BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

LICENCIATURA EN MATEMÁTICAS APLICADAS

PROCESOS AUTO-RECURSIVOS DE ORDEN UNO, SU RELACIÓN CON LAS MARTINGALAS Y SU APLICACIÓN EN LA PREDICCIÓN DE CICLONES EN MÉXICO

TESIS PARA OBTENER EL TÍTULO DE:

LICENCIADA EN MATEMÁTICAS APLICADAS

PRESENTA:

LOURDES PÉREZ AMARO

DIRECTORES DE TESIS:

DR. VÍCTOR HUGO VÁZQUEZ GUEVARA
DR. HUGO ADÁN CRUZ SUÁREZ

Dedicado a mi papás
Marina Amaro J. y Samuel Pérez J.
y a el gran ángel que vive en mi corazón
Carlos Amaro J. $^+$

Agradecimientos

A Dios, por la oportunidad de vivir y por la familia maravillosa que me ha dado.

A mis papás: Marina y Samuel, por todo su apoyo, confianza, comprensión y cariño que me han brindado desde pequeña.

A mis hermanos: Maricela y Marco Antonio, por todos los momentos que desde pequeños hemos compartido.

A mi pequeño sobrino: Johan, por todos los momentos de alegría que comparte conmigo.

A mis abuelitos: Agustina, Blanca, Amado y Concepción^{+}, por todo el cariño y confianza que han puesto en mi.

A toda la familia Amaro, por todos los momentos alegres y tristes que hemos vivido, porque a pesar de ciertas situaciones seguimos y seguiremos siendo una familia.

A toda la familia Pérez, por los buenos momentos que hemos compartido y por sus ánimos para concluir la Licenciatura.

A mis amigos y compañeros de la facultad, en especial a Noemi, Chente, Juan Carlos, Ana Laura, Bety, Alex y Camilo, por el apoyo a lo largo de la Licenciatura y por los buenos momentos que compartimos.

A mis amigas del futbol: Chely, Lili y Liz, por compartir conmigo la pasión al futbol, además de ser unas buenas amigas.

A mi novio: Rafa, por compartir conmigo y con mi familia bonitos momentos y por estar conmigo cuando lo he necesitado.

De manera especial, agradezco principalmente a mis directores de tesis:

Al Dr. Víctor Hugo Vázquez G., por los conocimientos brindados y por todo el tiempo dedicado para la realización de esta tesis, además de agradecerle por la amistad y confianza que me ha dado.

Al Dr. Hugo Adán Cruz S., por cada una de sus enseñanzas a lo largo de la Licenciatura y por dedicar parte de su tiempo para la mejora de esta tesis.

Y finalmente agradezco a mis sinodales:

Al Dr. Francisco Solano Tajonar Sanabria, al Dr. Bulmaro Juárez Hernández y a la Dr. Lidia Aurora Hernández Rebollar, por sus correcciones y aportaciones en esta tesis.

A cada uno de ustedes
¡Muchas Gracias!

Introducción

Una serie de tiempo es una sucesión de observaciones generadas por un proceso estocástico, cuyo conjunto de índices se toma en relación con el tiempo. Por lo tanto, la inferencia que se realice será acerca de las características del proceso estocástico generador de la serie observada. Consideraremos series de tiempo discretas, con la característica adicional de que las observaciones se hagan en intervalos de tiempo iguales.

Es importante notar que una serie de tiempo observada no es más que una realización de un proceso estocástico, lo cual significa que bien pudo haberse observado otra realización del mismo proceso, pero cuyo comportamiento fuese distinto del que se observó en la realidad.

Son diversas las aplicaciones que se pueden citar, en distintas áreas del conocimiento, tales como, en Meteorología, Economía, Física, Demografía, Mercadotecnía, etc. (ver [8],[9],[10],[11] y [12]). En esta tesis presentaremos el análisis de una serie de tiempo en Meteorología.

Entre los procesos de series de tiempo están los procesos auto-recursivos. Estos procesos se caracterizan por tener una forma funcional donde el valor de la variable al tiempo t depende de valores pasados de la misma variable a tiempos $t-1, t-2, .., t-p$, es decir, si $\{M_t, t \in \mathbb{N}\}$ es un proceso estacionario, el valor que toma la variable M_t depende de $M_{t-1}, M_{t-2}, ..., M_{t-p}$.

En esta tesis estudiaremos solamente el proceso auto-recursivo de orden uno, donde la variable M_t viene determinada únicamente por su valor pasado inmediato, M_{t-1}, además haremos una estimación de los parámetros del proceso por el método de mínimos cuadrados ligados a los resultados de martingalas con el objetivo de pronosticar a la variable M_{t+1} usando el proceso auto-recursivo de orden uno y la estimación de los parámetros.

En cuanto a los contenidos de esta tesis, el primer capítulo se inicia con algunos conceptos básicos de la teoría de probabilidad tales como espacio de probabilidad y variables aleatorias, también se enuncian los conceptos de esperanza condicional respecto a un evento, a una variable aleatoria discreta, a una variable aleatoria arbitraria y respecto a una σ-álgebra, además de que se enuncian las propiedades generales de la esperanza condicional que nos sirven para demostrar fácilmente ciertos resultados; y terminamos este capítulo con el concepto de proceso estocástico.

En el segundo capítulo se introduce el concepto de Martingala y se dan algunos resultados básicos como lo son el Teorema de Paro, las Desigualdades de Doob, el Teorema de convergencia de Doob, el Teorema de Robbins-Siegmund, la Ley de los grandes números y el Teorema de Límite Central; y concluimos dicho capítulo con resultados de los procesos auto-recursivos de orden uno.

En el tercer capítulo hacemos una aplicación de los procesos auto-recursivos de orden uno, trabajando con el software estadístico R, con datos extraídos del Servicio Meteorológico Nacional y datos artificiales. Los datos extraídos del Servicio Meteorológico son acerca del número de ciclones que han impactado a México cada año desde 1970-2010, la aplicación de los procesos de orden uno con estos datos es con el objetivo de pronosticar y hacer una comparación del número de ciclones que hubo en el año 2011 y pronosticar el número de ciclones que habrá en el año 2012.

Índice general

Capítulo 1

Espacios de probabilidad

La teoría de la probabilidad es la parte de las matemáticas que se encarga del estudio de los fenómenos o experimentos aleatorios. Se entiende por experimento aleatorio todo aquel experimento tal que cuando se le repite bajo las mismas condiciones iniciales, el resultado que se obtiene no siempre es el mismo. Bajo estas circunstancias, la teoría de la probabilidad tiene el objetivo de modelar matemáticamente cualquier experimento aleatorio de interés.

1.1. Espacio de probabilidad

El modelo matemático creado durante el primer tercio del siglo XX para estudiar los experimentos aleatorios es llamado espacio de probabilidad.

Definición 1. *Un* ***espacio de probabilidad*** *es una terna ordenada* $(\Omega, \mathcal{F}, P)$*, en donde* Ω *es un conjunto arbitrario,* $\mathcal{F}$ *es una* σ*-álgebra de subconjuntos de* Ω*, y* P *es una medida de probabilidad definida sobre* $\mathcal{F}$*.*

A continuación explicamos cada uno de estos elementos:

- **Espacio muestral**

El conjunto Ω es llamado **espacio muestral** y tiene como objetivo agrupar a todos los posibles resultados de cualquier experimento aleatorio. No es imprescindible darle esta interpretación al conjunto Ω, y matemáticamente se le considera entonces como un conjunto arbitrario.

- **σ-álgebra**

Definición 2. *Una colección no vacía $\mathcal{F}$ de subconjuntos de Ω es llamada una **σ-álgebra** si cumple las siguientes condiciones:*

1. $\Omega \in \mathcal{F}$,
2. *Si* $A \in \mathcal{F}$, *entonces* $A^c \in \mathcal{F}$ *y*
3. *Si* $\{A_n, n \in \mathbb{N}\} \subset \mathcal{F}$, *entonces* $\bigcup_{n=1}^{\infty} A_n \subset \mathcal{F}$.

A la pareja $(\Omega, \mathcal{F})$ se le llama **espacio medible** y a los elementos de $\mathcal{F}$ se les llama **eventos** o **conjuntos medibles**.

Nota 1. $\mathcal{F} = \{\emptyset, \Omega\}$ *es la σ-álgebra más pequeña que podemos asociar a Ω y $\mathcal{F} = 2^{\Omega}$ es la σ-álgebra más grande.*

Definición 3. *Sea $\mathscr{C}$ una colección no vacía de subconjuntos de Ω. La **σ-álgebra generada por** $\mathscr{C}$, denotada por $\sigma(\mathscr{C})$, es la colección*

$$\sigma(\mathscr{C}) = \bigcap \{\mathcal{F} : \mathcal{F} \text{ es } \sigma\text{-álgebra y } \mathscr{C} \subseteq \mathcal{F}\}. \tag{1.1}$$

Es decir, la colección $\sigma(\mathscr{C})$ es la intersección de todas aquellas σ-álgebras que contienen a $\mathscr{C}$. A $\sigma(\mathscr{C})$ también se le llama **mínima σ-álgebra generada por** $\mathscr{C}$, y el adjetivo mínima es claro a partir del hecho de que es la σ-álgebra más pequeña que contiene a la colección $\mathscr{C}$. Es decir, si $\mathcal{F}$ es una σ-álgebra que contiene a $\mathscr{C}$, entonces forzosamente $\sigma(\mathscr{C}) \subseteq \mathcal{F}$. Observemos que $\mathscr{C} \subseteq \sigma(\mathscr{C})$ pues a la colección $\mathscr{C}$ se le han añadido posiblemente algunos otros subconjuntos para convertirla en la σ-álgebra $\sigma(\mathscr{C})$.

Definición 4. *Sea $\mathscr{C} = \{(-\infty, x] : x \in \mathbb{R}\}$. A la mínima σ-álgebra generada por esta colección se le llama **σ-álgebra de Borel de** $\mathbb{R}$, y se le denota por $\mathcal{B}(\mathbb{R})$*

$$\mathcal{B}(\mathbb{R}) = \sigma(\mathscr{C}) = \sigma\{(-\infty, x] : x \in \mathbb{R}\}. \tag{1.2}$$

A los elementos de $\mathcal{B}(\mathbb{R})$ se les llama **conjuntos borelianos**. De esta forma se puede asociar la σ-álgebra $\mathcal{B}(\mathbb{R})$ al conjunto de números reales, y obtener así el espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Nota 2. $\mathcal{B}(\mathbb{R})$ *puede ser generada por cualquier colección de intervalos.*

- **Medida de probabilidad**

Definición 5. *Sea* $(\Omega, \mathcal{F})$ *un espacio medible. Una* ***medida de probabilidad*** *sobre el espacio medible, es una función* $P : \Omega \to [0,1]$ *que satisface:*

1. $P(\Omega) = 1$,
2. $P(A) \geq 0$ *para cualquier* $A \in \mathcal{F}$ *y*
3. *Si* $A_1, A_2, ... \in \mathcal{F}$ *son ajenos dos a dos, esto es,* $A_n \bigcap A_m = \emptyset$ *para* $n \neq m$*, entonces* $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$.

Estos axiomas fueron establecidos por A. N. Kolmogorov en 1933. En particular, la tercera propiedad se conoce con el nombre de σ-aditividad.

Considerando el espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ y $f : \mathbb{R} \to [0, \infty)$ una función no negativa y continua, tal que su integral sobre el intervalo $(-\infty, \infty)$ es uno. La función P definida por la siguiente integral es una medida de probabilidad,

$$P(A) = \int_A f(x)\, dx,$$

para cualquier conjunto de Borel A.

1.2. Variables aleatorias

Una variable aleatoria es una función del espacio muestral en el conjunto de números reales que además satisface cierta condición de medibilidad. Representa una traducción de cada uno de los resultados del espacio muestral en números reales. Mediante una variable aleatoria uno puede considerar que el posible experimento aleatorio en cuestión no produce como resultados elementos de Ω sino números reales. El concepto de variable aleatoria es fundamental en la teoría de la probabilidad, y una vez que enunciemos su definición, el término aparecerá con mucha frecuencia.

Definición 6. *Una* ***variable aleatoria*** *real es una función* $X : \Omega \to \mathbb{R}$ *tal que para cualquier conjunto Boreliano* $B \in \mathcal{B}(\mathbb{R})$*, se cumple que el conjunto* $(X^{-1}(B)) \in \mathcal{F}$*, es decir,* X *es una función* $\mathcal{F}$*-**medible**.*

Si B es un conjunto Boreliano, se usan los símbolos $(X^{-1}(B))$ y $\{X \in B\}$ para denotar el conjunto $\{\omega \in \Omega : X(\omega) \in B\}$.

Emplearemos las siglas v.a. para referirnos a una variable aleatoria.

Para comprobar que una función $X : \Omega \to \mathbb{R}$ es realmente una v.a., la definición requiere verificar la condición $(X^{-1}(B)) \in \mathcal{F}$ para cualquier conjunto Boreliano B. La siguiente proposición establece que no es necesario demostrar la condición de medibilidad para cualquier conjunto Boreliano B, sino que es suficiente tomar intervalos de la forma $(-\infty, x]$, para cada x en $\mathbb{R}$. Pero antes de enunciar la proposición, veamos el siguiente Lema y su demostración.

Lema 1. *$\mathscr{B} = \{B \in \mathcal{B}(\mathbb{R}) : (X^{-1}(B)) \in \mathcal{F}\}$ es una σ-álgebra de subconjuntos de $\mathbb{R}$.*

Demostración.

1. Tenemos que $\mathbb{R} \in \mathscr{B}$, pues $\mathbb{R} \in \mathcal{B}(\mathbb{R})$ y $(X^{-1}(\mathbb{R})) = \Omega \in \mathcal{F}$.

2. Sea $B \in \mathscr{B}$.
   Entonces $B \in \mathcal{B}(\mathbb{R})$ y $(X^{-1}(B)) \in \mathcal{F}$.
   Por lo tanto $B^c \in \mathcal{B}(\mathbb{R})$ y $(X^{-1}(B^c)) = (X^{-1}(B))^c \in \mathcal{F}$.
   Es decir, $B^c \in \mathscr{B}$.

3. Sea $\{B_n, n \in \mathbb{N}\}$ una sucesión en $\mathscr{B}$.
   Es decir, para cada número natural n, $B_n \in \mathcal{B}(\mathbb{R})$ y $(X^{-1}(B_n)) \in \mathcal{F}$.
   Entonces: $(\bigcup_{n=1}^{\infty} B_n) \in \mathcal{B}(\mathbb{R})$ y $(\bigcup_{n=1}^{\infty} (X^{-1}(B_n))) = (X^{-1}(\bigcup_{n=1}^{\infty} B_n)) \in \mathcal{F}$.
   Es decir, $(\bigcup_{n=1}^{\infty} B_n) \in \mathscr{B}$

$\square$

Proposición 1. *Una función $X : \Omega \to \mathbb{R}$ es una **variable aleatoria** si, y sólo si, para cada x en $\mathbb{R}$ se cumple que $X^{-1}(-\infty, x] \in \mathcal{F}$*

Demostración.

$\Rightarrow$) Si X es v.a., entonces claramente se cumple que para cualquier número real x el conjunto $X^{-1}(-\infty, x]$ es un elemento de $\mathcal{F}$.

$\Leftarrow$) Ahora, supongamos que para cada real x, el conjunto $X^{-1}(-\infty, x] \in \mathcal{F}$.

Sean $\mathscr{B}$ y $\mathscr{C}$ las colecciones:

$$\begin{aligned}\mathscr{B} &= \left\{B \in \mathcal{B}(\mathbb{R}) : \left(X^{-1}(B)\right) \in \mathcal{F}\right\} \quad \text{y} \\ \mathscr{C} &= \{(-\infty, x] : x \in \mathbb{R}\}.\end{aligned}$$

Entonces claramente $\mathscr{C} \subseteq \mathscr{B} \subseteq \mathcal{B}(\mathbb{R})$. La primera contención es por hipótesis, y la segunda es por definición de la colección $\mathscr{B}$. Por el Lema 1, $\mathscr{B}$ es una σ-álgebra de subconjuntos de $\mathbb{R}$. Entonces $\mathscr{B}$ es una σ-álgebra que contiene a $\mathscr{C}$.

Por lo tanto $\sigma(\mathscr{C}) = \mathcal{B}(\mathbb{R}) \subseteq \mathscr{B}$. Esto implica que $\mathscr{B} = \mathcal{B}(\mathbb{R})$, y entonces X es variable aleatoria. □

1.2.1. Función de distribución

Toda v.a. tiene asociada una función llamada función de distribución.

Definición 7. *La **función de distribución de una variable aleatoria** X es la función $F_X : \mathbb{R} \to [0,1]$, definida como sigue:*

$$F_X(x) = P\left(X^{-1}(-\infty, x]\right), x \in \mathbb{R}. \tag{1.3}$$

Cuando sea necesario especificar la v.a. en cuestión se escribe $F_X(x)$, pero en general se omite el subíndice X cuando no haya posibilidad de confusión. Observe que la función de distribución de una v.a. está definida sobre la totalidad del conjunto de números reales, y siendo una probabilidad, toma valores en el intervalo $[0,1]$.

Definición 8. *Una función $F : \mathbb{R} \to [0,1]$ es llamada **función de distribución de una variable aleatoria** si cumple las siguientes propiedades:*

1. $\lim_{x\to\infty} F(x) = 1$,
2. $\lim_{x\to-\infty} F(x) = 0$,
3. *Si* $x_1 < x_2$ *entonces* $F(x_1) \leq F(x_2)$ *y*
4. $F(x)$ *es continua por la derecha, es decir,* $\lim_{x\to a^+} F(x) = F(a)$.

