UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

PRONÓSTICO EN SERIES DE TIEMPO DE MEDIA MÓVIL FINITAS SUSTENTADO DESDE LOS ESPACIOS DE HILBERT

Boris Esteban Palacios Morales

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS
PROYECTO CURRICULAR DE MATEMÁTICAS
BOGOTÁ D.C.
2017

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

PRONÓSTICO EN SERIES DE TIEMPO DE MEDIA MÓVIL FINITAS SUSTENTADO DESDE LOS ESPACIOS DE HILBERT

Boris Esteban Palacios Morales

Trabajo presentado como parte de los requisitos para la obtención del título de Matemático por la Universidad Distrital Francisco José de Caldas.

Director:

Prof. Luis Fernando Villarraga Poveda

Dedicado a
mi familia

Agradecimientos

Agradezco principalmente a mi madre Juana Morales por el apoyo brindado durante todo el proceso de pregado, al ser el gran motor de este proyecto creyendo siempre en mí, sin importar las adversidades.

Agradezco a Carolina y Felipe por ser las personas que durante esta etapa de mi vida fueron el complemento perfecto al ser grandes amigos y compañeros de estudio, a la planta docente del pregrado curricular de matemáticas de la Universidad Distrital Francisco José de Caldas por toda su disposición y constancia en la academia, en especial a los profesores Luis Alejandro Másmela y Fernando Villarraga los cuales con su apoyo lograron la culminación de este trabajo.

Índice general

Resumen **1**

1. Introducción **2**

2. Preliminares **3**

2.1. Espacio de Probabilidad 3

2.2. Variables Aleatorias 5

2.3. Convergencia casi segura de variables aleatorias 7

2.4. Valor Esperado 9

2.5. El Espacio $\mathcal{L}^2(\Omega, \mathscr{F}, P)$ 11

2.5.1. $\mathcal{L}^2(\Omega, \mathscr{F}, P)$ es un espacio de Hilbert 12

2.6. El Teorema de la Proyección 15

3. Series de tiempo **20**

3.1. Estacionariedad 21

3.2. Procesos de Medias Móviles 25

3.2.1. Proceso de media móvil de orden q ($MA(q)$) 26

3.3. Estimadores 34

3.4. Predicción en Series de Tiempo $MA(q)$ 35

3.4.1. La Mejor Predicción Lineal en $\mathcal{L}^2(\Omega, \mathscr{F}, P)$ 36

3.4.2. Estimación de Parámetros con el Algoritmo de Innovaciones . . 39

4. Aplicación **41**

5. Conclusiones **46**

Resumen

El presente trabajo tiene como finalidad mostrar las bases matemáticas que permiten realizar predicciones a series estacionarias, más específicamente aquellas series que tienen un comportamiento $MA(q)$ llamados procesos de medias móviles de orden q estas series poseen algunas características especiales en su media, varianza, función de autocovarianzas, funciones de autocorrelación simple y parcial, así como el aspecto de sus correlogramas. Para ello son trabajados conceptos de teoría de la medida enfocados en el espacio $\mathcal{L}^2(\Omega, \mathcal{F}, P)$, un espacio dotado de un producto punto el cual cumple con todas las características necesarias para ser un espacio de Hilbert. Finalmente se realiza una pequeña simulación en R con la cual se busca contrastar toda la teoría trabajada con un conjunto de datos.

1 Introducción

A través de la historia el hombre siempre ha tratado de justificar fenómenos para conocer más sobre su naturaleza y adaptarse a la consecuencia que estos puedan traer, las series de tiempo nacen como un registro de observaciones a una situación de interés cuyas características son aleatorias y poco predecibles al ojo común. Para mitigar las consecuencias negativas que pueda traer cierto evento son necesarios los pronósticos los cuales tienen por hipótesis los registros de eventos ya ocurridos, un buen pronóstico esta basado en la cantidad y exactitud de los datos ya observados. Las series de tiempo están clasificadas como una de las herramientas matemáticas más útiles para realizar predicciones; su eficacia está fuertemente sustentada en sólidos conceptos en campos como teoría de la medida, probabilidad, estadística, álgebra lineal, la programación, entre otras.

2 Preliminares

La comprensión e interpretación de este trabajo se encuentra fundamentado en definiciones, proposiciones y teoremas que tratan conceptos de la teoría de la medida, algunas demostraciones serán referenciadas puesto que su demostración se encuentra en varios textos y algunas se escapan del objeto principal del trabajo, las definiciones en su mayoría son tomadas de [6], otros resultados estarán referenciados puntualmente a través del documento.

2.1 Espacio de Probabilidad

Al momento de definir un espacio en matemáticas, es de interés caracterizarlo para establecer propiedades y cerciorarse que la teoría sobre este es consistente y sólida. La teoría de la medida juega un papel fundamental en la caracterización de espacios, un ejemplo y caso especial son los espacios de probabilidad que cumplen ciertas características importantes que lo enriquecen de propiedades únicas en el conjunto de los espacios medibles.

Definición 2.1.1. *(σ-álgebra)*
Una colección no vacía de subconjuntos $\mathcal{F}$ sobre Ω es llamada σ-álgebra sí:

1. $\emptyset \in \mathcal{F}$.
2. *Si $F \in \mathcal{F}$ entonces $F^c \in \mathcal{F}$.*
3. *Si $F_i \in \mathcal{F}$, es una sucesión contable de conjuntos de $\mathcal{F}$, entonces $\bigcup_{i\in\mathbb{N}} F_i \in \mathcal{F}$.*

*Los miembros de $\mathcal{F}$ son llamados **eventos**, o subconjuntos $\mathcal{F}$-**medibles** de Ω, o simplemente subconjuntos **medibles** de Ω. El par $(\Omega, \mathcal{F})$ es llamado espacio de medible.*

Para cualquier conjunto X, el par $\{\emptyset, X\}$ es una σ-álgebra sobre X, esta es la menor σ-álgebra posible sobre X.

Ejemplo 2.1.1. (σ-álgebra de Borel)
La menor σ-álgebra sobre $\mathbb{R}$ que contiene a todos lo intervalos de la forma $(-\infty, a]$ con $a \in \mathbb{R}$ recibe el nombre de *σ-álgebra de Borel* y se denota por $\mathcal{B}$, al ser $\mathcal{B}$ una σ-álgebra entonces los siguientes conjuntos también están en $\mathcal{B}$

$$\begin{aligned}
(a,\infty) =& \mathbb{R} - (\infty, a],\\
(a,b] =& (-\infty, b] \cap (a,\infty),\\
(-\infty, a) =& \bigcup_{n=1}^{\infty} (-\infty, a - \frac{1}{n}],\\
[a,\infty) =& \mathbb{R} - (-\infty, a),\\
(a,b) =& (-\infty, b) \cap (a,\infty),\\
[a,b] =& \mathbb{R} - ((-\infty, a) \cup (b,\infty)),\\
\{a\} =& [a,a].
\end{aligned}$$

donde $a, b \in \mathbb{R}$ y $a \leq b$, así $(\mathbb{R}, \mathcal{B})$ es un espacio medible.Tomado de [1].

Definición 2.1.2. (Espacio de medida)
Una función $\mu : \mathcal{F} \to [0,\infty]$ se dice que es una medida si cumple que

$$\mu\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mu(F_i) \quad \text{para todo } F_i \in \mathcal{F}, \quad F_i \cap F_j = \emptyset, \quad i \neq j.$$

*la dupla $(\Omega, \mathcal{F})$ dotada por una **medida** μ, forma la tripla $(\Omega, \mathcal{F}, \mu)$ llamada **espacio de medida**.*

Definición 2.1.3. (Espacio de Probabilidad)
*Se dice que $P : \mathcal{F} \to [0,1]$ es una **medida de probabilidad** sobre $(\Omega, \mathcal{F})$ si*

1. $P(\Omega) = 1$.
2. $P\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} P(F_i) \quad \text{para todo } F_i \in \mathcal{F}, \quad F_i \cap F_j = \emptyset, \quad i \neq j.$

Estas son algunas propiedades importantes de un espacio de probabilidad:

- $P(A) \leq P(B), \quad A \subseteq B \quad A, B \in \mathcal{F}.$
- $P(A) = 1 - P(A^c), \quad A \in \mathcal{F}.$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B), \quad A, B \in \mathcal{F}.$

(Demostración en [8] p. 30)

Al momento de realizar predicciones es de sumo interés establecer que tanto pueda estar relacionado un evento de otro, es decir si un resultado puede afectar de manera significativa otro que quiera ser observado, es por ello que se estudia la probabilidad condicional, esta nos muestra la dependencia que existe entre los eventos observados.

Definición 2.1.4. *(Probabilidad Condicional)*
Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $B \in \mathcal{F}$ un evento con $P(B) > 0$. se define una nueva medida de probabilidad sobre $(\Omega, \mathcal{F})$ asumiendo que B ha ocurrido.
*La **Probabilidad de** A **condicionado sobre** (dado) B es*

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad A, B \in \mathcal{F} \;\; y \;\; P(B) > 0. \tag{2.1}$$

Ahora considérese $A_i \in \mathcal{F}, i = 1, 2, \cdots, n$, una partición sobre Ω tal que $P(A_i) > 0$ y sea $B \in \mathcal{F}$ un evento arbitrario, obtenemos de la probabilidad condicional los siguientes resultados.

- La **ley de la probabilidad total**:

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i), \tag{2.2}$$

- La **fórmula de Bayes**:

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{P(B)} = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}, \quad P(B) > 0. \tag{2.3}$$

2.2 Variables Aleatorias

Una variable aleatoria es una función que asigna un valor numérico al resultado observado de los eventos en un experimento aleatorio, esto sucede por que no existe una certeza del valor que esta puede tomar al estar afectada por el azar, este concepto se define a partir de la teoría de la medida.

