámetros de un modelo AR(12)

```
# Input: datos = Serie histórica
# Output: residuos = Residuos del modelo
# modelo = Predicción
# z = Parámetros del modelo
# intervalo = Intervalo de confianza
```

```
# erroresb = residuos del modelo
```

```
modulo<-(0)
```

```
zet<-(1)
```

```
# Primero se crea una función para calcular
```

```
# el error de un modelo AR(12)
```

```
objetivo<-function(y)
```

```
{
```

```
errores<-c(0)
```

```
for(i in 13:length(datos))
```

```
{
```

```
  f<-function(x)
```

```
  {
```

```
    modulo<-c(0)
```

```
    for(zet in 1:12)
```

```
      {
```

```
        modulo=modulo+
```

```
        datos[i-zet]*x[zet]
```

```
      }
```

```
    abs(datos[i]-modulo)
```

```
  }
```

```
  d<-c(0)
```

```
  for(j in 1:12)
```

```
  {
```

```
    d[j]=y[j]
```

```
  }
```

```
  errores[i-13+1]<-f(d)
```

```
}
```

```
sum(errores)
```

```
}
```

B.3. CÁLCULO DE LOS PARÁMETROS DE UN MODELO AR(12) 117

```
#Se calcula la matriz de Yule y Walker

A<-matrix(nrow=12,ncol=12)

for(j in 1:12)
  {
    for(i in 1:12)
      {
        ifelse( i<=j,A[i,j]<-acf(datos)$acf[j-i+1],
          A[i,j]<-acf(datos)$acf[i-j+1])
      }
  }

X<-matrix(nrow=12,ncol=1)

for(j in 1:12)
  {
    X[j]<-acf(datos)$acf[j+1]
  }

#Se encuentran los valores iniciales

y<-solve(A,X)

objetivo(y)

#Se optimiza el modelo

z<-optim(y,objetivo)$par

objetivo(z)

#Se muestran los parámetros optimizados
```

```

z

# Los residuos son calculados

erroresb<-c(0)

for(i in 13:length(datos))
  {
    f<-function(x)
      {
        modulo=c(0)
        for(zet in 1:12)
          {
            modulo=modulo+
            datos[i-zet]*x[zet]
          }
        datos[i]-modulo
      }
    d<-c(0)
    for(j in 1:12)
      {
        d[j]=z[j]
      }
    erroresb[i-13+1]<-f(d)
  }

```

B.4. Código de la prueba *Jarque – Bera*

```

# Input: crecimiento = Los crecimientos porcentuales del precio
# volatilidad = la volatilidad del precio
# Output: resultadoCre = resultados de la prueba
# para el crecimiento porcentual
# resultadoCre = resultados de la prueba para la volatilidad

```

```
# Se llama a la librería tseries

library(tseries)

resultadoCre <-jarque.bera.test(crecimiento)
resultadoVol<-jarque.bera.test(volatilidad)
```

B.5. Código de la prueba *Dickey – Fuller*

```
# Input: crecimiento = Crecimientos porcentuales del precio
# volatilidad = Volatilidad del precio
# Output: resultadoCre = Resultados de la prueba
# para el crecimiento porcentual
# resultadoCre = Resultados de la prueba para la volatilidad

# Se llama a la librería tseries

library(tseries)

resultadoCre <-adf.test(crecimiento)
resultadoVol <-adf.test(volatilidad)
```


Referencias

- [1] P. Artzner, F. Delbaen, J.-M. Eber y D. Heath, “Coherent Measures of Risk”, *Mathematical finance*, vol. 9, pp.203-228, 1999.
- [2] F. Bourguignon, M. Bussolo y L. A. Pereira da Silva, *The Impact of Macroeconomic Policies on Poverty and Income Distribution*, Estados Unidos, The International Bank for Reconstruction and Development / The World Bank, 2008.
- [3] G. Casella y R.L. Berger, *Statistical Inference*, Estados Unidos, Duxbury advanced series, 2002.
- [4] D.C. Champeney, *Fourier Transforms and their Physical Applications*. Reino Unido, Academic Press, 1973.
- [5] D. Collins, S. Rutherford, O. Ruthven, J. Morduch y A. Tato, *Las Finanzas de los Pobres*, Reino Unido, Debate, 2011.
- [6] Consejo Nacional de Evaluación de la Política de Desarrollo Social, (Noviembre, 2013). *Líneas de Bienestar y Canasta Alimentaria*. México. Disponible en: <http://www.coneval.gob.mx/Medicion/Paginas/Lineas-de-bienestar-y-canasta-basica.aspx>.
- [7] Consejo Nacional de Evaluación de la Política de Desarrollo Social, (Junio, 2014). *Metodología para la Medición de la Pobreza en los Municipios de México*, México. Disponible en: <http://www.coneval.gob.mx/Medicion/Paginas/Medici%C3%B3n/Nota-metodologica.aspx>.
- [8] Consejo Nacional de Evaluación de la Política de Desarrollo Social, (Junio, 2014). *Metodología para la Medición Multidimensional de la Pobreza en México*, México. Disponible en: <http://www.coneval.gob.mx/Medicion/Paginas/Medici%C3%B3n/Metodologia.aspx>.
- [9] Cruzada Nacional contra el Hambre, (Junio, 2014). *Qué es la cruzada*. México. Disponible en: <http://sinhambre.gob.mx/cruzada/que-es-la-cruzada>.