1.2.2. Tipos de variables aleatorias

Las variables aleatorias se clasifican en varios tipos dependiendo de las características de la correspondiente función de distribución. Veamos la definición de v.a discreta y continua:

Definición 9. *La variable aleatoria X se llama* ***discreta*** *si su correspondiente función de distribución $F(x)$ es una función constante a trozos.*

Sean $x_1, x_2, ...$ los puntos de discontinuidad de $F(x)$. En cada uno de estos puntos, el tamaño de la discontinuidad es $P(X = x_i) = F(x_i) - F(x_i-) > 0$. A la función $f(x)$ que indica estos incrementos se le llama **función de densidad** de X y se define como sigue:

$$f(x) = \begin{cases} P(X = x) & \text{si } x = x_1, x_2\ ,... \\ 0 & \text{en otro caso.} \end{cases} \tag{1.4}$$

La función de distribución se reconstruye de la forma siguiente:

$$F(x) = \sum_{u \leq x} f(u) \, .$$

Observe que la función de probabilidad $f(x)$ es una función no negativa que suma uno en el sentido $\sum_i f(x_i) = 1$. Recíprocamente, toda función de la forma (1.4) que cumpla estas dos propiedades se le llama función de probabilidad, sin que haya necesariamente una v.a. de por medio.

Veamos ahora el caso continuo.

Definición 10. *La variable aleatoria X se llama* ***continua*** *si su correspondiente función de distribución $F(x)$ es una función continua.*

Definición 11. *La variable aleatoria continua X con función de distribución $F(x)$ se llama* ***absolutamente continua****, si existe una función no negativa e integrable f tal que para cualquier valor de x se cumple*

$$F(x) = \int_{-\infty}^{x} f(u) \, du. \tag{1.5}$$

En tal caso a la función $f(x)$ se le llama función de densidad de X.

La función de densidad de una v.a. absolutamente continua es no negativa y su integral sobre toda la recta real es uno. Recíprocamente, toda función $f(x)$ no negativa que integre uno en $\mathbb{R}$ se llama función de densidad.

1.3. Esperanza condicional

La esperanza condicional es una herramienta esencial en el estudio de procesos estocásticos. Es decir, es importante para desarrollar la intuición necesaria detrás de este concepto.

Antes de enunciar las definiciones de esperanza condicional veamos la siguiente definición:

Definición 12. *Una variable aleatoria $X : \Omega \to \mathbb{R}$ es **integrable** si:*

$$E(|X|) = \int_{\Omega} |X| \, dP < \infty.$$

*Y una variable aleatoria $X : \Omega \to \mathbb{R}$ es de **cuadrado integrable** si:*

$$E\left(|X|^2\right) = \int_{\Omega} |X|^2 dP < \infty.$$

La familia de variables aleatorias integrables es denotado por $L^1(\Omega, \mathcal{F}, P)$, o simplemente L^1. Y la familia de variables aleatorias de cuadrado integrable es denotado por $L^2(\Omega, \mathcal{F}, P)$, o simplemente L^2.

- **Esperanza condicional de una variable aleatoria dado un evento**

Definición 13. *Para cualquier variable aleatoria integrable X y cualquier evento $B \in \mathcal{F}$ tal que $P(B) \neq 0$ la **esperanza condicional de** X **dado** B esta definida por:*

$$E(X|B) = \frac{1}{P(B)} \int_{B} X dP.$$

- **Esperanza condicional de una variable aleatoria dada una variable aleatoria discreta**

El siguiente paso implica condicionar sobre una v.a. discreta η con posibles valores $y_1, y_2, ...$ tal que $P\{\eta = y_n\} \neq 0$ para cada n.

Encontrar el valor de η equivale a encontrar cuál de los eventos $\{\eta = y_n\}$ ocurrirá o no. Condicionar respecto a η, por lo tanto, debe ser lo mismo que condicionar respecto a los eventos $\{\eta = y_n\}$, porque no sabemos por adelantado cual de estos eventos ocurrirá. Necesitamos considerar todas las posibilidades, implicando una sucesión de esperanzas condicionales

$$E\left(X|\left\{\eta = y_1\right\}\right), E\left(X|\left\{\eta = y_2\right\}\right), ...$$

Una manera conveniente de hacer esto es construir una nueva v.a. constante e igualarla a $E\left(X|\left\{\eta = y_n\right\}\right)$ en cada uno de los conjuntos $\{\eta = y_n\}$. Esto nos lleva a la siguiente definición:

Definición 14. *Sea X una variable aleatoria integrable y sea η una variable aleatoria discreta. Entonces la **esperanza condicional de** X **dada** η se define como la variable aleatoria $E\left(X|\eta\right)$ tal que:*

$$E\left(X|\eta\right)(w) = E\left(X|\left\{\eta = y_n\right\}\right) \text{ si } \eta(w) = y_n \ \forall n = 1, 2, ...$$

- **Esperanza condicional de una variable aleatoria dada una variable aleatoria arbitraria**

Definición 15. *Sea X una variable aleatoria integrable y sea η una variable aleatoria arbitraria. Entonces la **esperanza condicional de** X **dada** η se define como la variable aleatoria $E\left(X|\eta\right)$ tal que:*

1. *$E\left(X|\eta\right)$ es $\sigma(\eta)$-medible y*
2. *Para cada $A \in \sigma(\eta)$*

$$\int_A E\left(X|\eta\right) dP = \int_A X dP.$$

- **Esperanza condicional de una variable aleatoria dada una σ-álgebra**

La siguiente definición general de esperanza condicional está basada en la observación de que $E(X|\eta)$ depende sólo de la σ-álgebra generada por η, en lugar de los valores de η.

Definición 16. *Sea X una variable aleatoria integrable en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y sea $\mathcal{G}$ una σ-álgebra contenida en $\mathcal{F}$. Entonces la **esperanza condicional de X dada** $\mathcal{G}$ se define como la variable aleatoria $E(X|\mathcal{G})$ tal que:*

1. *$E(X|\mathcal{G})$ es $\mathcal{G}$-medible y*
2. *Para cada $A \in \mathcal{G}$*

$$\int_A E(X|\mathcal{G})\, dP = \int_A X dP.$$

- **Propiedades generales**

Proposición 2. *Sean a, b números reales arbitrarios, X, Y son variables aleatorias integrables en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y $\mathcal{G}, \mathcal{H}$ son σ-álgebras en Ω contenidas en $\mathcal{F}$. La esperanza condicional tiene las siguientes propiedades:*

1. *$E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$,*
2. *$E(E(X|\mathcal{G})) = E(X)$,*
3. *$E(XY|\mathcal{G}) = XE(Y|\mathcal{G})$ si X es $\mathcal{G}$-medible,*
4. *$E(X|\mathcal{G}) = E(X)$ si X es independiente de $\mathcal{G}$,*
5. *$E(E(X|\mathcal{G})|\mathcal{H}) = E(X|\mathcal{H})$ si $\mathcal{H} \subset \mathcal{G}$ y*
6. *Si $X \geq 0$, entonces $E(X|\mathcal{G}) \geq 0$ casi seguramente*

En la tercera propiedad asumimos que el producto XY es integrable.

Demostración. Ver [2]. □

1.4. Procesos estocásticos

Considere un sistema que puede caracterizarse por estar en cualquiera de un conjunto de estados previamente especificado. Suponga que el sistema evoluciona o cambia de un estado a otro a lo largo del tiempo de acuerdo a cierta ley de movimiento, y sea X_t el estado del sistema al tiempo t. Si se considera que la forma en la que el sistema evoluciona no es determinista, sino provocada por algún mecanismo aleatorio, entonces puede considerarse que X_t es una v.a. para cada valor del índice t. Esta colección de variables aleatorias es la definición de proceso estocástico, y sirve como modelo para representar la evolución aleatoria de un sistema a lo largo del tiempo. En general, las variables aleatorias que conforman un proceso no son independientes entre sí, sino que están relacionadas unas con otras de alguna manera particular. Más precisamente, la definición de proceso estocástico toma como base un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y puede enunciarse de la siguiente forma.

Definición 17. *Un* ***proceso estocástico*** *es una familia de variables aleatorias* $\{X_t : t \in T\}$*, donde* T *es un conjunto de índices.*

El rango común de las variables aleatorias se llama **espacio de estados** (S), $S \subseteq \mathbb{R}$.

El parámetro t se llamará **tiempo**.

Clasificación:

$$\text{Proceso estocástico}\begin{cases}\text{T es finito o numerable}\begin{cases}\text{S finito o numerable}\\ \text{S infinito no-numerable}\end{cases}\\ \text{T es infinito no-numerable}\begin{cases}\text{S finito o numerable}\\ \text{S infinito no-numerable}\end{cases}\end{cases}$$

Un proceso estocástico puede considerarse como una función de dos variables $X : T \times \Omega \to S$ tal que a la pareja (t, w) se le asocia el estado $X(t, w)$ lo cual también puede escribirse como $X_t(w)$. Para cada valor de t en T, el mapeo $w \to X_t(w)$ es una v.a., mientras que para cada w en Ω fijo, la función $t \to X_t(w)$ es llamada una trayectoria o realización del proceso. Es decir, a cada w del espacio muestral le corresponde una trayectoria del proceso.

Es por ello que a veces se define un proceso estocástico como una función aleatoria.

Si A es un conjunto de estados, el evento $X_t \in A$ corresponde a la situación en donde al tiempo t el proceso toma algún valor dentro del conjunto A. En particular, $(X_t = x)$ es el evento en donde al tiempo t el proceso se encuentra en el estado x.

Los diferentes tipos de procesos estocásticos se obtienen al considerar las distintas posibilidades para: el conjunto de índices, el espacio de estados, las características de las trayectorias, y principalmente las relaciones de dependencia entre las variables aleatorias que conforman el proceso. A lo largo del texto estudiaremos y especificaremos con mayor detalle el proceso estocástico denominado Martingala.

Capítulo 2

Martingalas

2.1. Martingalas a tiempo discreto

Existen varias aprobaciones para el término **martingala**. Una de ellas hace referencia a un tipo de proceso estocástico que básicamente cumple la identidad $E\left(X_{n+1}|X_0 = x_0, X_1 = x_1, ..., X_n = x_n\right) = x_n$. En otras palabras, esta igualdad significa que el estado promedio del proceso al tiempo futuro $n+1$ es el valor del proceso en su último momento observado, es decir, x_n. Esto es, se trata de una ley de movimiento aleatorio que es equilibrada pues en promedio el sistema no se mueve del último momento observado. A estos procesos también se les conoce como procesos de juegos justos pues si se considera una sucesión infinita de apuestas sucesivas y si X_n denota el capital de uno de los jugadores al tiempo n, entonces la propiedad de martingala establece que el juego es justo pues en promedio el jugador no pierde ni gana en cada apuesta. Parte de la motivación para el estudio de este tipo de procesos fue demostrar la inexistencia de estrategias ventajosas en este juego de apuestas. El concepto de martingala fue incorporado a la teoría de la probabilidad por Paul Lèvy, y buena parte de su desarrollo inicial fue realizado por Joseph Doob.

En este capítulo se presenta un estudio sobre la teoría de martingalas a tiempo discreto.

Emplearemos las siglas c.s. para referirnos a casi seguramente.

Antes de dar la definición de Martingala, enunciemos las siguientes definiciones:

Definición 18. *Una sucesión de σ-álgebras $\{\mathcal{F}_n, n \in \mathbb{N}\}$ sobre Ω tal que*

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset ... \subset \mathcal{F}_n \subset ... \subset \mathcal{F}$$

*es llamada una **filtración**.*

Donde $\mathcal{F}_n$ representa nuestro conocimiento acerca del experimento aleatorio hasta el tiempo n. Asi, $\mathcal{F}_n$ contiene a todos los eventos A, tal que al tiempo n es posible decir si el evento A ha ocurrido o no.

Definición 19. *Una sucesión de variables aleatorias $\{X_n, n \in \mathbb{N}\}$ es **adaptada** a una filtración $\{\mathcal{F}_n, n \in \mathbb{N}\}$ si X_n es $\mathcal{F}_n$-medible y una sucesión de variables aleatorias $\{X_n, n \in \mathbb{N}\}$ es **previsible** a una filtración $\{\mathcal{F}_n, n \in \mathbb{N}\}$ si X_n es $\mathcal{F}_{n-1}$-medible para cada $n = 1, 2, ...$*

Como ya se mencionó, el concepto de martingala describe un juego justo, lo cual se observa en la tercera condición de la siguiente definición.

Definición 20. *Una sucesión de variables aleatorias $\{X_n, n \in \mathbb{N}\}$ es llamada **martingala** con respecto a una filtración $\{\mathcal{F}_n, n \in \mathbb{N}\}$ si:*

1. *$\{X_n, n \in \mathbb{N}\}$ es integrable,*
2. *$\{X_n, n \in \mathbb{N}\}$ es adaptada a $\{\mathcal{F}_n, n \in \mathbb{N}\}$ y*
3. *Para cada $n = 1, 2, ...$*

$$E\left(X_{n+1}|\mathcal{F}_n\right) = X_n \ c.s. \tag{2.1}$$

*Cuando en lugar de (2.1) se cumple la desigualdad $E\left(X_{n+1}|\mathcal{F}_n\right) \geq X_n$ c.s., entonces el proceso es una **submartingala** y si $E\left(X_{n+1}|\mathcal{F}_n\right) \leq X_n$ c.s., entonces es una **supermartingala***

La desigualdad $E\left(X_{n+1}|\mathcal{F}_n\right) \geq X_n$, correspondiente a la definición de submartingala, equivale a un juego favorable al jugador y la desigualdad contraria, el caso de supermartingala, corresponde a un juego desfavorable al jugador.

Ejemplo 1. *Sea* $\{X_n, n \in \mathbb{N}\}$ *una sucesión de variables aleatorias independientes, idénticamente distribuidas e integrables con* $E(X_n) = 0$ *y sea* $\{\mathcal{F}_n, n \in \mathbb{N}\}$ *una filtración. Si* $S_n = X_1 + ... + X_n$*, entonces* S_n *es una martingala. Para comprobar esto, verifiquemos* (2.1)*:*

$$\begin{aligned} E(S_{n+1}|\mathcal{F}_n) &= E(S_n + X_{n+1}|\mathcal{F}_n) \\ &= S_n + E(X_{n+1}|\mathcal{F}_n) \\ &= S_n + E(X_n) \\ &= S_n. \end{aligned}$$

Ejemplo 2. *Consideremos nuevamente una sucesión de variables aleatorias independientes idénticamente distribuidas* $\{X_n, n \in \mathbb{N}\}$ *tal que* $P(X_n = 1) = \frac{1}{2}$ *y* $P(X_n = -1) = \frac{1}{2}$*, y con filtración* $\{\mathcal{F}_n, n \in \mathbb{N}\}$*. Considere la suma* $S_n = X_1 + ... + X_n$*, es decir,* S_n *es una martingala que representa el total de ganancias en una serie de* n *apuestas justas de una unidad monetaria.*

Suponga ahora que el monto de cada apuesta no es uno, sino una cantidad a_n *para la* n*-ésima apuesta. Supondremos que* a_n *es una v.a. que el jugador determina a partir de la información de las* $n-1$ *apuestas previas, y por lo tanto es una variable* $\mathcal{F}_{n-1}$*-medible, es decir, se trata de un proceso previsible. A la colección de variables aleatorias* $\{a_n\}$ *con esta característica se le llama una* ***estrategia de juego****. Bajo una de estas estrategias, el capital del jugador después de la* n*-ésima apuesta es ahora* $A_n = a_1X_1 + ... + a_nX_n$*, que es* $\mathcal{F}_n$*-medible. Bajo la hipótesis de que la estrategia de juego consta de variables aleatorias acotadas, se cumple que el proceso* $\{A_n, n \in \mathbb{N}\}$ *es integrable y cumple además la propiedad* (2.1) *pues*

$$\begin{aligned} E(A_{n+1}|\mathcal{F}_n) &= E(A_n + a_{n+1}X_{n+1}|\mathcal{F}_n) \\ &= A_n + a_{n+1}E(X_{n+1}|\mathcal{F}_n) \\ &= A_n + a_{n+1}E(S_{n+1} - S_n|\mathcal{F}_n) \\ &= A_n + a_{n+1}(E(S_{n+1}|\mathcal{F}_n) - S_n) \\ &= A_n. \end{aligned}$$

Esto quiere decir que bajo cualquier estrategia de juego, el proceso de ganancias $\{A_n\}$ *es una martingala siempre y cuando el proceso original* $\{X_n\}$ *lo sea. Para los apostadores es importante saber que no existe una estrategia* *de juego que convierta un juego justo en un juego favorable o desfavorable al jugador. El mismo análisis demuestra que si el proceso original $\{X_n\}$ es una submartingala o supermartingala y la estrategia de juego consta de variables aleatorias no negativas y acotadas, entonces el proceso $\{A_n\}$ sigue siendo una submartingala o supermartingala.*

2.2. Resultados básicos

- **Teorema de Paro**

En muchos juegos de azar tenemos la opción de retirarnos en cualquier momento. El número de rondas jugadas antes de retirarnos del juego será denotado por τ. Éste puede ser fijo, digamos $\tau = 10$ si uno decide avanzar o parar de jugar después de 10 rondas, pase lo que pase. Pero en general la desición de retirarse o no, se toma después de cada ronda dependiendo del conocimiento acumulado hasta entonces. Por lo tanto, τ es una v.a. con valores en $\mathbb{N} \cup \{\infty\}$, se incluye a ∞ sólo para cubrir la posibilidad teórica de que el jugador nunca pare. En cada paso n decidimos si paramos de jugar o no, es decir, si $\tau = n$. Por lo tanto el evento $\{\tau = n\}$ debe estar en la σ-álgebra $\mathcal{F}_n$ representando nuestro conocimiento hasta el tiempo n.