Definición 2.2.1. *(Función Medible)*
*Sea $(\Omega, \mathcal{F})$ y $(\Psi, \mathcal{G})$ dos espacios medibles y considérese la función $h : \Omega \to \Psi$. La aplicación h es una **función medible** de $(\Omega, \mathcal{F})$ a $(\Psi, \mathcal{G})$ si*

$$h^{-1}(B) = \{w : h(\omega) \in B\} \in \mathcal{F}, \quad \forall B \in \mathcal{G}. \tag{2.4}$$

Sí $h : (\Omega, \mathcal{F}) \to (\Psi, \mathcal{G})$ es medible y P es una medida de probabilidad sobre $(\Omega, \mathcal{F})$ entonces $Q : \mathcal{G} \to [0, 1]$, definida como:

$$Q(B) = P(h^{-1}(B)), \quad B \in \mathcal{G}, \tag{2.5}$$

*es una medida de probabilidad sobre $(\Psi, \mathcal{G})$, este es llamado **Probabilidad inducida por** h.*

Definición 2.2.2. (Variable Aleatoria)
*Si una medida de probabilidad P está asignada a un espacio medible $(\Omega, \mathcal{F})$ y $(\Psi, \mathcal{G}) = (\mathbb{R}, \mathcal{B})$ la función h en (2.4) es llamada **Variable Aleatoria** real.*

Así una variable aleatoria X es una función real definida sobre el espacio de probabilidad $(\Omega, \mathcal{F}, P)$, asociado a un experimento aleatorio.
Si dos variables aleatorias X, Y están definidas sobre $(\Omega, \mathcal{F}, P)$, y α es una constante real, haciendo uso de las propiedades de $\mathcal{F}$ y la definición de variable aleatoria es posible mostrar que $X + Y$, αX, XY, $\frac{X}{Y}$(si está definida), $\text{máx}\{X, Y\}$, $\text{mín}\{X, Y\}$ son también variables aleatorias.

Definición 2.2.3. *(Función de distribución)*
Sea X una variable aleatoria inducida por la función $X : \Omega \to \mathbb{R}$ sobre el espacio medible $(\mathbb{R}, \mathcal{B})$ la función F definida sobre $\mathbb{R}$ por medio de

$$F(B) = P(X^{-1}(B)), \quad B \in \mathcal{B}, \tag{2.6}$$

*es llamada **función de distribución** de la variable aleatoria X.*

Se trabajan dos conceptos importantes para caracterizar las variables aleatorias, la función característica y la variable aleatoria simple. La variable aleatoria simple se construye con suma de funciones características para luego mostrar que toda variable aleatoria definida positiva puede ser expresada como suma de variables aleatorias simples.

Definición 2.2.4. (Función Característica)
La variable aleatoria χ_A definida en Ω, con $A \subset \Omega$ como

$$\chi_A(\omega) = \begin{cases} 1 & si \quad w \in A \\ \\ 0 & si \quad w \notin A \end{cases}$$

*es llamada **función característica** sobre A.*

Definición 2.2.5. (Variable Aleatoria Simple)
*Una variable aleatoria X sobre $(\Omega, \mathcal{F}, P)$ cuyo rango consiste solamente en un conjunto finito de puntos es llamada **Variable Aleatoria Simple**, es decir si I es un conjunto finito y $A_i \in \mathcal{F}$, $i \in I$ es una partición de Ω entonces la expresión*

$$X(\omega) = \sum_{i \in I} x_i \chi_{A_i}(\omega), \quad x_i \in \mathbb{R}$$

define una variable aleatoria simple, donde $\{x_i\}$ son los valores distintos que toma la variable aleatoria simple.

Teorema 2.2.1. *Sea $X : \Omega \to \mathbb{R}$ tal que $X(\omega) \geq 0$ para todo $w \in \Omega$, entonces X es una variable aleatoria si y solo si existe una sucesión de variables aleatorias simples X_n sobre Ω, tal que*

a) $0 \leq X_1 \leq X_2 \leq \cdots \leq X$.

b) $X_n(\omega) \to X(\omega)$, *para cada w en Ω.*

(Demostración en [8], Teorema 5.1.1, p. 118.)

2.3 Convergencia casi segura de variables aleatorias

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, un conjunto $N \in \mathcal{F}$ tal que $P(N) = 0$ es llamado **conjunto nulo**. Una propiedad (p_0) que es verdadera sobre $(\Omega - N)$ se dice que se cumple en **casi toda parte** (c.t.p), **casi seguramente** (c.s), **para todo** w o **con probabilidad uno**.

Las variables aleatorias X, Y definidas sobre un conjunto A se dice que son iguales casi seguramente si el conjunto $\{\omega \in A : X(\omega) \neq Y(\omega)\}$ tiene probabilidad nula, y se nota $X \overset{c.s}{=} Y$.

Definición 2.3.1. *(Igualdad Casi Segura)*
Sean X, Y dos variables aleatorias definidas en $(\Omega, \mathcal{F}, P)$, $X \overset{c.s}{=} Y$ si

$$P(X = Y) = P(\{w : X(\omega) = Y(\omega)\}) = 1.$$

Definida así la igualdad casi segura de variables aleatorias es una relación de equivalencia

- **Reflexividad**
  $P(\{w : X(\omega) \neq X(\omega)\}) = P(\emptyset) = 0 \to X \overset{c.s}{=} X$
- **Simetría**
  Si $X \overset{c.s}{=} Y$ entonces

$$\begin{aligned} 1 = P(X = Y) &= P(\{w : X(\omega) = Y(\omega)\}) \\ &= P(\{w : Y(\omega) = X(\omega)\}) \end{aligned}$$

  así $Y \overset{c.s}{=} X$.

- **Transitividad**
  Sí $X \overset{c.s}{=} Y$ y $Y \overset{c.s}{=} Z$, sean $A = \{w : X(\omega) = Y(\omega)\}$ y $B = \{w : Y(\omega) = Z(\omega)\}$, entonces $P(A) = P(B) = 1$

$$P(A \cap B) = P(\{w : X(\omega) = Z(\omega)\}) = 1$$

  puesto que al ser eventos independientes entonces $P(A \cap B) = P(A)P(B) = 1$, así $X \overset{c.s}{=} Z$.

Definición 2.3.2. *(Convergencia casi segura de variables aleatorias)*
Una sucesión de variables aleatorias $\{X_n\}$ se dice que converge casi seguramente a una variable aleatoria X, si y solo si existe un conjunto N de medida nula tal que

$$\forall w \in \{\Omega - N\} : \lim_{n\to\infty} X_n(\omega) = X(\omega)$$

esto se nota como $\lim_{n\to\infty} X_n \overset{c.s}{=} X$.

Ejemplo 2.3.1. Sea $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}[0,1]$, $P[a,b] = b - a$; con la sucesión de variables aleatorias

$$X_n(\omega) = \begin{cases} n & si \quad 0 \leq \omega \leq \frac{1}{n} \\ 0 & si \quad \frac{1}{n} < \omega \leq 1 \end{cases}$$

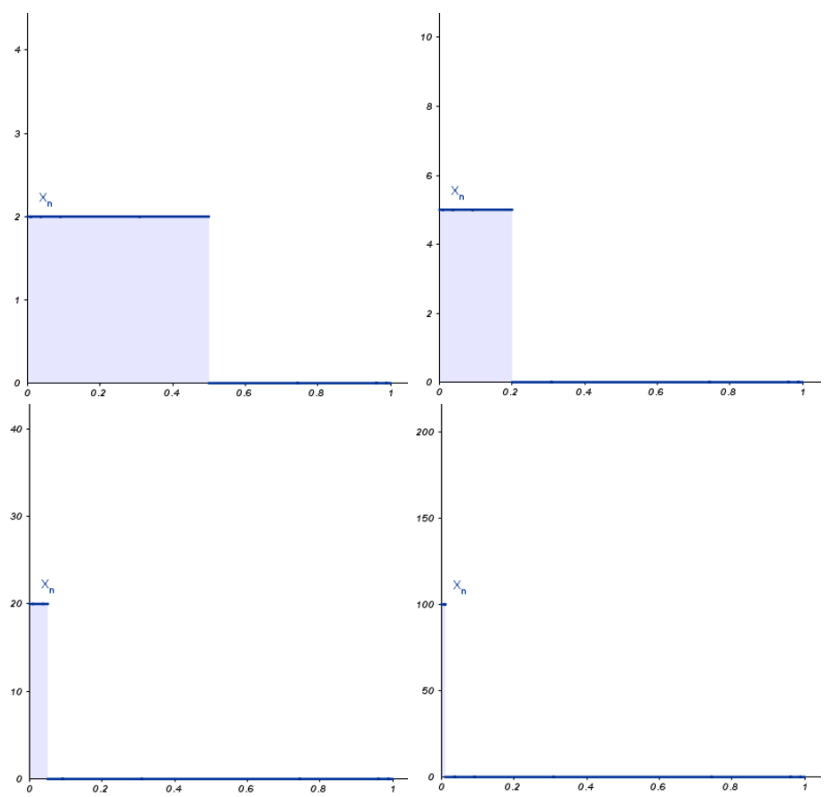


Figura 2.1: Grafica de X_2,X_5,X_{20},X_{100}.

la sucesión de variables aleatorias anterior converge casi seguramente a la constante cero, ya que la igualdad se tiene en el intervalo $(0,1]$, y este tiene probabilidad uno. El punto $\omega = 0$ es el único punto para el cual $X_n(\Omega)$ no converge a cero, por tanto $X_n(\Omega) \overset{c.s}{\to} 0$.

2.4 Valor Esperado

El valor esperado o esperanza matemática de una variable aleatoria representa el valor medio a largo plazo al realizar varias repeticiones es decir es un valor tendencia de un experimento aleatorio. Por ejemplo en el lanzamiento de una moneda corriente 10 veces, en un primer experimento cae 6 veces cara, en un segundo experimento 4, y en un tercer experimento 5, si se hiciera este experimento una cantidad considerada de veces se observaría que existe una tendencia en el número 5, esto tiene sentido basados en el hecho de que en un sólo lanzamiento la probabilidad de tener cara es $1/2$.