- [10] Equipo de Tareas de Alto Nivel sobre la Seguridad Alimentaria Mundial, (Junio, 2014). *Seguridad Alimentaria y Nutricional*, Estados Unidos. Disponible en: <http://www.un.org/es/issues/food/taskforce/>.
- [11] *Perspectivas Alimentarias*, Food Agriculture Organization, Estados Unidos, Mayo 2012.
- [12] A. Girón y E. Correa, *Asimetrías e Incertidumbre*, México, Instituto de Investigaciones Económicas, Universidad Nacional Autónoma de México, 2009.
- [13] V.M. Guerrero, *Análisis Estadístico y Pronóstico de Series de Tiempo Económicas*, México Jit Press, 2009.
- [14] B. Habib, A. Narayan, S. Olivieri y C. Sanchez, “The Impact of the Financial Crisis on Poverty and Income Distribution: Insights from Simulations in Selected Countries”, *Economic Premise*, vol. 7, 2010.
- [15] A.L. Haro, *Medición y Control de Riesgos Financieros*, México, Limusa, 2012.
- [16] P.G. Hoel, S.C. Port y C.J. Stone, *Introduction to Stochastic Processes*, Estados Unidos, Houghton Mifflin Boston, 1972.
- [17] Instituto Nacional de Estadística y Geografía, (Junio, 2014). *Encuesta Nacional de Ingresos y Gastos de los Hogares*, México. Disponible en: <http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/hogares/regulares/enigh/enigh2010/tradicional/default.aspx>.
- [18] K. Ito, “On a Formula Concerning Stochastic Differentials.” *Nagoya Mathematical Journal*, vol. 3, pp. 55-65, 1951.
- [19] P. Jorion, *Valor en Riesgo*, Estados Unidos, Limusa, 2003.
- [20] L.D. Kudriávtshev, *Curso de Análisis Matemático 2*, Rusia, Mir, 1984.
- [21] A.M. Law y W.D. Kelton, *Simulation Modeling and Analysis*, Estados Unidos, McGraw-Hill, 2000.
- [22] A. Ludwig, *Stochastic Differential Equations*, Canada, Wiley-Interscience, 1974.
- [23] J.E. Marsden y A. Tromba, *Cálculo Vectorial*, Estados Unidos, Pearson, 2004.

- [24] M. Matsumoto y T. Nishimura, “Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator”, *Transactions on Modeling and Computer Simulation*, vol. 8, pp. 3-30, 1998.
- [25] A.M. Mood, F.A. Graybill y D.C. Boes, *Introduction to the Theory of Statistics*, Estados Unidos, McGraw-hill, 1974.
- [26] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, Alemania, Springer, 2010.
- [27] R.T. Rockafellar y S. Uryasev, “Optimization of Conditional Value-at-Risk,” *Journal of risk*, vol. 2, pp. 21-42, 2000.
- [28] R.U. Seydel, *Tools for Computational Finance*, Alemania, Springer, 2012.
- [29] G. Szegö, “Measures of Risk”, *Journal of Banking & Finance*, vol. 26, pp. 1253-1272, 2002.
- [30] P. Teetor, *R Cookbook*, Estados Unidos, O’Reilly Media, 2011.
- [31] *R: Regulatory Compliance and Validation Issues A Guidance Document for the Use of R in Regulated Clinical Trial Environments*, The R Foundation for Statistical Computing c/o Department of Statistics and Mathematics Wirtschaftsuniversitat Wien, Austria, 2008.
- [32] O. Vasicek, “An Equilibrium Characterization of the Term Structure”, *Journal of Financial Economics*, vol. 5, pp. 177-188, 1977.
- [33] Z. Zhao, “Parametric and Nonparametric Models and Methods in Financial Econometrics,” *Statistics Surveys*, vol. 2, pp. 1-42, 2008.