Definición 21. *Una variable aleatoria τ con valores en $\mathbb{N} \cup \{\infty\}$ es llamada **tiempo de paro** (con respecto a una filtración $\{\mathcal{F}_n, n \in \mathbb{N}\}$) si para cada $n = 1, 2, ...$*

$$\{\tau = n\} \in \mathcal{F}_n.$$

Sea $\{X_n, n \in \mathbb{N}\}$ una sucesión de variables aleatorias adaptada a una filtración $\{\mathcal{F}_n, n \in \mathbb{N}\}$ y τ un tiempo de paro(con respecto a la filtración). Supongamos que $\{X_n\}$ representa nuestras ganancias (o pérdidas) después de n rondas de un juego. Si tú decides retirarte después de τ rondas, entonces tus ganancias totales debe ser X_τ. En este caso tus ganancias después de n rondas deben ser en realidad el proceso $X_{n\wedge\tau}$, donde $\tau \wedge n = \min(\tau, n)$. A este tipo de procesos, se les llama **procesos detenidos**.

El siguiente teorema denominado teorema de paro es un resultado clásico para las martingalas, demostramos que una martingala (submartingala o supermartingala) detenida sigue siendo una martingala (submartingala o supermartingala)

Teorema 1. *(Teorema de paro)*
Sea τ un tiempo de paro

1. *Si $\{X_n, n \in \mathbb{N}\}$ es una martingala, entonces $X_{n\wedge\tau}$ también es una martingala.*
2. *Si $\{X_n, n \in \mathbb{N}\}$ es una submartingala, entonces $X_{n\wedge\tau}$ también es una submartingala.*
3. *Si $\{X_n, n \in \mathbb{N}\}$ es una supermartingala, entonces $X_{n\wedge\tau}$ también es una supermartingala.*

Sólo demostraremos el caso cuando $\{X_n, n \in \mathbb{N}\}$ es una martingala, pues para submartingalas y supermartingalas se demuestra de la misma manera.

Demostración. En primer lugar, para toda $n \geq 0$, tenemos que $X_{n\wedge\tau}$ es integrable porque $X_{n\wedge\tau} = X_\tau 1_{(\tau<n)} + X_n 1_{(\tau\geq n)}$. Además, $\{\tau \geq n\} \in \mathcal{F}_{n-1}$ pues su complemento corresponde al evento $\{\tau < n\}$. Finalmente, tenemos que para toda $n \geq 0$

$$
\begin{aligned}
E\left(X_{(n+1)\wedge\tau}|\mathcal{F}_n\right) &= E\left(X_\tau 1_{(\tau<(n+1))} + X_{(n+1)}1_{(\tau\geq(n+1))}|\mathcal{F}_n\right) \\
&= X_\tau 1_{(\tau<(n+1))} + 1_{(\tau\geq(n+1))}E\left(X_{(n+1)}|\mathcal{F}_n\right) \\
&= X_\tau 1_{(\tau<(n+1))} + 1_{(\tau\geq(n+1))}X_n \\
&= X_\tau 1_{(\tau<n)} + X_n 1_{(\tau=n)} + 1_{(\tau\geq n)}X_n - X_n 1_{(\tau=n)} \\
&= X_\tau 1_{(\tau<n)} + 1_{(\tau\geq n)}X_n \\
&= X_{n\wedge\tau}.
\end{aligned}
$$

$\square$

- **Desigualdades de Doob para Martingalas**

En esta sección mostramos algunas desigualdades clásicas para martingalas como lo son las Desigualdades de Doob. Estas desigualdades nos darán las herramientas necesarias para el estudio de la convergencia de martingalas. Mostramos un resultado clásico conocido como el Teorema de convergencia de Doob para Martingalas, este teorema nos proporciona el límite de una martingala.

Proposición 3. *(Desigualdad máximal de Doob)*
Supongamos que $\{X_n, n \in \mathbb{N}\}$, es una submartingala no negativa con respecto a la filtración $\{\mathcal{F}_n, n \in \mathbb{N}\}$. Entonces para cada $\lambda > 0$

$$\lambda P\left(\max_{k\leq n}\{X_k\} \geq \lambda\right) \leq E\left(X_n 1_{(\max_{k\leq n}\{X_k\}\geq\lambda)}\right),$$

donde 1_A es la función indicadora del conjunto A.

Demostración. Denotamos por

$$X_n^* = \max_{k\leq n}\{X_k\}.$$

Para $\lambda > 0$ definimos

$$\tau = \begin{cases} \min\{k \leq n\} & \text{si existe } k \leq n,\ n \in \mathbb{N} \text{ tal que } X_k \geq \lambda \\ n & \text{en otro caso} \end{cases}$$

Entonces τ es un tiempo de paro tal que $\tau \leq n$ c.s.

Como $\{X_n, n \in \mathbb{N}\}$ es una submartingala,

$$\begin{aligned} E(X_{n+1}|\mathcal{F}_n) \geq X_n \quad &\Rightarrow \quad E(E(X_{n+1}|\mathcal{F}_n)) \geq E(X_n) \\ &\Rightarrow \quad E(X_{n+1}) \geq E(X_n) \end{aligned}$$

Luego,

$$E(X_n) \geq E(X_\tau)$$

Pero $E(X_\tau) = E\left(X_\tau 1_{(X_n^*\geq\lambda)}\right) + E\left(X_\tau 1_{(X_n^*<\lambda)}\right)$

Observemos que si $X_n^* \geq \lambda$, entonces $X_\tau \geq \lambda$. Por otra parte, si $X_n^* < \lambda$, entonces $\tau = n$ y también $X_\tau = X_n$. Además

$$\begin{aligned} E\left(X_\tau 1_{(X_n^*\geq\lambda)}\right) &\geq E\left(\lambda 1_{(X_n^*\geq\lambda)}\right) \\ &= \lambda E\left(1_{(X_n^*\geq\lambda)}\right) \\ &= \lambda P\{X_n^* \geq \lambda\} \end{aligned}$$

Entonces

$$E(X_n) \geq E(X_\tau) \geq \lambda P\{X_n^* \geq \lambda\} + E\left(X_n 1_{(X_n^* < \lambda)}\right)$$

Por lo tanto

$$\lambda P\{X_n^* \geq \lambda\} \leq E(X_n) - E\left(X_n 1_{(X_n^* < \lambda)}\right) = E\left(X_n 1_{(X_n^* \geq \lambda)}\right)$$

□

El siguiente Teorema se conoce como desigualdad máximal de Doob en L^2, pero antes de enunciarlo veamos el siguiente Lema:

Lema 2. *Si* $\eta : \Omega \to [0, \infty)$ *es una variable aleatoria no negativa de cuadrado integrable, entonces*

$$E\left(\eta^2\right) = 2\int_0^\infty tP(\eta > t)dt.$$

Demostración. Ver [2]. □

Teorema 2. *(Desigualdad máximal de Doob en L^2)*
Si $\{X_n, n \in \mathbb{N}\}$, *es una submartingala no negativa de cuadrado integrable con respecto a la filtración* $\{\mathcal{F}_n, n \in \mathbb{N}\}$. *Entonces*

$$E\left|\max_{k \leq n}\{X_k\}\right|^2 \leq 4E|X_n|^2. \tag{2.2}$$

Demostración. Denotamos por

$$X_n^* = \max_{k \leq n}\{X_k\}.$$

Luego

$$
\begin{aligned}
E|X_n^*|^2 &= 2\int_0^\infty tP\left(X_n^* > t\right)dt \\
&\quad \text{(Por el Lema anterior)} \\
&\leq 2\int_0^\infty E\left(X_n 1_{(X_n^* \geq t)}\right)dt \\
&\quad \text{(Por la Proposición 3)} \\
&= 2\int_0^\infty \left[\int_{\{X_n^* \geq t\}} X_n dP\right]dt \\
&= 2\int_0^\infty \left[\int_0^{X_n^*} X_n dP\right]dt \\
&= 2\int_\Omega X_n \left[\int_0^{X_n^*} dt\right]dP \\
&\quad \text{(Por el Teorema de Fubini)} \\
&= 2\int_\Omega X_n X_n^* dP \\
&= 2E\left(X_n X_n^*\right) \\
&\leq 2\left(E|X_n|^2\right)^{\frac{1}{2}}\left(E|X_n^*|^2\right)^{\frac{1}{2}} \\
&\quad \text{(Por la desigualdad de} \\
&\quad \text{Cauchy-Schwarz)} \\
\text{Entonces:} & \\
E|X_n^*|^2 &\leq 2\left(E|X_n|^2\right)^{\frac{1}{2}}\left(E|X_n^*|^2\right)^{\frac{1}{2}} \\
\Rightarrow \frac{E|X_n^*|^2}{\left(E|X_n^*|^2\right)^{\frac{1}{2}}} &\leq 2\left(E|X_n|^2\right)^{\frac{1}{2}} \\
\Rightarrow \left(E|X_n^*|^2\right)^{\frac{1}{2}} &\leq 2\left(E|X_n|^2\right)^{\frac{1}{2}} \\
\Rightarrow \left(E|X_n^*|^2\right) &\leq 4E|X_n|^2
\end{aligned}
$$

□

Definición 22. *Sea $\{X_n, n \in \mathbb{N}\}$ una sucesión adaptada de variables aleatorias y a, b dos números reales tal que $a < b$. Definimos una* ***estrategia de juego*** *$\alpha_1, \alpha_2, ...,$ por*

$$\alpha_1 = 0$$

y para $n = 1, 2, ...$

$$\alpha_{n+1} = \begin{cases} 1 & \textit{si } \alpha_n = 0 \textit{ y } X_n < a, \\ 1 & \textit{si } \alpha_n = 1 \textit{ y } X_n \leq b, \\ 0 & \textit{en otro caso.} \end{cases}$$

La Definición 22 es conocida como **la estrategia de cruzamiento**. Para cada $k = 1, 2, ...,$ tal que $\alpha_k = 1$ y $\alpha_{k+1} = 0$ es llamado **cruzamiento** del intervalo $[a, b]$. Los cruzamientos forman una sucesión (finita o infinita) creciente

$$u_1 < u_2 < ...$$

El número de cruzamientos realizados hasta el tiempo n, esto es, la mayor k tal que $u_k \leq n$ la denotaremos por $U_n[a, b]$ (definimos $U_n[a, b] = 0$ si no existe tal k).

Lema 3. *(Desigualdad de cruzamiento)*
Si $\{X_n, n \in \mathbb{N}\}$ es una martingala y $a < b$, entonces

$$(b - a) E(U_n[a, b]) \leq E\left((X_n - a)^-\right).$$

Por x^- denotamos la parte negativa de un número real x, es decir

$$x^- = \max\{0, -x\}.$$

Demostración. Ver [2]. □

- **Teorema de convergencia de Doob para Martingalas**

Teorema 3. *Supongamos que $\{X_n, n \in \mathbb{N}\}$ es una supermartingala con respecto a la filtración $\{\mathcal{F}_n, n \in \mathbb{N}\}$ tal que*

$$\sup_n E(|X_n|) < \infty$$

Entonces existe una variable aleatoria integrable X tal que

$$\lim_{n\to\infty} X_n = X \quad c.s.$$

Nota 3. *En particular, el Teorema es válido para martingalas porque toda martingala es una supermartingala. Esto también es válido para submartingalas, pues*

$\{X_n, n \in \mathbb{N}\}$ *es una submartingala si y sólo si* $\{-X_n, n \in \mathbb{N}\}$ *es una supermartingala.*

Nota 4. *Observemos que aunque las variables aleatorias* X_n *así como el límite* X *son integrables, solamente se afirma que* X_n *converge a* X *c.s. Notemos que no se asegura la convergencia en* L^1*.*

Demostración. Por el Lema 3

$$E\left(U_n\left[a,b\right]\right) \leq \frac{E\left(\left(X_n - a\right)^-\right)}{(b-a)} \leq \frac{M + |a|}{b-a} < \infty$$

donde

$$M = \sup_n E\left(|X_n|\right)$$

Como $U_n\left[a,b\right]$ es una sucesión no decreciente

$$E\left(\lim_{n\to\infty} U_n\left[a,b\right]\right) = \lim_{n\to\infty} E\left(U_n\left[a,b\right]\right) \leq \frac{M+|a|}{b-a} < \infty$$

Esto implica que

$$P\left\{\lim_{n\to\infty} U_n\left[a,b\right] < \infty\right\} = 1$$

para $a < b$.
Como el conjunto de números racionales es numerable

$$P\left\{\lim_{n\to\infty} U_n\left[a,b\right] < \infty \text{ para todo racional } a < b\right\} = 1 \tag{2.3}$$

Afirmamos que la sucesión X_n converge a X c.s.
Supongamos que

$$\liminf_n X_n < \limsup_n X_n$$

Entonces

$$U_n\left[a,b\right] \to \infty \quad \text{si} \quad \liminf_n X_n < a < b < \limsup_n X_n$$

Porque a, b pueden ser elegidos números racionales, esto contradice (2.3), probando la afirmación. Sólo resta probar que el límite X es una v.a. integrable,

$$\begin{aligned} E(|X|) &= E\left(\liminf_n |X_n|\right) \\ &\leq \liminf_n E(|X_n|) \quad \text{Por el Lema de Fatou} \\ &< \sup_n E(|X_n|) \\ &< \infty. \end{aligned}$$

$\square$

Corolario 1. *Sea $\{X_n, n \in \mathbb{N}\}$ una submartingala acotada superiormente por una constante M. Entonces, $\{X_n, n \in \mathbb{N}\}$ converge c.s. hacia X_∞ con $E(X_\infty|\mathcal{F}_n) \geq X_n$ c.s.*

Lo mismo ocurre si $\{X_n, n \in \mathbb{N}\}$ es una supermartingala acotada inferiormente por una constante m, entonces, $\{X_n, n \in \mathbb{N}\}$ converge c.s. hacia X_∞ con $E(X_\infty|\mathcal{F}_n) \leq X_n$ c.s.

Demostración. Ver [1]. $\square$

- **Teorema de Robbins-Siegmund**

Teorema 4. *(Teorema de Robbins-Siegmund)*
Sean $\{A_n, n \in \mathbb{N}\}, \{B_n, n \in \mathbb{N}\}$ y $\{C_n, n \in \mathbb{N}\}$ tres sucesiones de variables aleatorias positivas y adaptadas a $\mathcal{F}$. Supongamos que A_0 es finito c.s. y para todo $n \geq 0$

$$E(A_{n+1}|\mathcal{F}_n) \leq A_n + B_n - C_n \ \ c.s.$$

Entonces,

sobre $\Gamma = \{\sum_{n=0}^{\infty} B_n < \infty\}$, $\{A_n, n \in \mathbb{N}\}$ converge c.s. hacia una v.a A_∞ finita c.s. y

$$\sum_{n=0}^{\infty} C_n < \infty \ \ c.s.$$

Demostración. Ver [1]. $\square$

Corolario 2. *Sean* $\{A_n, n \in \mathbb{N}\}, \{B_n, n \in \mathbb{N}\}$ *y* $\{C_n, n \in \mathbb{N}\}$ *tres sucesiones de variables aleatorias positivas y adaptadas a* $\mathcal{F}$ *y sea* $\{a_n, n \in \mathbb{N}\}$ *una sucesión positiva estrictamente creciente, adaptada a* $\mathcal{F}$ *con* $a_n \to a_\infty$ *c.s. Supongamos que* A_0 *es finito c.s. y para todo* $n \geq 0$

$$E\left(A_{n+1}|\mathcal{F}_n\right) \leq A_n + B_n - C_n \;\; c.s.$$

Entonces, sobre $\Lambda = \left\{\sum_{n=0}^{\infty} \frac{B_n}{a_n} < \infty\right\}$, *tenemos*

$$\sum_{n=0}^{\infty} \frac{C_n}{a_n} < \infty \;\; c.s.$$

1. *Sobre* $\Lambda \cap \{a_\infty < \infty\}$, $\{A_n, n \in \mathbb{N}\}$ *converge c.s. hacia una v.a* A_∞ *finita c.s.*

2. *Sobre* $\Lambda \cap \{a_\infty = \infty\}$, $A_n = o\left(a_n\right)$ *y* $A_{n+1} = o\left(a_n\right)$ *c.s.*

Demostración. Ver [1]. □

- **Martingalas de cuadrado integrable**

Definición 23. *Sea* $\{X_n, n \in \mathbb{N}\}$ *una martingala de cuadrado integrable. A la sucesión* $\{\langle X\rangle_n, n \in \mathbb{N}\}$ *definida por*

$$\langle X\rangle_0 = 0 \quad y$$

Para toda $n > 0$

$$\langle X\rangle_{n+1} = E\left[\left(X_{n+1} - X_n\right)^2|\mathcal{F}_n\right] + \langle X\rangle_n$$

se le llama ***proceso creciente*** *asociado a* $\{X_n, n \in \mathbb{N}\}$.