Considérese un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, se define el valor esperado de una variable aleatoria simple como

Definición 2.4.1. *(Valor Esperado)*
Si X es una variable aleatoria simple tal que $|x_i| < \infty$, $i \in I$, su ***Valor Esperado*** *es*

$$E[X] = \sum_{i \in I} x_i P(A_i) = \int_\Omega X(\omega) P(d\omega) = \int X dP$$

Las siguientes propiedades son consecuencias de la definición anterior, si X y Y son variables aleatoria simples se tiene que

- $X \geq 0$, entonces $E[X] \geq 0$.
- $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y]$, con α, β constantes.
- $X \leq Y$, entonces $E[X] \leq E[Y]$.

Es posible extender la definición de valor esperado de variables aleatorias simples a variables aleatorias haciendo uso del Teorema 2.2.1, así:

Si $X : \Omega \to [0, \infty]$ es una variable aleatoria, se define

$$E[X] = \int_\Omega X(\omega) P(d\omega) = \int X dP = sup \int X_i dP.$$

Se toma así el supremo de todas las variables aleatorias simples X_i tales que $0 \leq X_i \leq X$.

Se presentan algunas propiedades y teoremas importantes que tiene la integral de una variable aleatoria. (Tomado de [4] p. 44)
Sean X, Y variables aleatorias con α, β constantes y $A \in \mathcal{F}$.

- **Integrabilidad Absoluta**
  $\int_A XdP$ es finita si y solo si
  $$\int_A |X|dP < \infty.$$
- **Linealidad**
  $$\int_A (\alpha X + \beta Y)dP = \alpha \int_A XdP + \beta \int_A YdP.$$
- **Aditividad sobre conjuntos**
  Si la colección $A_i \in \mathcal{F}$ es una partición de Ω, entonces
  $$\int_{\cup_n A_i} XdP = \sum_i \int_{A_i} XdP.$$
- **Positividad**
  Sí $X \geq 0$ c.s sobre A, entonces
  $$\int_A XdP \geq 0$$
- **Monoticídad**
  Sí $X_1 \geq X \geq X_2$ c.s sobre A, entonces
  $$\int_A X_1 dP \geq \int_A XdP \geq \int_A X_2 dP$$

Teorema 2.4.1. ***(Teorema de la convergencia monótona)***
Sí $\lim_{n\to\infty} X_n \overset{c.s}{=} X$ *y si para cada* n, $X_n \geq 0$ *es una suseción creciente que puede tomar un valor finito, entonces*
$$\lim_{n\to\infty} \int_A X_n dP = \int_A (\lim_{n\to\infty} X_n) dP = \int_A XdP.$$

Teorema 2.4.2. ***(Teorema de la convergencia Dominada)***
Sí $\lim_{n\to\infty} X_n \overset{c.s}{=} X$ *y si para cada* n, *existe* Y $|X_n| \leq Y, Y \geq 0$, $\int_A YdP < \infty$ *entonces*
$$\lim_{n\to\infty} \int_A X_n dP = \int_A (\lim_{n\to\infty} X_n) dP = \int_A XdP.$$

Teorema 2.4.3. ***(Lema de Fatou)***
Sí $X_n \geq 0$ c.s sobre A, entonces

$$\int_A (\varliminf_{n\to\infty} X_n) dP \leq \varliminf_{n\to\infty} \int_A X_n dP.$$

Observación 1. Se define el límite inferior y superior de la siguiente manera

$$\varliminf_{n\to\infty} X_n = \sup_{n\geq 1} \inf_{k\geq n} X_k$$
$$\varlimsup_{n\to\infty} X_n = \inf_{n\geq 1} \sup_{k\geq n} X_k$$

2.5 El Espacio $\mathcal{L}^2(\Omega, \mathscr{F}, P)$

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad se denota por $\mathcal{L}^2(\Omega, \mathscr{F}, P)$ la colección de variables aleatorias reales X definidas sobre $(\Omega, \mathscr{F}, P)$ tal que $E[|X|^q] < \infty$ para $q \geq 1$. Si X esta en $L^q(\Omega, \mathscr{F}, P)$, se escribe $X \sim \mathcal{L}^q(\Omega, \mathscr{F}, P)$ o $X \sim L^q$. El caso $q = 2$ es el más importante debido a sus propiedades. Se define ahora el espacio $\mathcal{L}^2(\Omega, \mathcal{F}, P)$.

$$\mathcal{L}^2(\Omega, \mathcal{F}, P) = \{X : \Omega \to \mathbb{R}, X \ \ es \ \ v.a \ \ y \ \ \int_\Omega X(\omega)^2 P(d\omega) < \infty\}$$

El espacio $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ es la colección de todas las variables aleatorias definidas en el espacio $(\Omega, \mathcal{F}, P)$ que satisfacen la siguiente condición

$$E(X^2) = \int_\Omega X(\omega)^2 P(d\omega) < \infty$$

así $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ un subespacio vectorial de las variables aleatorias definidas sobre Ω.

1. **Elemento Neutro**
   Sea $X_0 : \Omega \to \mathbb{R}$ tal que $X_0(\omega) = 0$ para todo $\omega \in \Omega$ obtenemos que

$$E(X_o^2) = \int_\Omega X_0(\omega)^2 P(d\omega) = 0 < \infty$$

2. **Cerradura bajo la suma**
   Consideremos

$$\begin{aligned}(X+Y)^2 + (X-Y)^2 =& X^2 + 2XY + Y^2 + X^2 - 2XY + Y^2 \\ =& 2X^2 + 2Y^2\end{aligned}$$

por tanto $(X+Y)^2 \leq 2X^2 + 2Y^2$

$$E[(X+Y)^2] \leq 2E[X^2] + 2E[Y^2] < \infty$$

Para todo $X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$.

3. **Cerradura bajo la multiplicación por escalar**

$$E[(aX)^2] = a^2 E[X^2] < \infty$$

Para todo $a \in \mathbb{R}$ y $X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$.

2.5.1 $\mathcal{L}^2(\Omega, \mathscr{F}, P)$ es un espacio de Hilbert

Proposición 2.5.1. Para cualquier $X, Y \in \mathcal{L}^2(\Omega, \mathscr{F}, P)$

$$< X, Y >= E[XY] = \int_\Omega X(\omega)Y(\omega)dP(\omega) = \int_\Omega XYdP$$

es producto interno de $\mathcal{L}^2(\Omega, \mathscr{F}, P)$.

Demostración. Se verifican las propiedades del producto punto

- **Simetría**

$$< X, Y >= E[XY] = E[YX] =< Y, X >$$

- **Aditividad**

$$\begin{aligned} < X, Y+Z >=& E[X(Y+Z)] \\ =& E[X(Y+Z)] \\ =& E[XY + XZ] \\ =& E[XY] + E[XZ] \\ =& < X, Y > + < X, Z > \end{aligned}$$

la aditividad a izquierda es análoga.

- **Homogéneidad**

$$< \alpha X, Y >= E[\alpha XY] = \alpha E[XY] = \alpha < X, Y >$$

para todo $\alpha \in \mathbb{R}$.

- **Positividad**

$$< X, X >= E[XX] = E[X^2] = \int_\Omega X(\omega)^2 P(d\omega) \geq 0$$

Si $X = 0$, entonces

$$< X, X >= E[(0)(0)] = E[0] = 0$$

Si $< X, X >= 0$, entonces

$$0 = E[X^2] = \int_\Omega X(\omega)^2 dP$$

Se define el conjunto $A_n = \{w : X^2(\omega) > 1/n\}$, donde $X^2(\omega) \geq X^2(\omega)\chi_{A_n}(\omega)$

$$\begin{aligned} 0 &= \int_\Omega X(\omega)^2 dP \\ &\geq \int_\Omega X^2(\omega)\chi_{A_n}(\omega) dP \\ &= \int_{A_n} X^2(\omega) dP \\ &\geq \int_{A_n} \frac{1}{n} dP \\ &= \frac{P(A_n)}{n} \end{aligned}$$

por tanto $P(A_n) = 0$, veamos que $\bigcup_n A_n = \{\omega : X^2(\omega) > 0\}$ entonces

$$P(\{\omega : X^2(\omega) \neq 0\}) = P(\{\omega : X^2(\omega) > 0\}) = P(\bigcup_n A_n) \leq \sum_{n=1}^{\infty} P(A_n) = 0$$

por tanto $X \overset{c.s}{=} 0$.

Así la norma de $X \in \mathcal{L}^2$ queda determinada como

$$\|X\| = \sqrt{< X, X >} = \sqrt{E[x^2]}. \tag{2.7}$$

Definición 2.5.1. *(Convergencia en media cuadrática)*
Se dice que la sucesión $\{X_n\}$ de variables aleatorias converge a X en media cuadrática en $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ si

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

se nota $X_n \overset{m.s.}{\to} X$.

Definición 2.5.2. *(Convergencia en norma de $\mathcal{L}^2(\Omega, \mathcal{F}, P)$)*
Se dice que la sucesión $\{X_n\}$ de variables aleatorias converge a X en norma de $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ si solo si

$$\lim_{n\to\infty} \|X_n - X\|_2 = \lim_{n\to\infty} \left(\int_\Omega |X_n - X|^2 dP \right)^{\frac{1}{2}} = 0.$$

se nota $X_n \overset{\mathcal{L}^2}{\to} X$.

Observación 2. Se puede notar que

$$\lim_{n\to\infty} \|X_n - X\|_2^2 = \lim_{n\to\infty} E|X_n - X|^2 = 0$$

es decir que la convergencia en norma $\mathcal{L}^2$ es equivalente a la convergencia en media cuadrática de la definición 2.5.1.