Lema 4. *Sea* $\{X_n, n \in \mathbb{N}\}$ *una martingala de cuadrado integrable y sea* $\{\langle X\rangle_n, n \in \mathbb{N}\}$ *su proceso creciente. Si*

$$Y_n = X_n^2 - \langle X\rangle_n \;\; para\ toda \;\; n \geq 0,$$

entonces $\{Y_n, n \in \mathbb{N}\}$ *es una martingala donde* $E\left(X_n^2\right) = E\left(X_0^2\right) + E\left(\langle X\rangle_n\right)$

Demostración.

$$\begin{aligned}
E\left(Y_n|\mathcal{F}_{n-1}\right) &= E\left(X_n^2 - \langle X\rangle_n|\mathcal{F}_{n-1}\right)\\
&= E\left(X_n^2|\mathcal{F}_{n-1}\right) - E\left(\langle X\rangle_n|\mathcal{F}_{n-1}\right)\\
&= E\left(X_n^2|\mathcal{F}_{n-1}\right) - \langle X\rangle_n\\
&= E\left(X_n^2|\mathcal{F}_{n-1}\right) - \left[E\left[(X_n - X_{n-1})^2|\mathcal{F}_{n-1}\right] + \langle X\rangle_{n-1}\right]\\
&= E\left(X_n^2 - X_n^2 + 2X_nX_{n-1} - X_{n-1}^2|\mathcal{F}_{n-1}\right) - \langle X\rangle_{n-1}\\
&= E\left(2X_nX_{n-1}|\mathcal{F}_{n-1}\right) - E\left(X_{n-1}^2|\mathcal{F}_{n-1}\right) - \langle X\rangle_{n-1}\\
&= 2X_{n-1}E\left(X_n|\mathcal{F}_{n-1}\right) - X_{n-1}^2 - \langle X\rangle_{n-1}\\
&= 2X_{n-1}^2 - X_{n-1}^2 - \langle X\rangle_{n-1}\\
&= X_{n-1}^2 - \langle X\rangle_{n-1}\\
&= Y_{n-1}
\end{aligned}$$

Por lo tanto $\{Y_n, n \in \mathbb{N}\}$ es una martingala de esperanza

$$E\left(Y_n\right) = E\left(Y_0\right) = E\left(X_0^2 - \langle X\rangle_0\right) = E\left(X_0^2\right)$$

Entonces:

$$E\left(Y_n\right) = E\left(X_n^2 - \langle X\rangle_n\right) = E\left(X_0^2\right)$$

Así:

$$E\left(X_n^2\right) = E\left(X_0^2\right) + E\left(\langle X\rangle_n\right)$$

□

Corolario 3. *Sea $\{X_n, n \in \mathbb{N}\}$ una martingala acotada en L^2. Entonces $\{X_n, n \in \mathbb{N}\}$ converge c.s. y en L^2.*

Demostración. Ver [1]. □

- **Ley de los grandes números**

La ley de los grandes números para martingalas generaliza la ley de los grandes números para las variables aleatorias independientes e integrables. Denotamos por

$$\langle X\rangle_\infty = \lim_{n\to\infty} \langle X\rangle_n.$$

Teorema 5. *(Primera ley de los grandes números)*
Sea $\{X_n, n \in \mathbb{N}\}$ una martingala de cuadrado integrable y sea $\{\langle X\rangle_n, n \in \mathbb{N}\}$ su proceso creciente.

1. *sobre* $\{\langle X\rangle_\infty < \infty\}$, $\{X_n, n \in \mathbb{N}\}$ *converge c.s. hacia una variable aleatoria* X_∞ *de cuadrado integrable.*

2. *sobre* $\{\langle X\rangle_\infty = \infty\}$ *tenemos*

$$\lim_{n\to\infty} \frac{X_n}{\langle X\rangle_n} = 0 \ \ c.s.$$

Demostración. Sea $a \in \mathbb{R}$ con $a > 0$ y sea T_a el tiempo de entrada de $\left\{\langle X\rangle_{n+1}, n \in \mathbb{N}\right\}$ en el intervalo $[a, \infty)$, $T_a = \inf\left\{n \geq 0, \langle X\rangle_{n+1} \geq a\right\}$.
En primer lugar, vemos que T_a es un tiempo de paro, pues

$$\{T_a = n\} = \left\{\langle X\rangle_{n+1} \geq a, \langle X\rangle_j < a \ \text{ con } \ j \leq n\right\} \in \mathcal{F}_n$$

$$\text{y} \quad \langle X\rangle_{n\wedge T_a} < a$$

ya que si $n < T_a, \langle X\rangle_n < a$ y si $T_a < n, \langle X\rangle_{T_a} < a$.

Luego, por el Teorema 1, $X_{n\wedge T_a}$ es una martingala de cuadrado integrable y por el Lema 4 está acotada en L^2 porque

$$E\left(X_{n\wedge T_a}^2\right) = E\left(X_0^2\right) + E\left(\langle X\rangle_{n\wedge T_a}\right) < E\left(X_0^2\right) + a < \infty$$

$$\begin{aligned} &\Longrightarrow \ X_{n\wedge T_a} \ \text{ es acotada en } \ L^2 \\ &\Longrightarrow \ X_{n\wedge T_a} \ \text{ converge c.s. y en } \ L^2 \end{aligned}$$

Por lo tanto $X_{n\wedge T_a} \to X_\infty$ c.s.

Para $\{T_a = \infty\}$, $X_{n\wedge T_a} = X_n \to X_\infty$ c.s.
Pero, $\forall\ p \in \mathbb{N}$, $T_p = \infty \Leftrightarrow \langle X\rangle_{n+1} < p \ \forall\ n \geq 0$

Veamos

$$\{\langle X\rangle_\infty < \infty\} = \bigcup_{p=0}^{\infty} \{T_p = \infty\}$$

$$
\begin{aligned}
\subseteq) \text{ Sea } w \in \{\langle X\rangle_\infty < \infty\} &\Longrightarrow \langle X\rangle_\infty (w) < \infty \\
&\Longrightarrow \exists p \in \mathbb{N} : \langle X\rangle_\infty (w) < p \\
&\Longrightarrow \exists p \in \mathbb{N} : \forall n \ \in \mathbb{N}\, \langle X\rangle_{n+1} (w) \le \langle X\rangle_\infty (w) < p \\
&\Longrightarrow w \in \{T_p = \infty\} \subseteq \bigcup_{p=0}^{\infty} \{T_p = \infty\}
\end{aligned}
$$

$$
\begin{aligned}
\supseteq) \text{ Sea } w \in \bigcup_{p=0}^{\infty} \{T_p = \infty\} &\Longrightarrow w \in \{T_p = \infty\} \\
&\Longrightarrow T_p(w) = \infty \\
&\Longrightarrow \langle X\rangle_{n+1} (w) < p \\
&\Longrightarrow \langle X\rangle_{n+1} (w) \to \langle X\rangle_\infty (w) < \infty \\
&\Longrightarrow w \in \{\langle X\rangle_\infty < \infty\}
\end{aligned}
$$

Por lo tanto sobre $\{\langle X\rangle_\infty < \infty\}$, $\{X_n, n \in \mathbb{N}\}$ converge c.s. hacia una variable aleatoria X_∞ de cuadrado integrable

Para la parte 2., tenemos que

$$\forall \ n \ge 0 \ \ E\left(X_{n+1}^2 | \mathcal{F}_n\right) = X_n^2 + \Delta\langle X\rangle_{n+1}$$

donde $\Delta\langle X\rangle_{n+1} = \langle X\rangle_{n+1} - \langle X\rangle_n$, pues

$$
\begin{aligned}
X_n^2 + \Delta\langle X\rangle_{n+1} &= X_n^2 + \langle X\rangle_{n+1} - \langle X\rangle_n \\
&= X_n^2 + E\left((X_{n+1} - X_n)^2 | \mathcal{F}_n\right) \\
&= X_n^2 + E\left(X_{n+1}^2 - 2X_{n+1}X_n + X_n^2 | \mathcal{F}_n\right) \\
&= X_n^2 + E\left(X_{n+1}^2 | \mathcal{F}_n\right) - 2X_n E\left(X_{n+1} | \mathcal{F}_n\right) + X_n^2 \\
&= X_n^2 + E\left(X_{n+1}^2 | \mathcal{F}_n\right) - 2X_n X_n + X_n^2 \\
&= E\left(X_{n+1}^2 | \mathcal{F}_n\right)
\end{aligned}
$$

Además, $\forall \ 0 < \gamma < 1$, sean

$$A_n = \frac{X_n^2}{\left(\langle X\rangle_n\right)^{1+\gamma}} \quad \text{y} \quad B_n = \frac{\Delta\langle X\rangle_{n+1}}{\left(\langle X\rangle_{n+1}\right)^{1+\gamma}}$$

Entonces

$$\begin{aligned}
E\left(A_{n+1}|\mathcal{F}_n\right) &= E\left(\frac{X_{n+1}^2}{\left(\langle X\rangle_{n+1}\right)^{1+\gamma}}|\mathcal{F}_n\right) \\
&= \frac{1}{\left(\langle X\rangle_{n+1}\right)^{1+\gamma}}E\left(X_{n+1}^2|\mathcal{F}_n\right) \\
&= \frac{1}{\left(\langle X\rangle_{n+1}\right)^{1+\gamma}}\left(X_n^2+\Delta\langle X\rangle_{n+1}\right) \\
&= \frac{X_n^2}{\left(\langle X\rangle_{n+1}\right)^{1+\gamma}}+\frac{\Delta\langle X\rangle_{n+1}}{\left(\langle X\rangle_{n+1}\right)^{1+\gamma}} \\
&\leq A_n+B_n \quad \text{c.s.}
\end{aligned}$$

Sin embargo, tenemos para n suficientemente grande $\langle X\rangle_n \geq 1$ y podemos encontrar $\alpha > 0$ tal que

$$\sum_{n=0}^{\infty} B_n \leq \alpha + \int_1^{\infty}\frac{1}{x^{1+\alpha}}dx < \infty \quad \text{c.s.}$$

Luego, del Teorema 4 con $\{C_n, n\in\mathbb{N}\} = \{0\}$ tenemos que $\{A_n, n\in\mathbb{N}\}$ converge c.s. hacia una v.a. A_∞ finita c.s.

Podemos por lo tanto concluir que $\frac{X_n}{\langle X\rangle_n} \to 0$ c.s. porque $\langle X\rangle_n \to \infty$ c.s. y

$$\begin{aligned}
\left(\frac{X_n}{\langle X\rangle_n}\right)^2 &= \frac{X_n^2\left(\frac{\left(\langle X\rangle_n\right)^{1+\gamma}}{\left(\langle X\rangle_n\right)^{1+\gamma}}\right)}{\langle X\rangle_n^2} \\
&= \frac{A_n(\langle X\rangle_n)^{1+\gamma}}{\langle X\rangle_n^2} \\
&= \frac{A_n}{\left(\langle X\rangle_n\right)^{1-\gamma}}
\end{aligned}$$

también converge a cero c.s.

□

Antes de enunciar la segunda ley de los grandes números denotemos a

$$f_n = \frac{\langle X\rangle_{n+1} - \langle X\rangle_n}{\langle X\rangle_{n+1}} = \frac{\Delta\langle X\rangle_{n+1}}{\langle X\rangle_{n+1}}$$

donde $\{f_n, n \in \mathbb{N}\}$ es adaptada a $\{\mathcal{F}_n, n \in \mathbb{N}\}$ con $0 < f_n < 1$.

Teorema 6. *(Segunda Ley de los grandes números)*
Sea $\{X_n, n \in \mathbb{N}\}$ una martingala de cuadrado integrable y sea $\{\langle X\rangle_n, n \in \mathbb{N}\}$ su proceso creciente. Si $\langle X\rangle_n \to \infty$ c.s., tenemos para todo $\gamma > 0$

$$\left(\frac{X_n^2}{\langle X\rangle_n}\right) = o\left((\log\langle X\rangle_n)^{1+\gamma}\right) \quad c.s.$$

$$\sum_{k=1}^{n} f_k \left(\frac{X_k^2}{\langle X\rangle_k}\right) = o\left(\left(\log\langle X\rangle_{n+1}\right)^{1+\gamma}\right) \quad c.s.$$

Es decir

$$\frac{X_n^2}{\langle X\rangle_n (\log\langle X\rangle_n)^{1+\gamma}} \to 0 \quad c.s.$$

$$\sum_{k=1}^{n} f_k \left(\frac{X_k^2}{\langle X\rangle_k \left(\log\langle X\rangle_{n+1}\right)^{1+\gamma}}\right) \to 0 \quad c.s.$$

Demostración. Sean

$$A_n = \frac{X_n^2}{\langle X\rangle_n}, B_n = \frac{\Delta\langle X\rangle_{n+1}}{\langle X\rangle_{n+1}} = f_n, \quad \text{y} \quad C_n = f_n A_n$$

Luego,

$$\begin{aligned} E\left(A_{n+1}|\mathcal{F}_n\right) &= E\left(\frac{X_{n+1}^2}{\langle X\rangle_{n+1}}|\mathcal{F}_n\right) \\ &= \frac{E\left(X_{n+1}^2|\mathcal{F}_n\right)}{\langle X\rangle_{n+1}} \\ &= \frac{X_n^2 + \Delta\langle X\rangle_{n+1}}{\langle X\rangle_{n+1}} \\ &= \frac{X_n^2}{\langle X\rangle_{n+1}} + f_n \quad \text{c.s.} \end{aligned}$$

y por otro lado, tenemos que

$$\begin{aligned}
A_n + B_n - C_n &= A_n - f_n A_n + B_n \\
&= A_n (1 - f_n) + f_n \\
&= A_n \left(1 - \left(\frac{\langle X\rangle_{n+1} - \langle X\rangle_n}{\langle X\rangle_{n+1}}\right)\right) + f_n \\
&= A_n \left(\frac{\langle X\rangle_n}{\langle X\rangle_{n+1}}\right) + f_n \\
&= \frac{X_n^2}{\langle X\rangle_n}\left(\frac{\langle X\rangle_n}{\langle X\rangle_{n+1}}\right) + f_n \\
&= \frac{X_n^2}{\langle X\rangle_{n+1}} + f_n \quad \text{c.s.}
\end{aligned}$$

Por lo tanto $E(A_{n+1}|\mathcal{F}_n) = A_n + B_n - C_n$ c.s. y en particular

$$E(A_{n+1}|\mathcal{F}_n) \leq A_n + B_n - C_n \quad \text{c.s.}$$

Sin embargo si $a_n = \left(\log \langle X\rangle_{n+1}\right)^{1+\gamma}$ con $\gamma > 0$, la sucesión $\{a_n, n \in \mathbb{N}\}$ es estrictamente creciente y tiende a infinito c.s. entonces para n suficientemente grande, $a_n \geq 2$.

Luego,

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{B_n}{a_n} &= \sum_{n=0}^{\infty} \frac{\Delta\langle X\rangle_{n+1}}{\langle X\rangle_{n+1} a_n} \\
&= \sum_{n=0}^{\infty} \frac{\Delta\langle X\rangle_{n+1}}{\langle X\rangle_{n+1}\left(\log \langle X\rangle_{n+1}\right)^{1+\gamma}} \\
&\leq \alpha + \int_2^{\infty} \frac{dx}{x(\log x)^{1+\gamma}} \quad \text{con } \alpha > 0 \\
&< \infty \quad \text{c.s.}
\end{aligned}$$

Por el Corolario 2 resulta entonces que $A_n = o(a_{n-1})$ c.s. y

$$\sum_{n=0}^{\infty} \frac{C_n}{a_n} < \infty \ \text{ c.s.}$$

Hasta aqui tenemos que $\{a_n, n \in \mathbb{N}\}$ una sucesión de variables aleatorias positiva y estrictamente creciente hacia infinito y $\{C_n, n \in \mathbb{N}\}$ una sucesión de variables aleatorias tal que $\sum_{k=1}^{n} \frac{C_k}{a_k}$ converge c.s. Entonces, por el Lema de Kronecker (ver [1]), tenemos que

$$\sum_{k=1}^{n} \frac{C_k}{a_n} \to 0 \quad \text{c.s.}$$

es decir

$$\sum_{k=1}^{n} \frac{f_k A_k}{a_n} \to 0 \quad \text{c.s.}$$

y por lo tanto

$$\sum_{k=1}^{n} f_k \left(\frac{X_k^2}{\langle X \rangle_k \left(\log \langle X \rangle_{n+1}\right)^{1+\gamma}} \right) \to 0 \quad \text{c.s.}$$

□

- **Teorema del límite central**

El teorema del límite central para las variables aleatorias independientes y de cuadrado integrable se generaliza también a un teorema de límite central para martingalas.

Teorema 7. *Sea* $\{X_n, n \in \mathbb{N}\}$ *una martingala de cuadrado integrable y sea* $\{\langle X \rangle_n, n \in \mathbb{N}\}$ *su proceso creciente. Sea* $\{a_n, n \in \mathbb{N}\}$ *una sucesión determinista, positiva, creciente hacia infinito. Supongamos que:*

1. *Existe un límite determinista* $\ell \geq 0$ *tal que*

$$\frac{\langle X \rangle_n}{a_n} \xrightarrow{\mathbb{P}} \ell$$

2. *La condición de Lindeberg se satisface, es decir, para todo* $\epsilon > 0$

$$\frac{1}{a_n} \sum_{k=1}^{n} E\left[|\Delta X_k|^2 I_{(|\Delta X_k| \geq \epsilon\sqrt{a_n})} |\mathcal{F}_{k-1} \right] \xrightarrow{\mathbb{P}} 0 \quad \textit{cuando} \quad n \to \infty$$

Entonces, tenemos

$$\frac{1}{\sqrt{a_n}} X_n \xrightarrow{d} \mathcal{N}(0, \ell) \quad \textit{cuando} \quad n \to \infty$$

Además, si $\ell > 0$, tenemos

$$\sqrt{a_n} \frac{X_n}{\langle X \rangle_n} \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{\ell}\right) \quad \textit{cuando} \quad n \to \infty.$$

Nota 5. *Si existe $\alpha > 2$ tal que*

$$\frac{1}{a_n} \sum_{k=1}^{n} E\left[|\Delta X_k|^{\alpha} | \mathcal{F}_{k-1}\right] \xrightarrow{\mathbb{P}} 0 \quad \textit{cuando} \quad n \to \infty.$$

Con el fin de proponer una prueba simple del TLCM, se colocará en el marco restrictivo, donde para todo $n \geq 0$

$$E\left[|\Delta X_{n+1}|^3 | \mathcal{F}_n\right] \leq 1 \quad \text{c.s.}$$

con $a_n = n$. Esta limitación de momento permite evitar un argumento de truncamiento suficientemente fuerte sobre la martingala $\{X_n, n \in \mathbb{N}\}$.

Demostración. Por el Teorema de Lévy y pasando a la función característica, es suficiente mostrar que, para todo $t \in \mathbb{R}$,

$$\lim_{n\to\infty} E\left[e^{\left(\frac{itX_n}{\sqrt{n}}\right)}\right] = e^{\left(-\frac{t^2 \ell}{2}\right)}$$

Sea

$$L_n(t) = e^{\left(itX_n + \frac{t^2}{2}\langle X \rangle_n\right)} \quad \text{para todo} \quad t \in \mathbb{R}$$

Vamos primero a demostrar que la esperanza de $L_n(t)$ converge uniformemente a 1 en t con la condición de que $|t| = O(\sqrt{n})$.

Sea

$$\Phi(t) = e^{\left(itx + \frac{t^2 y}{2}\right)} \quad \text{para todo} \quad x \in \mathbb{R} \quad \text{y} \quad y \in [0,1]$$

Entonces

$$\begin{aligned}
\Phi^{'}(t) &= (ix+ty)\Phi(t)\\
\Phi^{2}(t) &= y\Phi(t)+(ix+ty)^{2}\Phi(t)\\
\Phi^{3}(t) &= y(ix+ty)\Phi(t)+2(ix+ty)(y)\Phi(t)+(ix+ty)^{3}\Phi(t)\\
&= (ix+ty)(3y+(ix+ty)^{2})\Phi(t)
\end{aligned}$$

Luego, por la formula de Taylor con resto integral, tenemos que

$$\begin{aligned}
\Phi(t) &= \Phi(0)+t\Phi^{'}(0)+\frac{t^2}{2}\Phi^{2}(0)+\frac{1}{2}\int_0^t \Phi^{3}(s)(t-s)^2 ds\\
&= 1+t(ix)+\frac{t^2}{2}(y+i^2x^2)+R(t,x,y)\\
&= 1+itx+\frac{t^2}{2}(y-x^2)+R(t,x,y)
\end{aligned}$$

Se puede acotar superiormente $R(t,x,y)$ porque
$\Phi^3(t)=(ix+ty)(3y+(ix+ty)^2)\Phi(t)$. Si $|t|\leq 1$, tenemos

$$\begin{aligned}
|\Phi^{3}(t)| &= |(ix+ty)(3y+(ix+ty)^{2})e^{(itx)}e^{\frac{t^2y}{2}}|\\
&= |(ix+ty)(3y-x^2+2itxy+t^2y^2)|e^{\frac{t^2y}{2}}\\
&= |3ixy-ix^3-2tx^2y+it^2xy^2+3ty^2-tx^2y+2it^2xy^2+t^3y^3|e^{\frac{t^2y}{2}}\\
&= |-2tx^2y+3ty^2-tx^2y+t^3y^3+i(3xy-x^3+t^2xy^2+2t^2xy^2)|e^{\frac{t^2y}{2}}\\
&\leq (|-2tx^2y+3ty^2-tx^2y+t^3y^3|+|3xy-x^3+t^2xy^2+2t^2xy^2|)e^{\frac{t^2y}{2}}\\
&\leq (|2tx^2y|+|3ty^2|+|tx^2y|+|t^3y^3|+|3xy|+|x^3|+|t^2xy^2|+|2t^2xy^2|)e^{\frac{t^2y}{2}}\\
&\leq (2x^2+3+x^2+1+3|x|+|x|^3+|x|+2|x|)e^{\frac{t^2y}{2}} \quad \text{porque } y\in[0,1]\\
&= (|x|^3+3x^2+6|x|+4)e^{\frac{t^2y}{2}}\\
&= (|x|^3+3x^2+3|x|+1+3|x|+3)e^{\frac{t^2y}{2}}\\
&= ((|x|+1)^3+3(|x|+3))e^{\frac{t^2y}{2}}\\
&\leq ((|x|+1)^3+3(|x|+1)^3)e^{\frac{t^2y}{2}}\\
&= (4(|x|+1)^3)e^{\frac{t^2y}{2}}
\end{aligned}$$

Luego, si $t \geq 0$

$$
\begin{aligned}
|R(t,x,y)| &= \frac{1}{2}\left|\int_0^t \Phi^3(s)(t-s)^2 ds\right| \\
&\leq \left|\int_0^t \Phi^3(s)(t-s)^2 ds\right| \\
&\leq \int_0^t \left|\Phi^3(s)(t-s)^2\right| ds \\
&= \int_0^t |\Phi^3(s)|(t-s)^2 ds \\
&\leq \int_0^t 4(|x|+1)^3 e^{\left(\frac{s^2}{2}\right)}(t-s)^2 ds \\
&\leq \int_0^t 4(|x|+1)^3 e^{\left(\frac{t^2}{2}\right)}(t-s)^2 ds \\
&= 4(|x|+1)^3 e^{\left(\frac{t^2}{2}\right)} \int_0^t (t^2-2ts+s^2) ds \\
&= 4(|x|+1)^3 e^{\left(\frac{t^2}{2}\right)}(t^3-t^3+\frac{t^3}{3}) \\
&= \frac{4}{3}(|x|+1)^3 t^3 e^{\left(\frac{t^2}{2}\right)}
\end{aligned}
$$

y si $t < 0$

$$
\begin{aligned}
|R(t,x,y)| &= \frac{1}{2}\left|\int_0^t \Phi^3(s)(t-s)^2 ds\right| \\
&\leq \left|\int_0^t \Phi^3(s)(t-s)^2 ds\right| \\
&= \left|-\int_t^0 \Phi^3(s)(t-s)^2 ds\right| \\
&\leq \int_t^0 \left|\Phi^3(s)(t-s)^2\right| ds \\
&= \int_t^0 |\Phi^3(s)|(t-s)^2 ds
\end{aligned}
$$

$$\begin{aligned}
&\leq \int_t^0 4(|x|+1)^3 e^{\left(\frac{s^2}{2}\right)}(t-s)^2 ds \\
&\leq \int_t^0 4(|x|+1)^3 e^{\left(\frac{t^2}{2}\right)}(t-s)^2 ds \\
&= 4(|x|+1)^3 e^{\left(\frac{t^2}{2}\right)} \int_t^0 (t^2 - 2ts + s^2) ds \\
&= 4(|x|+1)^3 e^{\left(\frac{t^2}{2}\right)}(-t^3 + t^3 - \frac{t^3}{3}) \\
&= -\frac{4}{3}(|x|+1)^3 t^3 e^{\left(\frac{t^2}{2}\right)}
\end{aligned}$$

Por lo tanto $R(t,x,y)$ está acotada superiormente, es decir

$$|R(t,x,y)| \leq \frac{4}{3}(|x|+1)^3 |t|^3 e^{\left(\frac{t^2}{2}\right)}$$

Ahora, sea

$$Z_n(t) = e^{(it\Delta X_n + \frac{t^2}{2}\Delta\langle X\rangle_n)} \quad \text{para todo} \quad n \geq 0 \quad \text{y} \quad t \in \mathbb{R}$$

Sabemos que $\Delta\langle X\rangle_{n+1} = E\left(\Delta X_{n+1}{}^2|\mathcal{F}_n\right)$, entonces usando la desigualdad de Jensen, tenemos que $E\left(\Delta X_{n+1}{}^2|\mathcal{F}_n\right)^{\frac{3}{2}} \leq E\left(\Delta X_{n+1}{}^3|\mathcal{F}_n\right) \leq 1$ c.s., entonces $E\left(\Delta X_{n+1}{}^2|\mathcal{F}_n\right) \leq 1$ c.s.

Si tomamos $x = \Delta X_{n+1}$ y $y = \Delta\langle X\rangle_{n+1}$, es claro que $y \in [0,1]$ y

$$\begin{aligned}
E(Z_{n+1}(t)|\mathcal{F}_n) &= E[e^{(it\Delta X_{n+1} + \frac{t^2}{2}\Delta\langle X\rangle_{n+1})}|\mathcal{F}_n] \\
&= E(1 + it\Delta X_{n+1} + \frac{t^2}{2}(\Delta\langle X\rangle_{n+1} - \Delta X_{n+1}{}^2)|\mathcal{F}_n) + r_n(t)
\end{aligned}$$

con $r_n(t) = E[R(t, \Delta X_{n+1}, \Delta\langle X\rangle_{n+1})|\mathcal{F}_n)]$.

Todo se simplifica porque

$$E(\Delta X_{n+1}|\mathcal{F}_n) = E[X_{n+1} - X_n|\mathcal{F}_n] = E[X_{n+1}|\mathcal{F}_n] - X_n = X_n - X_n = 0$$

y

$$\begin{aligned}
E[\Delta\langle X\rangle_{n+1} - \Delta X_{n+1}{}^2|\mathcal{F}_n] &= E[\Delta\langle X\rangle_{n+1}|\mathcal{F}_n] - E[\Delta X_{n+1}{}^2|\mathcal{F}_n] \\
&= E[E\left(\Delta X_{n+1}{}^2|\mathcal{F}_n\right)|\mathcal{F}_n] - E[\Delta X_{n+1}{}^2|\mathcal{F}_n] \\
&= E[\Delta X_{n+1}{}^2|\mathcal{F}_n] - E[\Delta X_{n+1}{}^2|\mathcal{F}_n] \\
&= 0
\end{aligned}$$

Así,

$$E(Z_{n+1}(t)|\mathcal{F}_n) = 1 + r_n(t)$$

Sin embargo, tenemos

$$\begin{aligned}
E[(1+|\Delta X_{n+1}|)^3|\mathcal{F}_n] &= E[1 + 3|\Delta X_{n+1}| + 3(\Delta X_{n+1})^2 + (|\Delta X_{n+1}|)^3|\mathcal{F}_n] \\
&= 1 + 3E[|\Delta X_{n+1}||\mathcal{F}_n] + 3E[(\Delta X_{n+1})^2|\mathcal{F}_n] \\
&\quad + E[(|\Delta X_{n+1}|)^3|\mathcal{F}_n] \\
&\leq 1 + 3E[|\Delta X_{n+1}||\mathcal{F}_n] + 3 + 1 \\
&\leq 5 + 3E[(|\Delta X_{n+1}|)^3|\mathcal{F}_n] \\
&\leq 8
\end{aligned}$$

entonces

$$(1+|\Delta X_{n+1}|)^3 \leq 8$$

Señalamos entonces que

$$|r_n(t)| \leq \frac{32|t|^3}{3}e^{\frac{t^2}{2}} \quad \text{para todo} \quad n \geq 0 \quad \text{y} \quad |t| \leq 1$$

Como buscamos obtener un resultado asintótico sobre $\{X_n, n \in \mathbb{N}\}$, podemos suponer que $X_0 = 0$. Tenemos entonces

$$\begin{aligned}
L_n(t)Z_{n+1}(t) &= e^{\left(itX_n + \frac{t^2}{2}\langle X\rangle_n\right)} e^{\left(it\Delta X_{n+1} + \frac{t^2}{2}\Delta\langle X\rangle_{n+1}\right)} \\
&= e^{\left(itX_n + \frac{t^2}{2}\langle X\rangle_n + itX_{n+1} - itX_n + \frac{t^2}{2}\langle X\rangle_{n+1} - \frac{t^2}{2}\langle X\rangle_n\right)} \\
&= e^{\left(itX_{n+1} + \frac{t^2}{2}\langle X\rangle_{n+1}\right)} \\
&= L_{n+1}(t)
\end{aligned}$$

Por lo tanto

$$\begin{aligned} E[L_{n+1}(t)|\mathcal{F}_n] &= E[L_n(t)Z_{n+1}(t)|\mathcal{F}_n] \\ &= L_n(t)E[Z_{n+1}(t)|\mathcal{F}_n] \\ &= L_n(t)(1+r_n(t)) \\ &= L_n(t)+L_n(t)r_n(t) \end{aligned}$$

Por otra parte, vimos que $\langle X\rangle_n \leq n$ de donde

$$|L_n(t)| = |e^{(itX_n)}e^{\left(\frac{t^2}{2}\langle X\rangle_n\right)}| = |e^{\left(\frac{t^2}{2}\langle X\rangle_n\right)}| \leq e^{\left(\frac{nt^2}{2}\right)}$$

Además tenemos que

$$\begin{aligned} E[L_n(t)] &= E[E[L_n(t)|\mathcal{F}_{n-1}]] \\ &= E[L_{n-1}(t)+L_{n-1}(t)r_{n-1}(t)] \\ &= E[L_{n-1}(t)]+E[L_{n-1}(t)r_{n-1}(t)] \end{aligned}$$

Entonces,

$$\begin{aligned} |E[L_n(t)]-1| &= |E[L_{n-1}(t)]+E[L_{n-1}(t)r_{n-1}(t)]-1| \\ &= |E[L_{n-2}(t)]+E[L_{n-2}(t)r_{n-2}(t)]+E[L_{n-1}(t)r_{n-1}(t)]-1| \\ &\vdots \\ &= |1+E[r_0(t)]+E[L_1(t)r_1(t)]+E[L_2(t)r_2(t)]+... \\ &\quad +E[L_{n-2}(t)r_{n-2}(t)]+E[L_{n-1}(t)r_{n-1}(t)]-1| \\ &\leq |E[r_0(t)]|+|E[L_1(t)r_1(t)]|+|E[L_2(t)r_2(t)]|+... \\ &\quad +|E[L_{n-2}(t)r_{n-2}(t)]|+|E[L_{n-1}(t)r_{n-1}(t)]| \\ &\leq \frac{32|t|^3}{3}e^{\frac{t^2}{2}}+\frac{32|t|^3}{3}e^{\frac{t^2}{2}}e^{\frac{t^2}{2}}+\frac{32|t|^3}{3}e^{\frac{t^2}{2}}e^{\frac{2t^2}{2}}+... \\ &\quad +\frac{32|t|^3}{3}e^{\frac{t^2}{2}}e^{\frac{(n-2)t^2}{2}}+\frac{32|t|^3}{3}e^{\frac{t^2}{2}}e^{\frac{(n-1)t^2}{2}} \\ &= \frac{32|t|^3}{3}[e^{\frac{t^2}{2}}+e^{\frac{2t^2}{2}}+e^{\frac{3t^2}{2}}+... \\ &\quad +e^{\frac{(n-1)t^2}{2}}+e^{\frac{(n)t^2}{2}}] \end{aligned}$$

$$\begin{aligned}
&\leq \frac{32|t|^3}{3}[e^{\frac{nt^2}{2}} + e^{\frac{nt^2}{2}} + e^{\frac{nt^2}{2}} + ... \\
&\quad + e^{\frac{nt^2}{2}} + e^{\frac{nt^2}{2}}] \\
&= \frac{32n|t|^3}{3}(e^{\frac{nt^2}{2}})
\end{aligned}$$

Si cambiamos t por $\frac{t}{\sqrt{n}}$, obtenemos que

$$\begin{aligned}
\lim_{n\to\infty}\left|E\left[L_n\left(\frac{t}{\sqrt{n}}\right)\right]-1\right| &\leq \lim_{n\to\infty}\left(\frac{32n|t|^3}{3n^{3/2}}(e^{\frac{nt^2}{2n}})\right) \\
&= \lim_{n\to\infty}\left(\frac{32|t|^3}{3n^{1/2}}(e^{\frac{t^2}{2}})\right) \\
&= 0
\end{aligned}$$

Entonces

$$\lim_{n\to\infty} E\left[L_n\left(\frac{t}{\sqrt{n}}\right)\right] = \lim_{n\to\infty} e^{\left(\frac{itX_n}{\sqrt{n}}+\frac{t^2}{2n}\langle X\rangle_n\right)} = 1$$

Luego, sea

$$\Lambda_n(t) = e^{\left(\frac{t^2}{2n}\langle X\rangle_n\right)} \quad \text{para todo} \quad n \geq 1$$

Es claro que la sucesión $\{\Lambda_n, n \in \mathbb{N}\}$ está acotada, por lo tanto es uniformemente integrable porque $\langle X\rangle_n \leq n$. Además converge en probabilidad a $e^{\frac{t^2\ell}{2}}$ entonces converge en L^1 a $e^{\frac{t^2\ell}{2}}$, es decir

$$\lim_{n\to\infty} E\left[\left|\Lambda_n(t) - e^{\frac{t^2\ell}{2}}\right|\right] = 0 \tag{2.4}$$

Finalmente,

$$\begin{aligned}
\left|\Lambda_n(t) - e^{\left(\frac{t^2\ell}{2}\right)}\right| &= \left|e^{\left(\frac{t^2}{2n}\langle X\rangle_n\right)} - e^{\left(\frac{t^2\ell}{2}\right)}\right| \\
&= \left|e^{\left(\frac{itX_n}{\sqrt{n}}\right)}\right|\left|e^{\left(\frac{t^2}{2n}\langle X\rangle_n\right)} - e^{\left(\frac{t^2\ell}{2}\right)}\right| \\
&= \left|e^{\left(\frac{itX_n}{\sqrt{n}}\right)}e^{\left(\frac{t^2}{2n}\langle X\rangle_n\right)} - e^{\left(\frac{itX_n}{\sqrt{n}}\right)}e^{\left(\frac{t^2\ell}{2}\right)}\right| \\
&= \left|L_n\left(\frac{t}{\sqrt{n}}\right) - e^{\left(\frac{itX_n}{\sqrt{n}}+\frac{t^2\ell}{2}\right)}\right|
\end{aligned}$$

y por la ecuación (2.4) y como $E\left[L_n\left(\frac{t}{\sqrt{n}}\right)\right] \to 1$ resulta por lo tanto que

$$\lim_{n\to\infty} E\left[e^{\left(\frac{itX_n}{\sqrt{n}}+\frac{t^2\ell}{2}\right)}\right] = 1.$$

□

2.3. Procesos auto-recursivos

Los procesos auto-recursivos desempeñan un papel central en la modelación de numerosos fenómenos concretos.