Para probar que $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ es un espacio de Hilbert, se debe establecer la completitud, luego se necesita probar que $X_n \in \mathcal{L}^2, n = 1, 2, ...$, y $\|X_n - X_m\| \to 0$ como $m, n \to \infty$, entonces existe $X \in \mathcal{L}^2$ tal que $X_n \overset{m.s.}{\to} X$.

Proposición 2.5.2. Si $X_n \in \mathcal{L}^2$ y $\|X_{n+1} - X_n\| \leq 2^{-n}, n = 1, 2, ...$, entonces existe una variable aleatoria X en $\mathcal{L}^2(\Omega, \mathscr{F}, P)$ tal que $X_n \overset{c.s}{\to} X$.

Demostración. Sea $X_0 = 0$. Entonces $X_n = \sum_{j=1}^{n} (X_j - X_{j-1})$. Se define $\sum_{j=1}^{\infty} |X_j - X_{j-1}|$

$$\begin{aligned}
E\Big[\sum_{j=1}^{\infty} |X_j - X_{j-1}|\Big] &= \int_\Omega \sum_{j=1}^{\infty} |X_j - X_{j-1}| dP \\
&= \sum_{j=1}^{\infty} \int_\Omega |X_j - X_{j-1}| dP \quad \textit{(Teorema de la convergencia Mónotona)} \\
&= \sum_{j=1}^{\infty} E[|X_j - X_{j-1}|] \\
&\leq \sum_{j=1}^{\infty} \|X_j - X_{j-1}\| \qquad \textit{(Desigualdad de Cauchy-Schwarz)} \\
&= \|X_1\| + \sum_{j=1}^{\infty} \|X_j - X_{j-1}\| \\
&\leq \|x_1\| + \sum_{j=1}^{\infty} \frac{1}{2^j} \\
&< \infty
\end{aligned}$$

así que existe $X' < \infty$ tal que $\lim_{n\to\infty} \sum_{j=1}^{n} |X_j - X_{j-1}| \overset{c.s}{=} X'$, por tanto

$\lim_{n\to\infty} \sum_{j=1}^{n} (X_j - X_{j-1}) = \lim_{n\to\infty} X_n$ también converge casi seguramente a algún X.

Teorema 2.5.1. *$\mathcal{L}^2(\Omega, \mathscr{F}, P)$ es completo.*

Demostración. Si X_n es una sucesion de Cauchy en $\mathcal{L}^2$ se escoge la subsucesión $(X_{n_k} : k = 1, 2, \cdots)$ tal que $n_k < n_{k+1}$ y además

$$\|X_n - X_m\| \leq 2^{-k} \ \ para \ \ n, m > n_k$$

Por la proposición anterior, existe una variable aleatoria X tal que $X_{n_k} \to X$ casi seguramente cuando $k \to \infty$.Ahora

$$\|X_n - X\|^2 = \int |X_n - X|^2 dP = \int \varliminf_{n\to\infty} |X_n - X_{n_k}|^2 dP$$

usando el Lema de Fatou (2.4.3),

$$\|X_n - X\|^2 \leq \varliminf_{k\to\infty} \|X_n - X_{n_k}\|^2$$

el $\varliminf_{k\to\infty} \|X_n - X_{n_k}\|^2$ puede hacerse tan pequeño como se quiera tomando un n lo suficientemente grande ya que $\{X_n\}$ es una sucesión de Cauchy. En consecuencia a esto $\lim_{n\to\infty} \|X_n - X\|^2 = 0$.

Se comprueba ahora que $X \in \mathcal{L}^2$ con la siguiente desigualdad

$$E[X^2] = \|X\| \leq \|X_n - X\| + \|X_n\|$$

El lado derecho de esta desigualdad es finito para un n suficientemente grande.

2.6 El Teorema de la Proyección

En matemáticas el término proyección posee varias definiciones e interpretaciones dependiendo el campo donde se este trabajando. En este documento se utilizará la definición que se encuentra en el álgebra lineal más precisamente el de proyección ortogonal. Una herramienta que nos permite determinar cuando dos variables aleatorias son ortogonales es el valor esperado de sus productos, verificado esto se procede a hallar la proyección la cual es única, esto esta garantizado por el teorema de la proyección.

Considérese X_1, X_2 dos variables aleatorias en $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ y queremos estimar una variable Y usando una combinación lineal de X_1, X_2 de modo que buscamos una

variable $\hat{Y} = \alpha_1 X_1 + \alpha_2 X_2$ como se busca la mejor aproximación se busca minimizar la siguiente función

$$\begin{aligned} S = \|(Y - \hat{Y})\|^2 &= E[(Y - \hat{Y})^2] \\ &= E[(Y - \alpha_1 X_1 - \alpha_2 X_2)^2] \\ &= E[Y^2 - 2\alpha_1 Y X_1 - 2\alpha_2 Y X_2 + \alpha_1^2 X_1^2 + \alpha_2^2 X_2^2 + \alpha_1\alpha_2 X_1 X_2] \\ &= E[Y^2] - 2\alpha_1 E[Y X_1] - 2\alpha_2 E[Y X_2] + \alpha_1^2 E[X_1^2] + \alpha_2^2 E[X_2^2] + \alpha_1\alpha_2 E[X_1 X_2], \end{aligned}$$

Un método puede ser el uso de la derivada con respecto a α_1 y α_2, pero es posible hacer uso de conceptos del álgebra lineal si se define

$$\mathscr{M} = \{\hat{Y} \in \mathcal{L}^2(\Omega, \mathcal{F}, P) : \hat{Y} = \alpha_1 X_1 + \alpha_2 X_2\}$$

así para todo $\hat{Y} \in \mathscr{M}$ se tendrá que $< Y - \hat{Y}, X_i >$ con $i = 1, 2$ así

$$< Y - \alpha_1 X_1 - \alpha_2 X_2, X_i >= E[(Y - \alpha_1 X_1 - \alpha_2 X_2) X_i] = 0 \qquad i = 1, 2.$$

esto establece el siguiente sistema

$$\begin{aligned} \alpha_1 E[X_1^2] + \alpha_2 E[X_2 X_1] &= E[Y X_1], \\ \alpha_1 E[X_1 X_2] + \alpha_2 E[X_2^2] &= E[Y X_2]. \end{aligned}$$

el cual puede ser resuelto conociendo la naturaleza de las variables aleatorias, se mostrará que resolver el sistema anterior es equivalente a minimizar la función S.

Definición 2.6.1. *(Complemento Ortogonal)*
El complemento ortogonal de un subconjunto $\mathscr{M}$ de $\mathscr{H}$ un espacio de Hilbert se define como el conjunto $\mathscr{M}^\perp$ formado por todos los elementos de $\mathscr{H}$ los cuales son ortogonales a cada uno de los elementos de $\mathscr{M}$, es decir

$$x \in \mathscr{M}^\perp \text{ si y sólo sí } < x, y >= 0 \text{ para todo } y \in \mathscr{M}$$

se nota $x \perp y$.

lo anterior extiende la definición de ortogonalidad al determinar cuando un elemento es ortogonal a un conjunto de interés, esto tiene relevancia en las variables aleatorias puesto que se puede establecer cuanta dependencia pueda existir en un conjunto de variables aleatorias respecto a una en particular.

Definición 2.6.2. *(Subespacio Cerrado)*
Un subespacio $\mathscr{M}$ de un espacio de Hilbert $\mathscr{H}$ se dice que es un subespacio cerrado de $\mathscr{H}$ si para toda sucesión $x_n \in \mathscr{M}$ y $\lim_{n\to\infty} \|x_n - x\| = 0$ implica que $x \in \mathscr{M}$.

Teorema 2.6.1. *(El Teorema de la Proyección)*
Sí $\mathscr{M}$ es un subespacio cerrado de un espacio de Hilbert $\mathscr{H}$, entonces

i) Existe un único elemento $\hat{x} \in \mathscr{M}$ tal que

$$\|x - \hat{x}\| = \inf_{y \in \mathscr{M}} \|x - y\| \tag{2.8}$$

ii) $\|x - \hat{x}\| = \inf_{y \in \mathscr{M}} \|x - y\|$ *si y solo si* $\hat{x} \in \mathscr{M}$ *y* $(x - \hat{x}) \in \mathscr{M}^{\perp}$.

el elemento $\hat{x}$ es llamado la proyección ortogonal de x sobre $\mathscr{M}$.

Demostración. i) Si $d = \inf_{y \in \mathscr{M}} \|x - y\|^2$ entonces existe una sucesión $\{y_n\}$ de elementos de $\mathscr{M}$ tal que $\|x - y\|^2 \to d$. Aplicando la ley del paralelogramo, y usando el hecho de que $(y_m + y_n)/2 \in \mathscr{M}$, se puede escribir

$$\begin{aligned} 0 \leq \|y_m - y_n\|^2 &= -4\|(y_m + y_n)/2 - x\|^2 + 2(\|y_n - x\|^2 + \|y_m - x\|^2) \\ &\leq -4d + 2(\|y_n - x\|^2 + \|y_m - x\|^2) \\ &\to 0 \textit{ como } m, n \to \infty \end{aligned}$$

En consecuencia, por el criterio de Cauchy, existe $\hat{x} \in \mathscr{H}$ tal que $\|y_n - \hat{x}\| \to 0$. Ya que $\mathscr{M}$ es cerrado entonces $\hat{x} \in \mathscr{M}$, y por la continuidad del producto interno

$$\|x - \hat{x}\|^2 = \lim_{n \to \infty} \|x - y_n\|^2 = d.$$

para establecer la unicidad, suponga que $\hat{y} \in \mathscr{M}$ y que $\|x - \hat{y}\|^2 = \|x - \hat{x}\|^2 = d$. Entonces aplicando la ley del paralelogramo de nuevo

$$\begin{aligned} 0 \leq \|\hat{x} - \hat{y}\|^2 &= -4\|(\hat{x} + \hat{y})/2 - x\|^2 + 2(\|\hat{x} - x\|^2 + \|\hat{y} - x\|^2) \\ &\leq -4d + 4d = 0 \end{aligned}$$

así que $\hat{y} = \hat{x}$. i)) Si $\hat{x} \in \mathscr{M}$ y $(x - \hat{x}) \in \mathscr{M}^{\perp}$ entonces $\hat{x}$ es el único elemento de $\mathscr{M}$ definido en i) ya que para cualquier $y \in \mathscr{M}$,

$$\begin{aligned} \|x - y\|^2 &=< x - \hat{x} + \hat{x} - y, x - \hat{x} + \hat{x} - y > \\ &= \|x - \hat{x}\|^2 + \|\hat{x} - y\|^2 \\ &\geq \|x - \hat{x}\|^2, \end{aligned}$$

con igualdad si y solo si $y = \hat{x}$. por otro lado si $\hat{x} \in \mathscr{M}$ y $(x - \hat{x}) \notin \mathscr{M}^{\perp}$ entonces $\hat{x}$ no es el elemento más cercano de $\mathscr{M}$ respecto a x ya que

$$\overline{x} = \hat{x} + ay/\|y\|^2$$

es más cercano, donde y es cualquier elemento de $\mathscr{M}$ tal que $< x - \hat{x}, y > \neq 0$ y $a =< x - \hat{x}, y >$.Para ver esto escribimos

$$\begin{aligned}
\|x - \overline{x}\|^2 &=< x - \hat{x} + \hat{x} - \overline{x}, x - \hat{x} + \hat{x} - \overline{x} > \\
&= \|x - \overline{x}\|^2 + |a| / \|y\|^2 + \mathbf{Re} < x - \hat{x}, \hat{x} - \overline{x} > \\
&= \|x - \overline{x}\|^2 - |a| / \|y\|^2 \\
&< \|x - \overline{x}\|^2
\end{aligned}$$

(Demostración tomada de [3], pág 51) En virtud del teorema anterior definimos la función proyección $P_{\mathscr{M}}$ de $\mathscr{H}$ sobre $\mathscr{M}$ la cual para cada $x \in \mathscr{H}$ se garantiza la existencia y unicidad de $\hat{x} \in \mathscr{M}$ tal que $(x - \hat{x}) \in \mathscr{M}^{\perp}$, así

$$P_{\mathscr{M}} x = \hat{x}, \quad x \in \mathscr{H}. \tag{2.9}$$

el elemento $\hat{x}$ es llamado la proyección ortogonal de x sobre $\mathscr{M}$. Se presentan algunas propiedades importantes de la función proyección:

Proposición 2.6.1. Sea $\mathscr{H}$ un espacio de Hilbert y sea $P_{\mathscr{M}}$ la función proyección sobre el subespacio cerrado $\mathscr{M}$ entonces

i) Para cada $x \in \mathscr{H}$ hay una única representación como suma de elementos de $\mathscr{M}$ y $\mathscr{M}^{\perp}$, es decir
$$x = P_{\mathscr{M}} x + P_{\mathscr{M}^{\perp}} x \tag{2.10}$$

ii) Las funciones $P_{\mathscr{M}}$ y $P_{\mathscr{M}^{\perp}}$ son lineales.

iii) $\|x\|^2 = \|P_{\mathscr{M}} x\|^2 + \|P_{\mathscr{M}^{\perp}} x\|^2$.

iv) $P_{\mathscr{M}} x_n \to P_{\mathscr{M}} x$ si $\|x_n - x\| \to 0$.

v) $x \in \mathscr{M}$ si y solo si $P_{\mathscr{M}} x = x$.

vi) $x \in \mathscr{M}^{\perp}$ si y solo si $P_{\mathscr{M}} x = 0$.

Definición 2.6.3. *(Espacio generado cerrado)*
El espacio generado cerrado notado $\overline{sp}\{x_t, t \in T\}$ de un conjunto $\{x_t, t \in T\}$ de un espacio de Hilbert $\mathscr{H}$ se define como el subespacio cerrado más pequeño de $\mathscr{H}$ el cual contiene todos los elementos de $x_t, t \in T$.

Ejemplo 2.6.1. Sí $\mathscr{M} = \overline{sp}\{X_1, \cdots, X_n\}$, donde $\mathscr{M} \in L^2(\Omega, \mathcal{F}, P)$ entonces para cada $X \in L^2(\Omega)$, $P_{\mathscr{M}} X$ es un único elemento de la forma

$$P_{\mathscr{M}} X = \alpha_1 X_1 + \cdots + \alpha_n X_n$$

tal que

$$< X - P_{\mathscr{M}} X, X_j >= 0 \qquad X_j \in \mathscr{M} \qquad j = 1, \cdots, n$$

haciendo uso de las propiedades de el producto punto obtenemos

$$\begin{aligned} < X - P_{\mathscr{M}} X, X_j > &= 0 \qquad X_j \in \mathscr{M} \qquad j = 1, \cdots, n \\ < X, X_j > - < P_{\mathscr{M}} X, X_j > &= 0 \\ < P_{\mathscr{M}} X, X_j > &=< X, X_j > \\ \sum_{i=1}^{n} \alpha_i < X_i, X_j > &=< X, X_j > \end{aligned} \tag{2.11}$$

el teorema de la proyección garantiza la existencia de la solución para $\alpha_1, \cdots, \alpha_n$, y la unicidad por la función proyección $P_{\mathscr{M}}$.

Definición 2.6.4. *(Conjunto Ortonormal)*
Un conjunto $\{e_t, t \in T\}$ de elementos de un espacio con producto punto se dice que es ortonormal si para cada $s, t \in T$,

$$< e_s, e_t >= \begin{cases} 1 & si \quad s = t, \\ \\ 0 & si \quad s \neq t. \end{cases}$$

Ejemplo 2.6.2. Cualquier secuencia $\{Z_t, t \in T\}$ de variables aleatorias normales estándar es un conjunto ortonormal en $\mathcal{L}^2(\Omega, \mathcal{F}, P)$.

- $s \neq t$

$$< Z_s, Z_t >= E[Z_s Z_t] = E[Z_s][Z_t] = (0)(0) = 0$$

- $s = t$

$$\begin{aligned} < Z_s, Z_t >=< Z_s, Z_s > &= E[Z_s Z_s] \\ &= E[Z_s^2] \\ &= V[Z_s] + (E[Z_s])^2 \\ &= 1 + (0)^2 \\ &= 1 \end{aligned}$$

Teorema 2.6.2. *Sí $\{e_1, \cdots, e_k\}$ es un conjunto ortonormal de un espacio de Hilbert $\mathscr{H}$ y $\mathscr{M} = \overline{sp}\{e_1, \cdots, e_k\}$ entonces*

$$P_{\mathscr{M}} x = \sum_{i=1}^{k} < x.e_i > e_i \qquad \forall x \in \mathscr{H}. \tag{2.12}$$

(Demostración en [3], pág 56)

Observación 3. La sucesión de escalares $\{< x, e_i >\}$ son usualmente llamados coeficientes de Fourier de x relativos al conjunto $\{e_1, \cdots, e_k\}$.

3 Series de tiempo

Una serie de tiempo es una lista de variables temporales dadas en magnitudes como minutos, horas, días o años, que se encuentran asociadas a un valor numérico el cual depende del objeto de estudio. Las series de tiempo son un buen método de representación de datos, gráficamente muestran características importantes del fenómeno estudiado. Por ejemplo la cantidad de clientes que asisten a un supermercado cada día, la manufacturación anual de cierto producto, la cantidad de accidentes en una ciudad cada mes son todas situaciones que pueden ser representadas como series de tiempo. Las gráficas que se encuentran en esta sección fueron realizadas con el software R.

Definición 3.0.1. *(Serie de Tiempo)*
Una serie de tiempo es un proceso estocástico cuya familia de variables aleatorias $\{X_t, t \in T\}$ *están definidas sobre un espacio de probabilidad* $\mathcal{L}^2(\Omega, \mathcal{F}, P)$*, donde cada* X_t *es una observación fijada en un instante* t*.*

Ejemplo 3.0.1. (Emisiones de $C0_2$ en Colombia)
La siguiente tabla con datos extraídos de [10], (Datos oficiales del banco mundial) muestran las emisiones de CO_2 (toneladas métricas per cápita) en Colombia.

1960	0,996	1971	1,341	1982	1,582	1993	1,769	2004	1,289
1961	1,073	1972	1,360	1983	1,666	1994	1,835	2005	1,408
1962	1,111	1973	1,423	1984	1,614	1995	1,592	2006	1,436
1963	1,179	1974	1,507	1985	1,560	1996	1,591	2007	1,407
1964	1,168	1975	1,450	1986	1,550	1997	1,680	2008	1,516
1965	1,195	1976	1,504	1987	1,562	1998	1,682	2009	1,600
1966	1,192	1977	1,519	1988	1,590	1999	1,419	2010	1,659
1967	1,217	1978	1,573	1989	1,583	2000	1,434	2011	1,649
1968	1,273	1979	1,639	1990	1,673	2001	1,373	2012	1,704
1969	1,305	1980	1,599	1991	1,636	2002	1,339	2013	1,893
1970	1,288	1981	1,567	1992	1,745	2003	1,362	2014	1,901

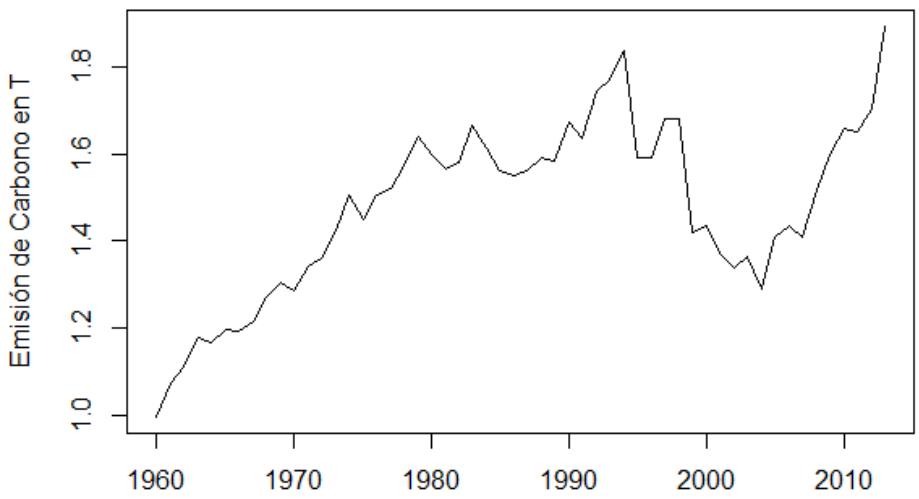


Figura 3.1: Centro de Análisis de Información sobre Dióxido de Carbono en Colombia, División de Ciencias Ambientales del Laboratorio Nacional de Oak Ridge (Tennessee, Estados Unidos).