Consideremos el proceso auto-recursivo de orden uno

$$M_{n+1} = \theta M_n + \varepsilon_{n+1}$$

donde el estado inicial M_0 es elegido arbitrariamente y la sucesión $\{\varepsilon_n, n \in \mathbb{N}\}$ es i.i.d., centrada ($E[\varepsilon_{n+1}|\mathcal{F}_n] = 0$) y de varianza $\sigma^2 > 0$ ($E[{\varepsilon_{n+1}}^2|\mathcal{F}_n] = \sigma^2$). Supongamos además que M_0 y $\{\varepsilon_n, n \in \mathbb{N}\}$ son independientes. Este proceso constituye una de las más simples series de tiempo. Su estudio teórico depende de la tricotomía $|\theta| < 1$, $|\theta| = 1$ y $|\theta| > 1$. El proceso será **estable** si $|\theta| < 1$, **inestable** si $|\theta| = 1$ y **explosivo** si $|\theta| > 1$. En los tres casos, estimamos los parámetros desconocidos θ y σ^2 al tiempo n mediante

$$\widehat{\theta}_n = \frac{\sum_{k=1}^{n} M_k M_{k-1}}{\sum_{k=0}^{n-1} M_k^2} \tag{2.5}$$

y

$$\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{k=1}^{n}(M_k - \widehat{\theta}_{k-1}M_{k-1})^2 \tag{2.6}$$

Consideremos la siguiente sucesión de variables aleatorias

$$X_n = \sum_{k=1}^{n} M_{k-1}\varepsilon_k$$

Observemos que X_n es una martingala ya que

$$\begin{aligned}
E[X_{n+1}|\mathcal{F}_n] &= E[X_n + M_n\varepsilon_{n+1}|\mathcal{F}_n] \\
&= X_n + M_n E[\varepsilon_{n+1}|\mathcal{F}_n] \\
&= X_n \quad \text{c.s.}
\end{aligned}$$

con proceso creciente

$$\begin{aligned}
\langle X\rangle_n &= \sum_{k=1}^{n} E[(X_k - X_{k-1})^2|\mathcal{F}_{k-1}] \\
&= \sum_{k=1}^{n} E\left[\left(\sum_{j=1}^{k} M_{j-1}\varepsilon_j - \sum_{j=1}^{k-1} M_{j-1}\varepsilon_j\right)^2 |\mathcal{F}_{k-1}\right] \\
&= \sum_{k=1}^{n} E\left[M_{k-1}^2\varepsilon_k^2|\mathcal{F}_{k-1}\right] \\
&= \sum_{k=1}^{n} M_{k-1}^2 E\left[\varepsilon_k^2|\mathcal{F}_{k-1}\right] \\
&= \sigma^2\sum_{k=1}^{n} M_{k-1}^2
\end{aligned}$$

Luego, obtenemos la siguiente descomposición

$$\begin{aligned}
\sigma^2 \frac{X_n}{\langle X \rangle_n} &= \frac{\sum_{k=1}^n M_{k-1}\varepsilon_k}{\sum_{k=1}^n M_{k-1}^2} \\
&= \frac{\sum_{k=1}^n M_{k-1}(M_k - \theta M_{k-1})}{\sum_{k=1}^n M_{k-1}^2} \\
&= \frac{\sum_{k=1}^n M_{k-1}M_k - \theta \sum_{k=1}^n M_{k-1}^2}{\sum_{k=1}^n M_{k-1}^2} \\
&= \frac{\sum_{k=1}^n M_{k-1}M_k}{\sum_{k=0}^{n-1} M_k^2} - \theta \\
&= \widehat{\theta}_n - \theta
\end{aligned}$$

En el caso estable, tenemos:

$$\begin{aligned}
M_1 &= \theta M_0 + \varepsilon_1 \\
M_2 &= \theta M_1 + \varepsilon_2 = \theta^2 M_0 + \theta\varepsilon_1 + \varepsilon_2 \\
M_3 &= \theta M_2 + \varepsilon_3 = \theta^3 M_0 + \theta^2\varepsilon_1 + \theta\varepsilon_2 + \varepsilon_3 \\
&\vdots \\
M_n &= \theta^n M_0 + \theta^{n-1}\varepsilon_1 + \theta^{n-2}\varepsilon_2 + \ldots + \theta\varepsilon_{n-1} + \varepsilon_n \\
&= \theta^n M_0 + \sum_{k=1}^n \theta^{n-k}\varepsilon_k
\end{aligned}$$

Supongamos para simplificar que $M_0 = 0$. Tenemos, por la desigualdades de Cauchy-Schwarz

$$M_n^2 \leq \frac{1}{1-|\theta|} \sum_{k=1}^n |\theta|^{n-k} \varepsilon_k^2$$

Entonces

$$\begin{aligned}
\sum_{k=1}^{n} M_k^2 &\leq \sum_{k=1}^{n} \frac{1}{1-|\theta|} \sum_{j=1}^{k} |\theta|^{k-j} \varepsilon_j^2 \\
&= \frac{1}{1-|\theta|} \sum_{k=1}^{n} \sum_{j=1}^{k} |\theta|^{k-j} \varepsilon_j^2 \\
&= \frac{1}{1-|\theta|} \sum_{k=1}^{n} \sum_{j=1}^{k} |\theta|^{k-j} \varepsilon_j^2 \\
&= \frac{1}{1-|\theta|} \left[|\theta|^0 \sum_{k=1}^{n} \varepsilon_k^2 + |\theta|^1 \sum_{k=1}^{n-1} \varepsilon_k^2 + |\theta|^2 \sum_{k=1}^{n-2} \varepsilon_k^2 + \ldots + |\theta|^{n-1} \varepsilon_1^2 \right] \\
&\leq \frac{1}{1-|\theta|} \sum_{j=0}^{\infty} |\theta|^j \sum_{k=1}^{n} \varepsilon_k^2 \\
&= \left(\frac{1}{1-|\theta|} \right)^2 \sum_{k=1}^{n} \varepsilon_k^2
\end{aligned}$$

Sea

$$S_n = \sum_{k=0}^{n} M_k^2 \qquad \text{y} \qquad L_n = \sum_{k=1}^{n} \varepsilon_k^2$$

tenemos, por la desigualdad anterior, que $S_n = O(L_n)$, además $L_n = O(n)$, entonces $S_n \leq ML_n \leq kn$, así $S_n \leq kn$ a partir de cierto n, es decir $S_n = O(n)$ c.s., lo cual implica que $S_n \to \infty$ c.s. y también $\langle X \rangle_n \to \infty$ c.s., por consiguiente $\langle X \rangle_n = O(n)$ c.s., luego de la primera LGNM tenemos que $X_n = o(n)$ c.s.

Por otra parte, tenemos

$$
\begin{aligned}
S_n &= \sum_{k=0}^{n} M_k^2 \\
&= M_0^2 + \sum_{k=1}^{n} M_k^2 \\
&= \sum_{k=1}^{n} (\theta M_{k-1} + \varepsilon_k)^2 \\
&= \theta^2 \sum_{k=1}^{n} M_{k-1}^2 + 2\theta \sum_{k=1}^{n} M_{k-1}\varepsilon_k + \sum_{k=1}^{n} \varepsilon_k^2 \\
&= \theta^2 S_{n-1} + 2\theta X_n + L_n
\end{aligned}
$$

entonces

$$
\begin{aligned}
S_n - \theta^2 S_{n-1} &= 2\theta X_n + L_n \\
S_n(1-\theta^2) &= \theta^2 M_n^2 + 2\theta X_n + L_n
\end{aligned}
$$

Sea $V_n = \frac{S_n}{n}$ para todo $n \geq 1$, entonces

$$
V_n(1-\theta^2) = \frac{\theta^2 M_n^2}{n} + \frac{2\theta X_n}{n} + \frac{L_n}{n}
$$

Pero como $M_n^2 = o(n)$ y $X_n = o(n)$ c.s. deducimos que:

$$
V_n \to \ell \text{ c.s., donde } \ell = \frac{\sigma^2}{1-\theta^2}
$$

Entonces

$$\begin{aligned}
\lim_{n\to\infty}\frac{\langle X\rangle_n}{n} &= \lim_{n\to\infty}\frac{\sigma^2\sum_{k=1}^{n}M_{k-1}^2}{n}\\
&= \lim_{n\to\infty}\frac{\sigma^2\left(S_n - M_n^2\right)}{n}\\
&= \lim_{n\to\infty}\left(\sigma^2 V_n - \frac{\sigma^2 M_n^2}{n}\right)\\
&= \frac{\sigma^4}{1-\theta^2} \quad \text{c.s.}
\end{aligned}$$

Concluimos entonces, de la primera LGNM, que

$$\widehat{\theta}_n \to \theta \quad \text{c.s.} \tag{2.7}$$

Además la condición de Lindeberg es satisfecha, ver [1]. Por lo tanto tenemos por el TLCM que

$$\sqrt{n}(\widehat{\theta}_n - \theta) = \sqrt{n}\frac{X_n}{\langle X\rangle_n}\sigma^2 \xrightarrow{d} \mathcal{N}\left(0, 1-\theta^2\right) \quad \text{cuando} \quad n \to \infty \tag{2.8}$$

Finalmente tenemos que

$$\begin{aligned}
\lim_{n\to\infty}\widehat{\sigma}_n^2 &= \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}(M_k - \widehat{\theta}_{k-1}M_{k-1})^2\right]\\
&= \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\left(M_k - \left(\frac{\sigma^2 X_{k-1}}{\langle X\rangle_{k-1}} + \theta\right)M_{k-1}\right)^2\right]\\
&= \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\left(M_k - \theta M_{k-1} - \frac{\sigma^2 X_{k-1}M_{k-1}}{\langle X\rangle_{k-1}}\right)^2\right]\\
&= \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\left(\varepsilon_k - \frac{\sigma^2 X_{k-1}M_{k-1}}{\langle X\rangle_{k-1}}\right)^2\right]\\
&= \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2\right] - \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\frac{2\varepsilon_k\sigma^2 X_{k-1}M_{k-1}}{\langle X\rangle_{k-1}}\right]\\
&\quad + \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\frac{\sigma^4 X_{k-1}^2 M_{k-1}^2}{\langle X\rangle_{k-1}^2}\right]
\end{aligned}$$

Pero

$$\begin{aligned}
\lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\frac{\sigma^4 X_{k-1}^2 M_{k-1}^2}{\langle X\rangle_{k-1}^2}\right] &= \lim_{n\to\infty}\left[\frac{\sigma^2}{n}\sum_{k=1}^{n}\frac{\sigma^2 M_{k-1}^2}{\langle X\rangle_k}\frac{X_{k-1}^2}{\langle X\rangle_{k-1}}\frac{\langle X\rangle_k}{\langle X\rangle_{k-1}}\right] \\
&= \lim_{n\to\infty}\left[\frac{\sigma^2}{n}\sum_{k=1}^{n} f_{k-1}\frac{X_{k-1}^2}{\langle X\rangle_{k-1}}\frac{\langle X\rangle_k}{\langle X\rangle_{k-1}}\right] \\
&= \lim_{n\to\infty}\left[\frac{\sigma^2}{n}\sum_{k=1}^{n} f_{k-1}\frac{X_{k-1}^2}{\langle X\rangle_{k-1}}\right] \\
&\quad \left(\text{Porque } \frac{\langle X\rangle_k}{\langle X\rangle_{k-1}} = \frac{O(k)}{O(k-1)} \to 1\right) \\
&= \lim_{n\to\infty}\left[\sigma^2(0)\right] \\
&\quad (\text{Por la segunda LGNM y} \\
&\quad \text{porque } \frac{(\log n)^{1+\gamma}}{n} \to 0 \text{ para } \gamma > 0) \\
&= 0,
\end{aligned}$$

$$\begin{aligned}
\lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\frac{2\varepsilon_k\sigma^2 X_{k-1}M_{k-1}}{\langle X\rangle_{k-1}}\right] &\leq \lim_{n\to\infty}\left[2\left|\frac{1}{n}\sum_{k=1}^{n}\frac{\varepsilon_k\sigma^2 X_{k-1}M_{k-1}}{\langle X\rangle_{k-1}}\right|\right] \\
&\leq \lim_{n\to\infty}\left[2\sqrt{\frac{1}{n^2}}\sqrt{\sum_{k=1}^{n}\varepsilon_k^2}\sqrt{\sum_{k=1}^{n}\frac{\sigma^4 X_{k-1}^2 M_{k-1}^2}{\langle X\rangle_{k-1}^2}}\right] \\
&= \lim_{n\to\infty}\left[2\sqrt{\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2}\sqrt{\frac{1}{n}\sum_{k=1}^{n}\frac{\sigma^4 X_{k-1}^2 M_{k-1}^2}{\langle X\rangle_{k-1}^2}}\right] \\
&= \lim_{n\to\infty}\left[2(\sigma)(0)\right] \\
&= 0,
\end{aligned}$$

$$\text{y} \quad \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2\right] = \sigma^2$$

Por lo tanto

$$\widehat{\sigma}_n^2 \to \sigma^2 \quad \text{c.s.} \tag{2.9}$$

Capítulo 3

Aplicación del proceso auto-recursivo

3.1. Datos de los ciclones que han impactado a México

Ciclón tropical es un término meteorológico usado para referirse a un sistema de tormentas caracterizado por una circulación cerrada alrededor de un centro de baja presión y que produce fuertes vientos y abundante lluvia.

El término "tropical" se refiere tanto al origen geográfico de estos sistemas, que se forman casi exclusivamente en las regiones tropicales del planeta, como a su formación en masas de aire tropical de origen marino. El término "ciclón" se refiere a la naturaleza ciclónica de las tormentas, con una rotación en el sentido contrario al de las agujas del reloj en el hemisferio norte y en el sentido de las agujas del reloj en el hemisferio sur.

Oficialmente, la "entrada en tierra" se produce cuando el centro de una tormenta (el centro del ojo, no su extremo), alcanza tierra. Naturalmente, las condiciones de tormenta pueden sentirse en la costa y en el interior mucho antes de la llegada. En realidad, para una tormenta moviéndose hacia el interior, las áreas de entrada en tierra experimentan la mitad de la misma antes de la llegada del centro del ojo.

La evolución de un ciclón tropical puede llegar a desarrollar cuatro etapas:

Perturbación Tropical: Zona de inestabilidad atmosférica asociada a la existencia de un área de baja presión, la cual propicia la generación incipiente de vientos convergentes cuya organización eventual provoca el desarrollo de una depresión tropical.

Depresión Tropical: Los vientos se incrementan en la superficie, producto de la existencia de una zona de baja presión. Dichos vientos alcanzan una velocidad sostenida menor o igual a 62 kilómetros por hora.

Tormenta Tropical: El incremento continuo de los vientos provoca que éstos alcancen velocidades sostenidas entre los 63 y 118 km/h. Las nubes se distribuyen en forma de espiral. Cuando el ciclón alcanza esta intensidad se le asigna un nombre preestablecido por la Organización Meteorológica Mundial.

Huracán: Los vientos máximos sostenidos alcanzan o superan los 119 km/h. El área nubosa cubre una extensión entre los 500 y 900 km de diámetro, produciendo lluvias intensas. El ojo del huracán alcanza normalmente un diámetro que varía entre 24 y 40 km, sin embargo, puede llegar hasta cerca de 100 km. En esta etapa el ciclón se clasifica por medio de la escala Saffir-Simpson, como se indica en la tabla.

Categoría	Vientos Máximos (km/h)	Marea de Tormenta que Normalmente Ocaciona (m)	Características de los Posibles Daños Materiales e Inundaciones
1	118.1-154	1.2-1.5	Caída de pequeños árboles; algunas inundaciones en carreteras costeras en sus zonas más bajas.
2	154.1-178	1.6-2.5	Tejados, puertas y ventanas dañados; desprendimiento de árboles.
3	178.1-210	2.6-4.0	Grietas en pequeñas construcciones; inundaciones en terrenos bajos y planos.
4	210.1-250	4.1-5.5	Desprendimiento de techos en viviendas; erosiones importantes en playas, cauces de ríos y arroyos. Daños inminentes en los servicios de agua potable y saneamiento.
5	Mayor a 250	Mayor a 5.5	Daño muy severo y extenso en ventanas y puertas. Falla total de techos en muchas residencias y edificios industriales.

Los aspectos destructivos de los ciclones tropicales, que marcan su intensidad, se deben principalmente a cuatro aspectos: viento, oleaje, marea de tormenta y lluvia.

Las tormentas que alcanzan fuerza tropical reciben un nombre, para ayudar a la hora de formular demandas del seguro, ayudar a advertir a la gente de la llegada de una tormenta y además para indicar que se trata de fenómenos importantes que no deben ser ignorados. Estos nombres se toman de listas que varían de región a región y son renovadas. Cada año, los nombres de tormentas que hayan sido especialmente destructivas (si ha habido alguna) son “retirados” y se eligen nuevos nombres para ocupar su lugar.