3.1 Estacionariedad

Definición 3.1.1. *(Función de Autocovarianza)*
Si $\{X_t, t \in T\}$ es un proceso tal que para cada $Var(X_t) = V[X_t] < \infty$ para cada $t \in T$ entonces se define la función de autocovarianzas $\gamma_X(\cdot,\cdot)$ de $\{ X_t\}$ por

$$\begin{aligned}\gamma_X(r,s) = Cov(X_r, X_s) &= E[(X_r - E[X_r])((X_s - E[X_s])] \\ &= E[(X_r - \mu_{X_r})((X_s - \mu_{X_s})] \\ &= E[X_r X_s] - \mu_{X_r}\mu_{X_s} \\ &= E[X_r X_s] - E[X_r]E[X_s] \qquad r,s \in T\end{aligned}$$

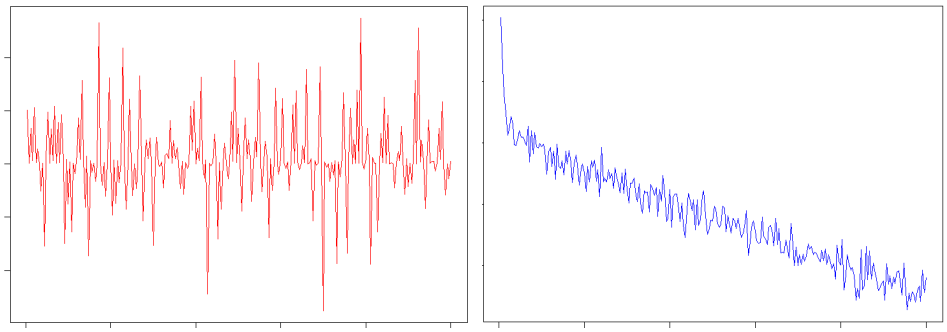
Definición 3.1.2. *(Serie débilmente estacionaria)*
La serie de tiempo $\{X_t, t \in \mathbb{Z}\}$ con el conjunto de índices $\mathbb{Z} = \{0, \pm 1 \pm 2\}$ se dice débilmente si cumple las siguientes propiedades

i) $E[X_t] = \mu_t = m$ *para todo* $t \in \mathbb{Z}$.

ii) $V[X_t] = \sigma_t^2 = p$ *para todo* $t \in \mathbb{Z}$.

iii) $Cov(t, t+h) = (t, t-h) = \gamma_X(h) = \gamma_X(-h)$

con m, p constantes, además la última propiedad nos afirma que γ_X es una función que sólo depende de h y no depende de t.



(a) Media constante (b) Varianza constante

Figura 3.2

Definición 3.1.3. *(Función de autocorrelación Simple (fas))*
Para un proceso estacionario la función de autocorrelación simple se define como

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = Corr(X_{t+h}, X_t) \qquad \forall t, h \in \mathbb{Z}$$

Se comprueba que $\gamma_X(0) = Cov(X_t, X_t) = Var(X_t) = \sigma^2$, en virtud de la definición 3.1.2.

Definición 3.1.4. *(Serie estrictamente estacionaria)*
La serie de tiempo $\{X_t, t \in \mathbb{Z}\}$ con el conjunto de índices $\mathbb{Z} = \{0, \pm 1 \pm 2\}$ se dice estrictamente estacionaria si la familia de distribuciones $(X_{t_1}, X_{t_2}, \cdots, X_{t_k})$ y $(X_{t_1+h}, X_{t_2+h}, \cdots, X_{t_k+h})$ son la misma para todo entero k y todo $t_1, \cdots, t_k, h \in \mathbb{Z}$.

Toda serie de tiempo estrictamente estacionaria es estacionaria, pero el recíproco no es cierto.

Ejemplo 3.1.1. Sea $\{X_t, t \in \mathbb{Z}\}$ distribuida

$$X_t = \begin{cases} X_t \sim N(1,1) & si \quad t \quad es \quad par \\ X_t \sim Exp(1) & si \quad t \quad es \quad impar \end{cases}$$

Se combrueba que $\{X_t\}$ es estacionaria

i) $E[X_t] = 1$ para todo $t \in \mathbb{Z}$, ya que

$$E[X_t] = \begin{cases} 1 & si \quad t \quad es \quad par \\ 1 & si \quad t \quad es \quad impar \end{cases}$$

ii) $V[X_t] = 1$ para todo $t \in \mathbb{Z}$, ya que

$$V[X_t] = \begin{cases} 1 & si \quad t \quad es \quad par \\ 1 & si \quad t \quad es \quad impar \end{cases}$$

iii) $\gamma_X(h) = 0$ para todo $h, t \in \mathbb{Z}$, ya que

$$\begin{aligned}\gamma_X(h) = Cov(X_t, X_{t+h}) &= E[X_t X_{t+h}] - E[X_t]E[X_{t+h}] \\ &= E[X_t]E[X_{t+h}] - E[X_t]E[X_{t+h}] \\ &= 0\end{aligned}$$

sin embargo no es estrictamente ya que (X_1, X_2, X_3, X_4) y (X_2, X_3, X_4, X_5) tienen distintas distribuciones por tanto no puede ser estrictamente estacionaria.

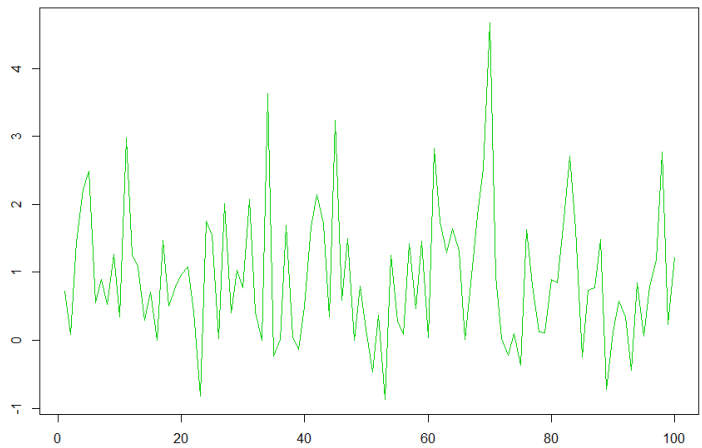


Figura 3.3: Serie $\{X_t\}$ estacionaria, no estrictamente estacionaria.

Definición 3.1.5. *(Serie de tiempo Gaussiana)*
Un proceso $\{X_t\}$ es una serie de tiempo Gaussiana si y solo si las funciones de distribución de $\{X_t\}$ son todas normales multivariadas.

Observación 4. Bajo la hipótesis de normalidad la estacionareidad débil coincide con la estricta.

Una propiedad importante de los procesos estacionarios es el de tener crecimientos estacionarios, es decir si $\{Z_t\}$ es estacionario con media μ_Z, varianza σ_Z^2 y función de autocovarianzas $\gamma_Z(h)$ entonces $\{W_t = Z_t - Z_{t-1}\}$ también lo es, veamos

i) $E[W_t] = E[Z_t - Z_{t-1}] = E[Z_t] - E[Z_{t-1}] = \mu_Z - \mu_Z = 0$ para todo $t \in \mathbb{Z}$

ii)

$$\begin{aligned}V[W_t] &= E[(W_t - \mu_W)^2] \\ &= E[(W_t)^2] \\ &= E[(Z_t - Z_{t-1})^2] \\ &= E[Z_t^2] + E[Z_{t-1}^2] - 2E[Z_t Z_{t-1}] \\ &= 2(V[Z_t] + E[Z_t]^2) - 2E[Z_t Z_{t-1}]\end{aligned}$$

$$
\begin{aligned}
&= 2(\sigma_Z^2 + \mu_Z^2) - 2E[Z_t Z_{t-1}] \\
&= 2(\sigma^2 - E[Z_t Z_{t-1}] + \mu_Z^2) \\
&= 2(\sigma^2 - \gamma_Z(1))
\end{aligned}
$$

iii)

$$
\begin{aligned}
Cov(W_t, W_{t+h}) = \gamma_W(h) &= E[(W_t - \mu_W)(W_{t+h} - \mu_W)] \\
&= E[W_t W_{t+h}] \\
&= E[(Z_t - Z_{t-1})(Z_{t+h} - Z_{t+h-1})] \\
&= E[Z_t Z_{t+h} - Z_t Z_{t+h-1} - Z_{t-1} Z_{t+h} + Z_{t-1} Z_{t+h-1}] \\
&= 2\gamma_Z(h) - \gamma_Z(h-1) - \gamma_Z(h+1)
\end{aligned}
$$

el cual no depende de t, por tanto $\{W_t\}$ es estacionario.

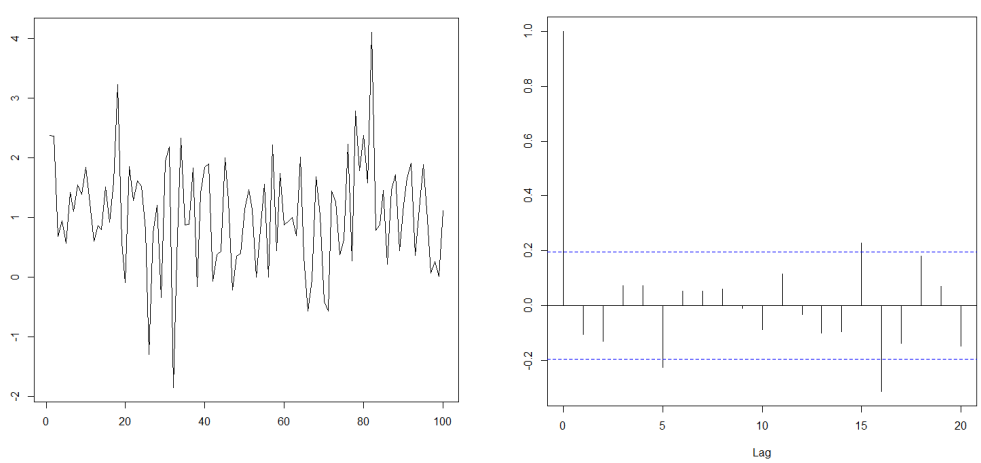
Definición 3.1.6. *(Ruido Blanco)*
Un proceso $\{Z_t, T \in \mathbb{Z}\}$ se dice que es ruido blanco si

i) $E[Z_t] = 0$ *para todo* $t \in \mathbb{Z}$.

ii) $V[Z_t] = \sigma_t^2$ *para todo* $t \in \mathbb{Z}$.

iii) $Cov(t, t+h) = 0$.

es decir sus valores son independientes e idénticamente distribuidos, su notación es $Z_t \sim RB(0, \sigma^2)$.



(a) $Z_t \sim N(0,1)$ (b) Función de aurocorrelacón simple

Figura 3.4: Serie de Tiempo Gaussiana

3.2 Procesos de Medias Móviles

Definición 3.2.1. *(Operador Rezago)*
Sea $\{X_t, T \in \mathbb{Z}\}$ una serie de tiempo el operador rezago (retraso) B, actúa sobre la serie retardando esta un periodo anterior

$$B(X_t) = X_{t-1}.$$

definido así el operador rezago adquiere las siguientes propiedades

- $B^2(X_t) = B(B(X_t)) = B(X_{t-1}) = X_{t-2}$
- $B^p(X_t) = X_{t-p}$
- $B^0(X_t) = I(X_t) = X_t$

Un polinomio de de grado q en el operador rezago B es un operador compuesto por una combinación lineal de las potencias de B

$$\theta_q(B) = \theta_0 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$$

este actúa sobre una variable aleatoria de una serie de tiempo de la forma

$$\begin{aligned}
\theta_q(B)(X_t) &= (\theta_0 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q)X_t \\
&= \theta_0(X_t) + \theta_1 B(X_t) + \theta_2 B^2(X_t) + \cdots + \theta_q B^q(X_t) \\
&= \theta_0 X_t + \theta_1 X_{t-1} + \theta_2 X_{t-1} + \cdots + \theta_q X_{t-q} \\
&= \sum_{j=0}^{q} \theta_j X_{t-j}
\end{aligned}$$

Ejemplo 3.2.1.

$$\begin{aligned}
(2 + 4B^3 - B^4)X_t &= 2X_t + 4BX_t^3 - B^2 X_t \\
&= 2X_t + 4X_{t-3} - X_{t-4}
\end{aligned}$$

Definición 3.2.2. *(Operador Diferencia)*
Sea $\{X_t, T \in \mathbb{Z}\}$ una serie de tiempo, el operador diferencia (retardo) ∇, es igual al binomio $(1 - B)$.

$$\nabla(X_t) = (1 - B)X_t = X_t - BX_t = X_t - X_{t-1}.$$

definido así el operador diferencia y utilizando el binomio de Newton

$$\nabla^p(X_t) = (1 - B)^p X_t = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!}(-1)^j X_{t-j} \tag{3.1}$$

Ejemplo 3.2.2.

$$\begin{aligned}\nabla^2(X_t) &= (1-B)^2 X_t \\ &= (1-2B+B^2)X_t \\ &= X_t - 2X_{t-1} + X_{t-2}\end{aligned}$$

3.2.1 Proceso de media móvil de orden q ($MA(q)$)

Definición 3.2.3. *Una serie $\{X_t\}$ es un proceso $MA(q)$, (Un proceso de medias móviles de orden q) con $q \in \mathbb{N}$, sí la serie es de la forma*

$$X_t = \theta_0 Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-1} + \cdots + \theta_q Z_{t-q} \quad t \in \mathbb{Z} \tag{3.2}$$

donde $Z_t \sim RB(0, \sigma^2)$.

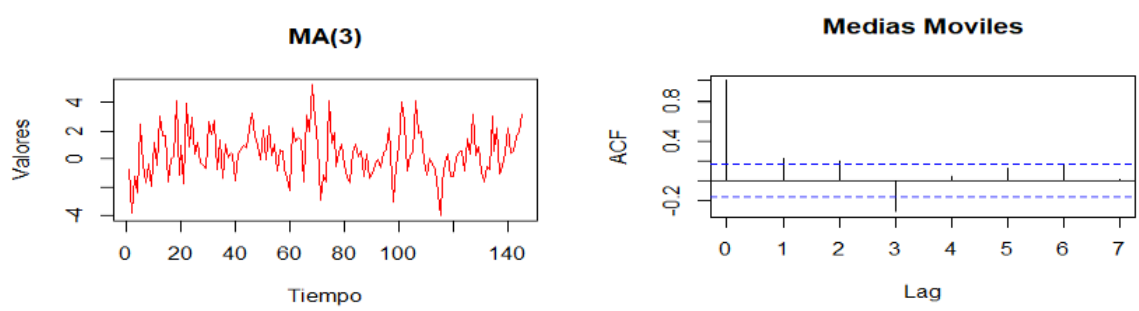
haciendo uso del operador retraso un proceso $MA(q)$ puede escribirse como

$$X_t = \theta_q(B) Z_t = \sum_{j=0}^{q} \theta_j Z_{t-j} \tag{3.3}$$

donde $\theta_q(B)$ es un polinomio de grado q.

Ejemplo 3.2.3. Una serie $MA(3)$ definida como

$$X_t = Z_t + (0{,}4)Z_{t-1} + (0{,}7)Z_{t-2} - (0{,}2)Z_{t-2}$$



(a) Proceso $MA(3)$ (b) Función de autocorrelacón simple

Figura 3.5

Una característica importante que se debe tener en cuenta a la hora de determinar si una serie es un proceso $MA(q)$ es el comportamiento de la función de autocorrelación simple, puesto que esta después de un valor "q"tiende de manera brusca a 0, en el ejemplo (3.2.3) se puede observar como a partir de $\gamma(3)$ los valores de la función de autocorrelación simple tienden a 0, esto se compruebe en el teorema (3.2.1).

Proposición 3.2.1. Un proceso $MA(q)$ con $\theta_0 = 1$ y $\theta_j = 0$ para $j < q$ es estacionario.

Demostración. Se debe comprobar que se cumplen las tres condiciones de estacionariedad.

i)

$$E[X_t] = E\left[\sum_{j=0}^{q} \theta_j Z_{t-j}\right] = \sum_{j=0}^{q} \theta_j E[Z_{t-j}] = 0 \tag{3.4}$$

ii)

$$\begin{aligned}
V[X_t] = E[X_t^2] &= E\left[\left(\sum_{j=0}^{q} \theta_j Z_{t-j}\right)^2\right] \\
&= E\left[\sum_{j=0}^{q} (\theta_j Z_{t-j})^2 + 2\sum_{0<i<j}^{q} \theta_i \theta_j Z_{t-j} Z_{t-i}\right] \\
&= \sum_{j=0}^{q} \theta_j^2 E[Z_{t-j}^2] + 2\sum_{0<i<j}^{q} \theta_i \theta_j E[Z_{t-j} Z_{t-i}] \\
&= \sum_{j=0}^{q} \theta_j^2 V[Z_t] + (E[Z_{t-j}])^2 + 2\sum_{0<i<j}^{q} \theta_i \theta_j (0) \\
&= \sigma^2 \sum_{j=0}^{q} \theta_j^2 \\
&= \sigma^2 (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2) \qquad (3.5)
\end{aligned}$$

iii)

$$\begin{aligned}
Cov(X_t, X_{t+k}) = \gamma_X(k) &= E[X_t X_{t+k}] - E[X_t]E[X_{t+k}] \\
&= E[X_t X_{t+k}] \\
&= E\left[\left(\sum_{i=0}^{q} \theta_i Z_{t-i}\right)\left(\sum_{j=0}^{q} \theta_{Zt+k-j}\right)\right] \\
&= E\left[\left(\sum_{j=0}^{q} \theta_j Z_{t-j}\right)\left(\sum_{i=-k}^{q-k} \theta_{i+k} Z_{t+k-(i+k)}\right)\right] \\
&= E\left[\left(\sum_{j=0}^{q} \theta_j Z_{t-j}\right)\left(\sum_{i=-k}^{q-k} \theta_{i+k} Z_{t-i}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= E\left[\sum_{j=0}^{q}\sum_{i=-k}^{q-k}\theta_j\theta_{i+k}Z_{t-j}Z_{t-i}\right] \\
&= \sum_{j=0}^{q}\sum_{i=-k}^{q-k}\theta_j\theta_{i+k}E[Z_{t-j}Z_{t-i}] \\
&= \sum_{j=0}^{q-k}\theta_j\theta_{j+k}E[Z_{t-j}^2] \qquad si\ i \neq j\ entonces\ E[Z_{t-j}Z_{t-i}] = 0 \\
&= \sum_{j=0}^{q-k}\theta_j\theta_{j+k}V[Z_{t-j}] \\
&= \sigma^2\sum_{j=0}^{q-k}\theta_j\theta_{j+k}
\end{aligned}
$$

el caso para $-k$, es análogo y se obtiene así la siguiente equivalencia

$$
\gamma_X(k) = \begin{cases} \sigma^2\displaystyle\sum_{j=0}^{q-|k|}\theta_j\theta_{j+|k|} & si \quad |k| \leq q, \\ 0 & si \quad |k| > q \end{cases} \tag{3.6}
$$

así la función de autocorrelación simple de un proceso $MA(q)$ queda determinada de la siguiente manera

$$
\rho_X(k) = \begin{cases} \dfrac{\displaystyle\sum_{j=0}^{q-|k|}\theta_j\theta_{j+|k|}}{1+\theta_1^2+\cdots+\theta_q^2} & si \quad |k| \leq q, \\ 0 & si \quad |k| > q \end{cases} \tag{3.7}
$$