Los efectos de un ciclón tropical también tienen un efecto positivo, traen consigo lluvias para las cosechas de temporada y proporcionan agua para el llenado de presas que permiten el abastecimiento de agua a las ciudades, el riego en zonas semiáridas y la generación de energía eléctrica. Una presa llena, puede funcionar hasta por más de un año.

En México el Servicio Meteorológico Nacional se encarga de captar, procesar e interpretar dichos datos con la finalidad de mantener informados a los centros de prevención y autoridades, así como a la población y sobre el surgimiento, trayectoria de desplazamiento, avance y otras características de los huracanes.

Entre los meses de mayo a noviembre, en México se presentan en promedio 23 ciclones tropicales con vientos mayores a 63 km/h. Del orden de 14 ciclones tropicales ocurren en el océano Pacífico y 9 en el Golfo de México y el mar Caribe. De ellos 4 inciden cada año sobre territorio nacional o se acercan a menos de 100 km, 2 desde el Pacífico y 2 desde el Atlántico.

Es importante conocer el número de ciclones tropicales que entran a México cada año, así como sus consecuencias y aplicar acciones para disminuir el riesgo por efectos de estos, tanto en los estados costeros como en algunos del interior pues, por ejemplo, se han presentado ciclones devastadores, como el caso de Gilbert, en el golfo de México en 1988, el cual provocó muertes principalmente en la ciudad de Monterrey (ciudad no costera del estado de Nuevo León) y pérdidas económicas considerables en la zona de Cancún, Q. Roo. En el primer caso, el río Santa Catarina sobrepasó su capacidad total, y en el segundo, el fuerte oleaje, más la acción de la marea de tormenta, removió la arena de las playas de Cancún. Otro caso importante fue en 1997 cuando apareció en el océano Pacífico el huracán Pauline, que provocó la muerte de varios cientos de personas en la costa de los estados de Oaxaca y Guerrero, resultando dañado principalmente el puerto de Acapulco, donde se produjeron flujos de escombros y de lodo, producto de las intensas lluvias que dejó a su paso el huracán sobre la zona montañosa cercana.

Es por ello que estamos interesados en estudiar sobre dicho tema. El estudio estadístico que se desarrolla en esta investigación trata sobre el número de ciclones que han impactado a México cada año y comprende 2 períodos, el primero de 1970-2010 y el segundo de 1990-2010. Los datos que se presentan en la siguiente figura fueron extraídos de la página del Servicio Meteorológico Nacional: http://smn.cna.gob.mx/

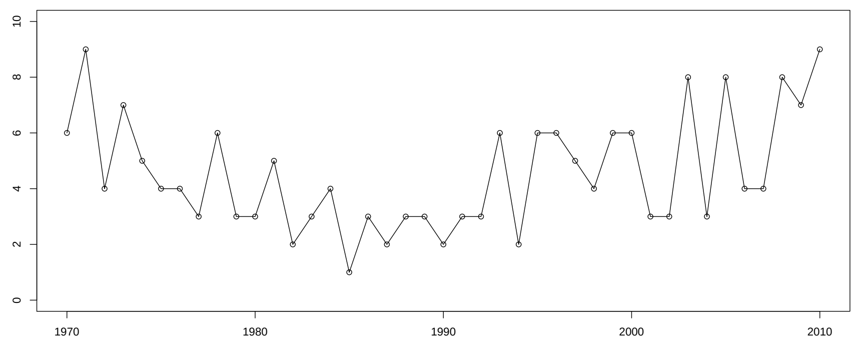

Figura 3.1: Número de ciclones que han impactado a México

Denotaremos a $M_i^{'}$ como el número de ciclones observados en el i-ésimo año a partir de 1970; es decir $i = 1, 2, ..., 41$ y a $M_i^{''}$ como el número de ciclones observados en el i-ésimo año a partir de 1990; es decir $i = 1, 2, ..., 21$. Por lo tanto, el tamaño de la muestra para $M_i^{'}$ es 41 y para $M_i^{''}$ es 21.

De las ecuaciones (2.5) y (2.6) obtenemos las siguientes expresiones recursivas para estimar los parámetros θ y σ^2 al tiempo $n = 41$ y $n = 21$.

$$\widehat{\theta}_i = \widehat{\theta}_{i-1} + \frac{1}{S_{i-1}} M_{i-1}(M_i - \widehat{\theta}_{i-1} M_{i-1}) \quad \text{y}$$

$$\widehat{\sigma}_i^2 = \left(1 - \frac{1}{i}\right)\left(\widehat{\sigma}_{i-1}^2\right) + \frac{1}{i}(M_i - \widehat{\theta}_i M_{i-1})^2 \tag{3.1}$$

donde $i \geq 2$, $\widehat{\theta}_1 = 0$ y $\widehat{\sigma}_1^2 = {M_1}^2$.

Cabe señalar que M_i toma los valores de $M_i^{'}$ y $M_i^{''}$ según sea el caso.

Hemos utilizado el software estadístico R para la estimación de los parámetros anteriores, ver el comportamiento de ε y predecir el número de ciclones en el año 2011 y 2012. Además de estimar los parámetros con datos artificiales y comprobrobar (2.8). Los algoritmos se encuentran en el Apéndice.

En la Figura 3.2 observamos las estimaciones de los parámetros θ y σ^2 cuando tenemos $n = 41$ datos (Algoritmo 1) y en la Figura 3.3 observamos las estimaciones de los parámetros cuando tenemos $n = 21$ datos (Algoritmo 2).

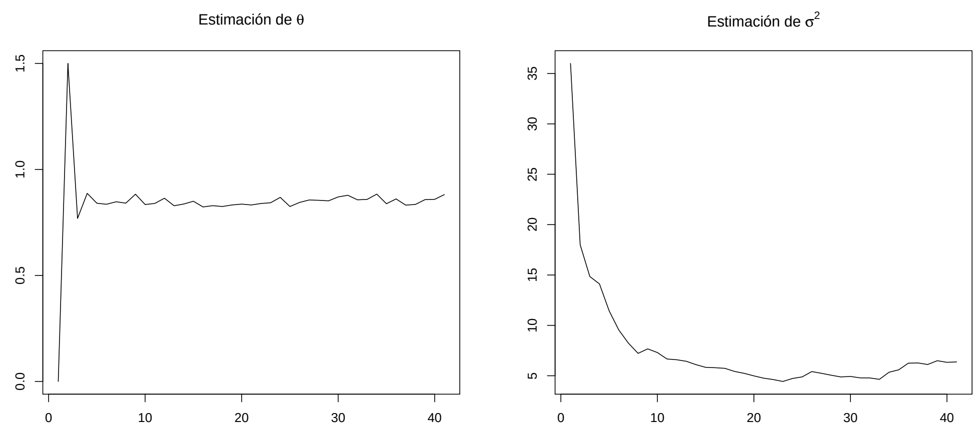

Figura 3.2: Estimación de los parámetros

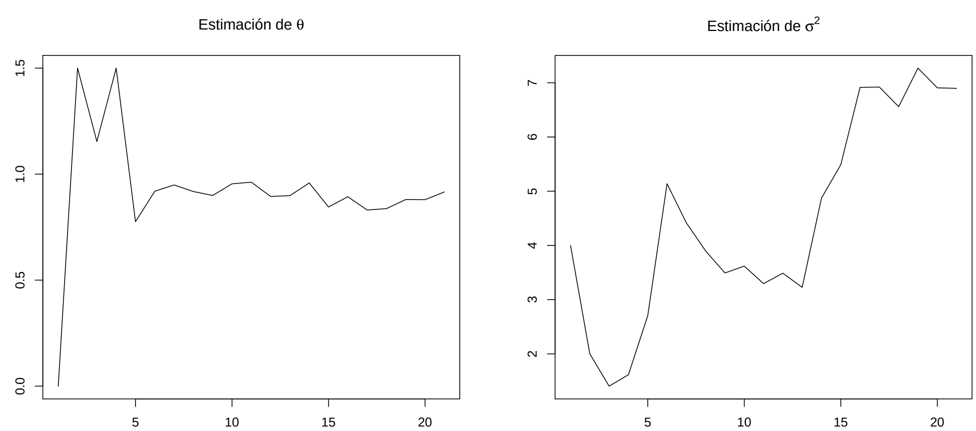

Figura 3.3: Estimación de los parámetros

En la siguiente tabla presentamos el último valor de los parámetros.

	$n = 41$	$n = 21$
$\widehat{\theta}_n$	0.8812834	0.915904
$\widehat{\sigma}_n^2$	6.373446	6.896962

Observamos que para ambos casos $|\widehat{\theta}_n| < 1$, por lo tanto el proceso es estable.

Recientemente, Bercu B. y Proia F. ver [6] consideraron que las variables aleatorias $\varepsilon_1, \varepsilon_2, ...$ puedan estar correlacionadas mediante un proceso autorecursivo de orden uno:

$$\varepsilon_{n+1} = \rho \varepsilon_n + V_{n+1}$$

en donde $|\rho| < 1$ y V_{n+1} es una sucesión de variables aleatorias i.i.d.,centrada y de varianza positiva. En este marco obtuvieron una prueba basada en la estadística de Durbin-Watson para contrastar las hipótesis

$H_0 : \rho = \rho_0$ contra $H_1 : \rho \neq \rho_0$

en donde ρ_0 es un valor fijo. Se consideró $\rho_0 = 0$ y se verificó que la sucesión ε_n es i.i.d.

Para el presente trabajo, primero estimamos los valores de los residuos en cada instante i, mediante la siguiente expresión:

$$\widehat{\varepsilon}_i = M_i - \widehat{\theta}_n M_{i-1}$$

donde $i \geq 2$ y $\widehat{\varepsilon}_1 = M_1$

Y utilizando la prueba de Bercu B. y Proia F. ya mencionada, obtenemos que para $n = 41$, rechazamos la hipótesis nula, es decir, $\widehat{\varepsilon}_1, \widehat{\varepsilon}_2, ..., \widehat{\varepsilon}_{41}$ no son independientes, sin embargo para $n = 21$, no rechazamos la hipótesis nula, así, tenemos independencia de $\widehat{\varepsilon}_1, \widehat{\varepsilon}_2, ..., \widehat{\varepsilon}_{21}$. Por lo tanto cosideraremos los datos del número de ciclones que impactaron a México desde 1990.

Luego, con la prueba Anderson-Darling no rechazamos la hipótesis nula que nos dice que los residuos siguen una Distribución Normal, con un nivel de significancia igual a 0.01, y observamos dicho comportamiento en la siguiente figura.

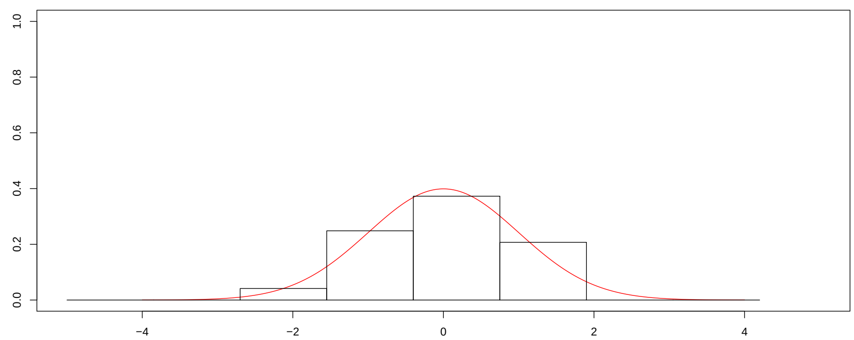

Figura 3.4: Distribución de los residuos

Además con la prueba t-Student no rechazamos nuestra hipótesis nula, con un nivel de significancia igual a 0.05, entonces tenemos que los residuos tienen media cero.

Podemos decir por lo tanto que $\varepsilon \sim N(0, \widehat{\sigma}^2_{21})$ (Algoritmo 2).

Ahora, gracias a que conocemos la distribución de ε, es posible predecir el número de ciclones en el año 2011 y 2012, mediante

$$\widehat{M_{22}} = \widehat{\theta}_{21} M_{21} + \varepsilon \quad \text{y} \quad \widehat{M_{23}} = \widehat{\theta}_{21} \widehat{M_{22}} + \varepsilon$$

donde $\varepsilon \sim N(0, \widehat{\sigma}^2_{21})$.

Así tenemos que el número de ciclones que impactaron a México en el año 2011 fueron 8 y nos damos cuenta, por el dato ya dado en el Servicio Meteorológico, que coincide con el número de ciclones que hubo en ese año. Basta comprobar que el número de ciclones que entren a México para el año 2012 sea

7 u 8, pero este dato sólo lo conoceremos al concluir dicho año (Algoritmo 2).

Ahora, consideraremos las primeras diferencias de nuestros 21 datos para tener un proceso estacionario (Algoritmo 3), como se muestra en la siguiente figura

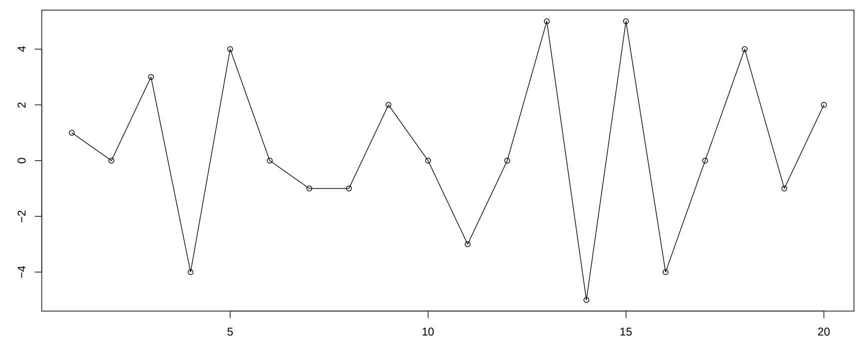

Figura 3.5: Primeras Diferencias

Luego en la siguiente figura observamos la estimación de los parámetros θ y σ^2, donde $\widehat{\theta}_{20} = -0.6363636$ y $\widehat{\sigma}^2_{20} = 4.629923$ (Algoritmo 3)

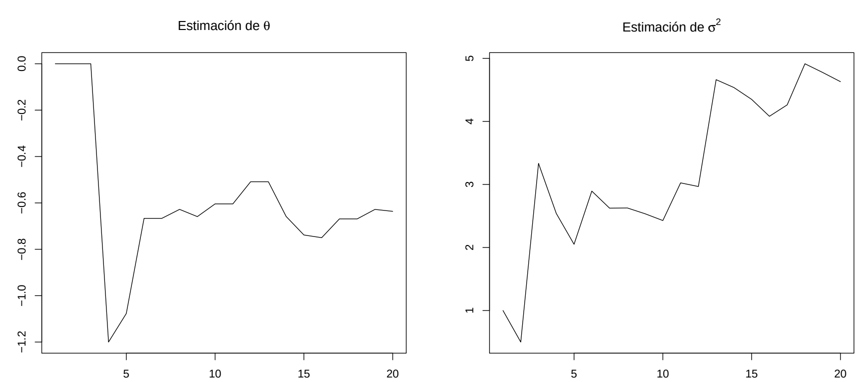

Figura 3.6: Estimación de los parámetros con Primeras Diferencias

y tenemos que $|\widehat{\theta}_{20}| = |-0.6363636| < 1$, así el proceso es estable.

Luego con la prueba de Bercu B. y Proia F., no rechazamos la hipótesis nula, es decir, obtenemos independencia de $\widehat{\varepsilon}_1, \widehat{\varepsilon}_2, ..., \widehat{\varepsilon}_{20}$ y con la prueba de Anderson-Darling tampoco rechazamos nuestra hipótesis nula y por lo tanto podemos decir que los residuos siguen una distribución Normal, observando dicho comportamiento en la siguiente figura

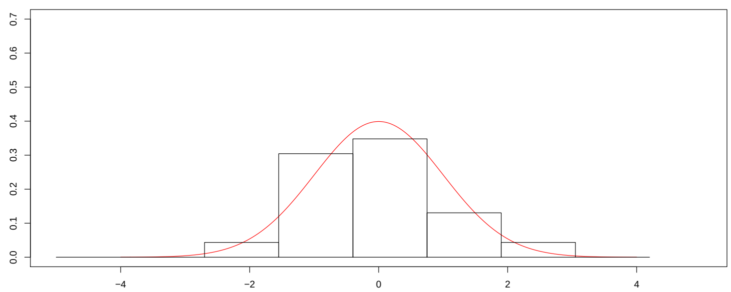

Figura 3.7: Distribución de los residuos con Primeras Diferencias

Además con la prueba t-Student no rechazamos nuestra hipótesis nula que nos dice que los residuos tienen media cero.

Podemos decir por lo tanto que $\varepsilon \sim N(0, \widehat{\sigma}^2_{20})$ (Algoritmo 3).

Ahora, nuestra predicción estará dada por

$$\widehat{M_{22}} = D_{21} + M_{21} \quad \text{y} \quad \widehat{M_{23}} = D_{22} + \widehat{M_{22}}$$

donde $\quad D_{21} = \widehat{\theta_{20}} D_{20} + \varepsilon, \quad D_{22} = \widehat{\theta_{20}} D_{21} + \varepsilon \quad$ y $\quad \varepsilon \sim N(0, \widehat{\sigma}^2_{20})$

y llegamos a que el número de ciclones que impactaron a México en el año 2011 fueron 8, con primeras diferencias, y coincide con el dato pronosticado anteriormente y con el dato del Servicio Meteorológico, sin embargo para el año 2012 pronosticamos 8 o 9 pero, como se habia comentado, este dato soló lo conoceremos al concluir dicho año (Algoritmo 3).

3.2. Datos artificiales

Nuestros datos artificiales estarán dados por el proceso auto-recursivo de orden uno

$$M_n = \theta M_{n-1} + \varepsilon_n$$

donde $\varepsilon_n \sim N(0, \sigma^2)$, al igual que M_1.

Consideremos $\theta = 0.3$ y $\sigma^2 = 2$, y obtenemos 30000 observaciones del proceso M_n.

De las ecuaciones recursivas (3.1) estimamos los parámetros θ y σ^2 (Algoritmo 4), observando una mejor estimación conforme n crece, como lo muestra la siguiente figura

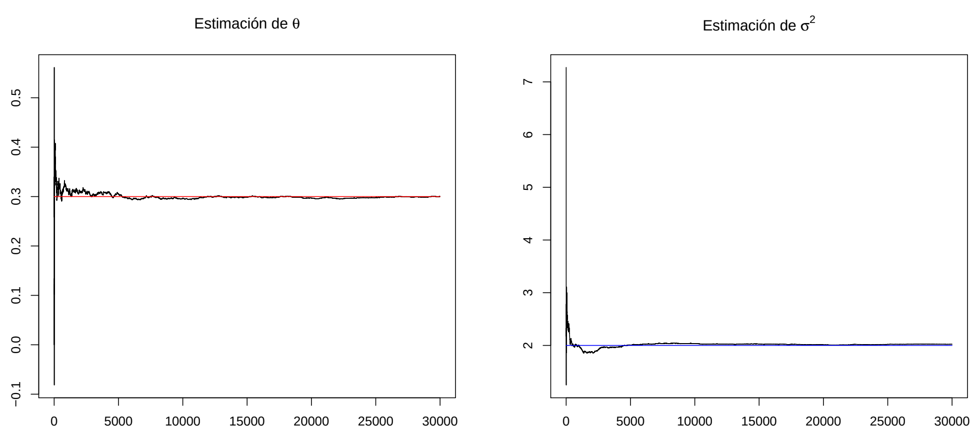

Figura 3.8: Estimación de los parámetros

Luego, obtenemos que $\widehat{\theta}_{30000} = 0.300324$ y $\widehat{\sigma}^2_{30000} = 2.023835$. Por lo tanto corroboramos las convergencias (2.7) y (2.9).

Finalmente, hacemos 1000 realizaciones de $\widehat{\theta}_{30000}$: $\widehat{\theta}^1_{30000}, \widehat{\theta}^2_{30000}, ..., \widehat{\theta}^{1000}_{30000}$ y definimos

$$Z_k = \frac{\sqrt{30000}[\widehat{\theta}^k_{30000} - .3]}{\sqrt{1-(.3)^2}} \quad \text{para } k = 1, 2, ..., 1000$$

Luego, representamos el siguiente histograma de $Z_1, Z_2, ..., Z_{1000}$

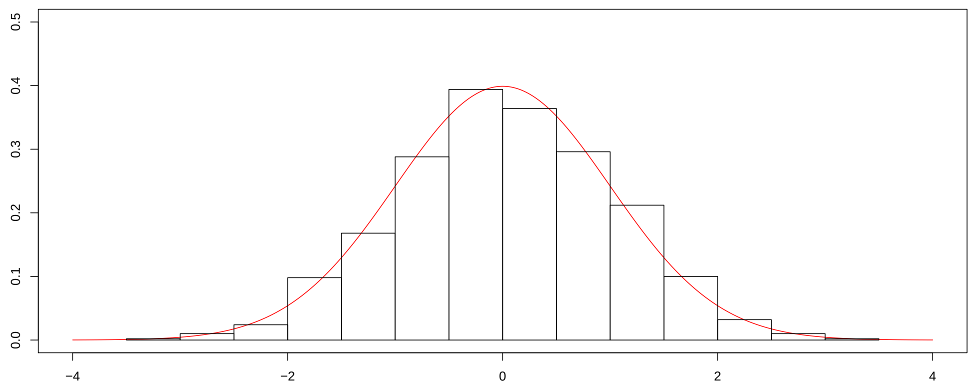

Figura 3.9: Distribución de Z

observando que Z aproximadamente tiene una distribución Normal estandar como se esperaba, por (2.8) (Algoritmo 5).

Conclusiones

Mediante el uso de la teoría básica de Martingalas y de los procesos auto-recursivos fué posible modelar el comportamiento del número de ciclones que impactaron a México de 1990-2010, con la finalidad de predecir el número de ciclones en el año 2011 y 2012.

Nuestra predicción hecha en el 2011 coincidió con el número de ciclones que hubo en ese año.

Además, mediante el uso de simulaciones computacionales, algunos de los resultados referentes a la estimación de los parámetros del modelo auto-recursivo de orden uno (convergencia c.s. y TLCM) fueron ilustrados.

Apéndice

Algoritmo 1:

```
p1<-function()
{
n<-41
s<-numeric()
est.teta<-numeric()
est.sigma2<-numeric()
Z<-numeric()
X<-numeric()
xx<-numeric()
yy<-numeric()
X<-read.csv(file="ciclones.csv",header=T)
for(i in 1:n)
{
xx[i]<-X[i,2]
}
for(i in 1:n)
{
yy[i]<-X[i,1]
}
plot(xx,yy,type="o",xlab="",ylab="",xlim=c(1970,2010),
  ylim=c(0,10))
s[1]<-(X[1,1])^2
est.teta[1]<-0
est.sigma2[1]<-(X[1,1])^2
for(i in 1:(n-1))
{
est.teta[i+1]<-est.teta[i]+(1/s[i])*X[i,1]*
  (X[(i+1),1]-est.teta[i]*X[i,1])
est.sigma2[i+1]<-((i/(i+1))*est.sigma2[i])+
  ((1/(i+1))*(X[(i+1),1]-est.teta[i+1]*X[i,1])^2)
s[i+1]=s[i]+(X[(i+1),1])^2
}
print(est.teta[n])
```

```
print(est.sigma2[n])
#par(mfrow=c(1,2))
#{plot(est.teta, type='l', main = expression
  (paste("Estimación de ",theta)) , xlab=" ", ylab=" ")}
#{plot(est.sigma2, type='l', main = expression
  (paste("Estimación de ",sigma^2)) , xlab=" ", ylab=" ")}
}
```

Algoritmo 2:

```
p2<-function()
{
n<-21
s<-numeric()
est.teta<-numeric()
est.sigma2<-numeric()
Z<-numeric()
X<-numeric()
xx<-numeric()
X<-read.csv(file="ciclones.csv",header=T)
for(i in 1:n)
{
X[i,1]<-X[20+i,1]
}
for(i in 1:n)
{
xx[i]<-X[i,1]
}
#plot(xx,type="l")
###ESTIMACIÓN DE THETA Y LA VARIANZA###
s[1]<-(X[1,1])^2
est.teta[1]<-0
est.sigma2[1]<-(X[1,1])^2
for(i in 1:(n-1))
{
est.teta[i+1]<-est.teta[i]+(1/s[i])*X[i,1]*
  (X[(i+1),1]-est.teta[i]*X[i,1])
est.sigma2[i+1]<-((i/(i+1))*est.sigma2[i])+
  ((1/(i+1))*(X[(i+1),1]-est.teta[i+1]*X[i,1])^2)
s[i+1]=s[i]+(X[(i+1),1])^2
}
print(est.teta[n])
print(est.sigma2[n])
#par(mfrow=c(1,2))
#{plot(est.teta, type='l', main = expression
  (paste("Estimación de ",theta)) , xlab=" ", ylab=" ")}
```

```
#{plot(est.sigma2, type='l', main = expression
  (paste("Estimación de ",sigma^2)) , xlab=" ", ylab=" ")}
###MEDIA Y VARIANZA DE LOS RESIDUOS###
eps<-numeric() #residuo
eps[1]<-0
for(i in 1:(n-1))
{
eps[i+1]<-X[(i+1),1]-(est.teta[n]*X[i,1])
}
res<-numeric()
res[1]<-X[1,1]
for(j in 2:n)
{
res[j]<-eps[j]
}
a<-mean(res)
b<-var(res)
print(a)
print(b)
###PRUEBAS Y DISTRIBUCIÓN DE LOS RESIDUOS###
library(nortest)
ad<-ad.test(res)
#print(ad)
ts<-t.test(res)
#print(ts)
{
mu<-0
sigma<-1
x<-c(-400:400)/100
fx<-(1/sqrt(2*pi*sigma))*exp((-1/2)*(((x-mu)/sigma)^2))
}
h.eps<-(res-a)/sqrt(b)
#plot(x,fx,xlim=c(-5,5),ylim=c(0,1),type="l",
  col="red",xlab=" ",ylab=" ")
#hist(h.eps,add=T,seq(-5, 5, by=1.15),xlim=c(-5,5),
  ylim=c(0,1),freq=F,xlab=NULL,ylab=NULL)
###PRONOSTICOS###
y<-numeric()
```

```
a<-0
m<-10000
for(i in 1:m)
{
y[i]<-rnorm(1,mean=0,sd=sqrt(b))
a<-a+y[i]
}
yy<-a/m
x22<-(est.teta[n])*(X[n,1])+yy
print(x22)
a<-0
for(i in 1:m)
{
y[i]<-rnorm(1,mean=0,sd=sqrt(b))
a<-a+y[i]
}
yy<-a/m
x23<-(est.teta[n])*x22+yy
print(x23)
}
```

Algoritmo 3:

```
p3<-function()
{
n<-21
s<-numeric()
est.teta<-numeric()
est.sigma2<-numeric()
Z<-numeric()
X<-numeric()
xx<-numeric()
X<-read.csv(file="ciclones.csv",header=T)
for(i in 1:n)
{
X[i,1]<-X[20+i,1]
}
x21<-X[21,1]
for(i in 1:(n-1))
{
X[i,1]<-X[i+1,1]-X[i,1]
}
for(i in 1:(n-1))
{
xx[i]<-X[i,1]
}
#plot(xx,type="o",xlab=" ",ylab=" ")
###ESTIMACIÓN DE THETA Y LA VARIANZA###
s[1]<-(X[1,1])^2
est.teta[1]<-0
est.sigma2[1]<-(X[1,1])^2
for(i in 1:(n-2))
{
est.teta[i+1]<-est.teta[i]+(1/s[i])*X[i,1]*
  (X[(i+1),1]-est.teta[i]*X[i,1])
est.sigma2[i+1]<-((i/(i+1))*est.sigma2[i])+
  ((1/(i+1))*(X[(i+1),1]-est.teta[i+1]*X[i,1])^2)
s[i+1]=s[i]+(X[(i+1),1])^2
}
```

```
print(est.teta[n-1])
print(est.sigma2[n-1])
#par(mfrow=c(1,2))
#{plot(est.teta, type='l', main = expression
  (paste("Estimación de ",theta)) , xlab=" ", ylab=" ")}
#{plot(est.sigma2, type='l', main = expression
  (paste("Estimación de ",sigma^2)) , xlab=" ", ylab=" ")}
###MEDIA Y VARIANZA DE LOS RESIDUOS###
eps<-numeric() #residuo
eps[1]<-0
for(i in 1:(n-2))
{
eps[i+1]<-X[(i+1),1]-(est.teta[n-1]*X[i,1])
}
res<-numeric()
res[1]<-X[1,1]
for(j in 2:(n-1))
{
res[j]<-eps[j]
}
a<-mean(res)
b<-var(res)
print(a)
print(b)
###PRUEBAS Y DISTRIBUCIÓN DE LOS RESIDUOS###
library(nortest)
ad<-ad.test(res)
#print(ad)
ts<-t.test(res)
#print(ts)
{
mu<-0
sigma<-1
x<-c(-400:400)/100
fx<-(1/sqrt(2*pi*sigma))*exp((-1/2)*(((x-mu)/sigma)^2))
}
h.eps<-(res-a)/sqrt(b)
#plot(x,fx,xlim=c(-5,5),ylim=c(0,0.7),type="l",
```

```
  col="red",xlab=" ",ylab=" ")
#hist(h.eps,add=T,seq(-5, 5, by=1.15),xlim=c(-5,5),
  ylim=c(0,0.7),freq=F,xlab=NULL,ylab=NULL)
###PRONOSTICOS###
y<-numeric()
a<-0
m<-10000
for(i in 1:m)
{
y[i]<-rnorm(1,mean=0,sd=sqrt(b))
a<-a+y[i]
}
yy<-a/m
d21<-(est.teta[n-1])*(X[n-1,1])+yy
x22<-d21+x21
print(x22)
a<-0
for(i in 1:m)
{
y[i]<-rnorm(1,mean=0,sd=sqrt(b))
a<-a+y[i]
}
yy<-a/m
d22<-(est.teta[n-1])*d21+yy
x23<-d22+x22
print(x23)
}
```

Algoritmo 4:

```
p4<-function()
{
theta<-0.3
sigma<-sqrt(2)
s<-numeric()
est.teta<-numeric()
est.sigma2<-numeric()
X<-rnorm(1, mean=0, sd=sigma)
s[1]<-X^2
est.teta[1]<-0
est.sigma2[1]<-X^2
n<-30000
for(i in 1:(n-1))
{
eps<-rnorm(1, mean=0, sd=sigma)
Y<-X
X<-theta*Y+eps
est.teta[i+1]<-est.teta[i]+(1/s[i])*Y*(X-est.teta[i]*Y)
est.sigma2[i+1]<-(i/(i+1))*est.sigma2[i]+(1/(i+1))*
  (X-est.teta[i+1]*Y)^2
s[i+1]=s[i]+X^2
}
print(est.teta[n])
print(est.sigma2[n])
sig<-numeric()
for(j in 1:n)
{
sig[j]<-sigma^2
}
th<-numeric()
for(j in 1:n)
{
th[j]<-theta
}
par(mfrow=c(1,2))
{plot(est.teta, type='l', main = expression
```

```
  (paste("Estimación de ",theta)) , xlab=" ", ylab=" ")
par(new=TRUE)
lines(th, type='l',col="red")}
{plot(est.sigma2, type='l', main = expression
  (paste("Estimación de ",sigma^2)) , xlab=" ", ylab=" ")
par(new=TRUE)
lines(sig, type='l',col="blue")}
}
```

Algoritmo 5:

```
p5<-function()
{
theta<-0.3
sigma<-sqrt(2)
mu<-0
s<-numeric()
est.teta<-numeric()
est.sigma2<-numeric()
Z<-numeric()
n<-30000
nr<-1000
for(k in 1:nr)
{
X<-rnorm(1, mean=mu, sd=sigma)
s[1]<-X^2
est.teta[1]<-0
est.sigma2[1]<-X^2
for(i in 1:(n-1))
{
eps<-rnorm(1, mean=0, sd=sigma)
Y<-X
X<-theta*Y+eps
est.teta[i+1]<-est.teta[i]+(1/s[i])*Y*(X-est.teta[i]*Y)
est.sigma2[i+1]<-(i/(i+1))*est.sigma2[i]+(1/(i+1))*
  (X-est.teta[i+1]*Y)^2
s[i+1]=s[i]+X^2
}
#print(est.teta[n])
#print(est.sigma2[n])
Z[k]<-(sqrt(n)*(est.teta[n]-theta))/sqrt(1-theta^2)
}
{
mu<-0
sigma<-1
x<-c(-400:400)/100
fx<-(1/sqrt(2*pi*sigma))*exp((-1/2)*(((x-mu)/sigma)^2))
```

```
}
plot(x,fx,xlim=c(-4,4),ylim=c(0,0.5),type="l",main="TCLM",
  col="red",xlab=" ",ylab=" ")
hist(Z,add=T,xlim=c(-4,4),ylim=c(0,0.5),freq=F,xlab=NULL,
  ylab=NULL)
}
```

Bibliografía

[1] Bernard Bercu Djalil Chafaï, *Modélisation stochastique et simulation Cours et applications*, Dunod, Sciences Sup, Mathématiques appliqueés pour le Master/SMAÍ

[2] Zdzislaw Brzeźniak and Tomasz Zastawniak, *Basic Stochastic Processes*, Springer

[3] Luis Rincón, *Introducción a los Procesos Estocásticos*, Versión: Enero 2011

[4] Luis Rincón, *Curso Intermedio de Probabilidad*, Versión: Octubre 2007

[5] Marek Capiński and Ekkehard Kopp, *Measure, Integral and Probability*, Second Edition, Springer

[6] Bernard B., Proia F., *A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first order autoregresive process*, ESAIMPS, Vol. 16, 2012

[7] Ruppert D., *Statistics and finance, an introduction*, Nueva York, Springer, Verlang 2004, 485 pp.

[8] Charu Chandra, Janis Grabis, *Application of multi-steps forecasting for restraining the bullwhip effect and improving inventory performance under autoregressive demand*, European Journal of Operational Research, Vol. 166, Issue 2, 16 October 2005, pages 337-350

[9] Roberts, S.J., *Variational Bayes for generalized autoregressive models*, Signal Processing, IEEE Transactions on, Vol. 50, Sep 2002, pages 2245-2257

[10] Jouchi Nakajima, Munehisa Kasuya, Toshiaki Watanabe, *Bayesian analysis of time-varying parameter vector autoregressive model for the Japanese economy and monetary policy*, Journal of the Japanese and International Economies, Vol. 25, Issue 3, Sep 2011, pages 225 - 245

[11] André Fujita, Patricia Severino, Joao Ricardo Sato and Satoru Miyano, *Granger Causality in Systems Biology: Modeling Gene Networks in Time Series Microarray Data Using Vector Autoregressive Models*, Advances in Bioinformatics and Computational Biology Lecture Notes in Computer Science, 2010, Volume 6268/2010, 13-24

[12] Xiangjun Zhang, *Image Interpolation by Adaptive 2-D Autoregressive Modeling and Soft-Decision Estimation*, Image Processing, IEEE Transactions on, Vol. 17 , Issue: 6, pages 887 - 896