Ejemplo 3.2.4. Un proceso $MA(1)$

$$
X_t = Z_t + \theta Z_{t-1}
$$

por (3.4), (3.5) se obtiene $E[X_t] = 0$, $V[X_t] = \sigma^2(1+\theta^2)$ respectivamente. Además por (3.6)

$$
\gamma_X = \begin{cases} \sigma^2(1+\theta^2) & si \quad k = 0 \\ \theta\sigma^2 & si \quad k = 1 \\ 0 & si \quad k > 1 \end{cases}
$$

la función de autocorrelación simple de un proceso $MA(1)$ por (3.7) es:

$$\rho_X = \begin{cases} 1 & si \quad k=0 \\ \frac{\theta}{1+\theta^2} & si \quad k=1 \\ 0 & si \quad k>1 \end{cases}$$

Ejemplo 3.2.5. Un proceso $MA(2)$

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$$

$E[X_t] = 0$, $V[X_t] = \sigma^2(1+\theta_1^2+\theta_2^2)$,

$$\gamma_X = \begin{cases} \sigma^2(1+\theta_1^2+\theta_2^2) & si \quad k=0 \\ \sigma^2(\theta_1+\theta_1\theta_2) & si \quad k=1 \\ \theta_2\sigma^2 & si \quad k=2 \\ 0 & si \quad k>2, \end{cases}$$

$$\rho_X = \begin{cases} 1 & si \quad k=0 \\ \frac{\theta_1+\theta_1^2\theta_2^2}{1+\theta_1^2+\theta_2^2} & si \quad k=1 \\ \frac{\theta_2}{1+\theta_1^2+\theta_2^2} & si \quad k=2 \\ 0 & si \quad k>2 \end{cases}$$

Definición 3.2.4. *(Función de Autocorrelación Parcial)*
La función de autocorrelación parcial de una serie de tiempo $\{X_t, t \in \mathbb{Z}\}$ se define como:

$$\begin{aligned} \alpha(1) &= Corr(X_2, X_1) = \rho(1) \\ \alpha(k) &= Corr(X_{k+1} - P_{\mathscr{M}} X_{k+1}, X_1 - P_{\mathscr{M}} X_1) \end{aligned} \tag{3.8}$$

donde $\mathscr{M} = \overline{sp}\{X_2, \cdots X_k\}$, los valores $\alpha(k)$ son llamados retardos k de la función de autocorrelación parcial.

La función de autocorrelación parcial muestra que tanta correlación existe entre dos variables X_t y X_k omitiendo la linealidad que pueda existir en las observaciones que están entre ellas es decir el conjunto $\{X_2, \cdots X_k\}$.

Es posible obtener una definición alternativa de la función de autocorrelación parcial utilizando las propiedades de el valor esperado. Sea $\{X_t\}$ un proceso con media cero y supongamos que ϕ_{kj}, $j = 1, \cdots, k$; $k = 1, \cdots$ son los coeficientes de la proyección de $\mathscr{M} = \overline{sp}\{X_1, X_2 \cdots X_k\}$ es decir

$$P_{\mathscr{M}} X_{k+1} = \sum_{j=1}^{k} \phi_{kj} X_{k+1-j}$$

por el ejemplo (2.6.1)

$$< X_{k+1} - P_{\mathscr{M}} X_{k+1}, X_j >= 0, \qquad j = k, \cdots, 1.$$

cuya solución se encuentra resolviendo el sistema

$$\begin{bmatrix} \rho(0) & \rho(1) & \rho(2) & \cdots & \rho(k-1) \\ \rho(1) & \rho(0) & \rho(1) & \cdots & \rho(k-2) \\ \rho(2) & \rho(1) & \rho(0) & \cdots & \rho(k-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \rho(k-3) & \cdots & \rho(0) \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \phi_{k3} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \vdots \\ \rho(k) \end{bmatrix}, \qquad k \geq 1. \tag{3.9}$$

Definición 3.2.5. *La función de autocorrelación parcial $\alpha(k)$ de $\{X_t\}$ de retraso k es*

$$\alpha(k) = \phi_{kk}, \qquad k \geq 1, \tag{3.10}$$

donde ϕ_{kk} es únicamente determinado por (3.9).

(Tomado de [3] pág 102)

Ejemplo 3.2.6. La función de autocorrelación parcial de un proceso $MA(1)$ $X_t = Z_t - \theta Z_{t-1}$ es:

$$\alpha(1) = \rho(1) = \frac{\theta}{1+\theta^2}$$

por la ecuación (3.12)

$$\begin{aligned} \lambda < X_2^2 > &=< X_3, X_2 > \\ \lambda E[X_2^2] &= E[X_3 X_2] \\ \lambda V[X_2] &= \gamma_X(1) \\ \lambda \sigma^2 (1+\theta^2) &= \theta \sigma^2 \\ \lambda &= \frac{\theta}{1+\theta^2} \end{aligned}$$

así $P_{\{X_2\}} X_3 = \frac{\theta}{1+\theta^2} X_2$. De manera análoga $P_{\{X_2\}} X_1 = \frac{\theta}{1+\theta^2} X_2$, por tanto

$$\alpha(2) = Corr\left(X_3 - \frac{\theta}{1+\theta^2} X_2, X_1 - \frac{\theta}{1+\theta^2} X_2 \right)$$

a pesar de encontrar los coeficientes de las proyecciones generalmente encontrar la correlación puede ser un trabajo tedioso, así que para $\alpha(2)$ se usará la definición (3.2.5) y la regla de Cramer para sistemas lineales

$$\alpha(2) = \frac{\begin{vmatrix} \rho(0) & \rho(1) \\ \rho(1) & \rho(2) \end{vmatrix}}{\begin{vmatrix} \rho(0) & \rho(1) \\ \rho(1) & \rho(0) \end{vmatrix}} = \frac{\rho(0)\rho(2) - \rho(1)^2}{\rho(0)^2 - \rho(1)^2}$$

$$= \frac{\frac{-\theta^2}{(1+\theta^2)^2}}{\frac{(1+\theta^2)^2-\theta^2}{(1+\theta^2)^2}}$$

$$= \frac{-\theta^2}{(1+\theta^2+\theta^4)}$$

es posible generalizar la función utilizando propiedades de los determinantes e inducción matemática

$$\alpha(k) = -\frac{(-\theta)^k(1-\theta^2)}{1-\theta^{2(k+1)}}.$$

ver ([7]).

Teorema 3.2.1. *Sí X_t es un proceso estacionario con media 0, y con función de autocovarianzas $\gamma_X(\cdot)$ tal que $\gamma_X(h) = 0$ para $|h| > q$ y $\gamma_X(q) \neq 0$, entonces $\{X_t\}$ es un proceso $MA(q)$, es decir que existe un proceso de ruido blanco Z_t tal que*

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q} \tag{3.11}$$

Demostración. Para cada t, definimos el subespacio $\mathscr{M}_t = \overline{sp}\{X_s, -\infty < s \leq t\}$ de $L^2(\Omega, \mathcal{F}, P)$, se define la variable

$$Z_t = X_t - P_{\mathscr{M}_{t-1}} X_t \tag{3.12}$$

$Z_t \in \mathscr{M}_t$, puesto que es la combinación lineal de elementos del conjunto de $\mathscr{M}_t$, además teorema 2.6.1 ii) $Z_t \in \mathscr{M}^{\perp}{}_{t-1}$. En consecuencia si $s < t$, se tiene que

$$Z_s \in \mathscr{M}_s \subset \mathscr{M}_{t-1}$$

por tanto $< Z_s, Z_t >= E[Z_s Z_t] = 0$, y por ser independientes $E[Z_s], E[Z_t] = 0$. Por otro lado si $\{X_t, t = 0, \pm 1 \cdots\}$ es un proceso estacionario entonces

$$\lim_{r\to\infty} P_{\overline{sp}\{X_j, n-r<j\leq n\}} X_{n+1} = P_{\overline{sp}\{X_j, -\infty<j\leq n\}} X_{n+1}. \tag{3.13}$$

entonces

$$P_{\overline{sp}\{X_s, s=t-n,\cdots,t-1\}} X_t \overset{m.s}{\to} P_{\mathscr{M}_{t-1}} X_t \quad \textit{como} \quad n \to \infty$$

utilizando las propiedades de estacionariedad y continuidad de la norma de $\mathcal{L}^2$

$$\begin{aligned}
\|Z_{t+1}\| &= \|X_{t+1} - P_{\mathscr{M}_t} X_{t+1}\| \\
&= \lim_{n\to\infty} \|X_{t+1} - P_{\overline{sp}\{X_s, s=t+1-n,\cdots,t-1\}} X_{t+1}\| \\
&= \lim_{n\to\infty} \|X_t - P_{\overline{sp}\{X_s, s=t-n,\cdots,t-1\}} X_t\| \\
&= \|X_t - P_{\mathscr{M}_{t-1}} X_t\| \\
&= \|Z_t\|
\end{aligned}$$

como $V[Z_t] = E[Z_t^2] - E[Z_t]^2$ y por la ecuación 2.7 se tiene que $\sigma^2 = \|Z_t\|^2$, así se concluye que $\{Z_t\} \sim RB(0, \sigma^2)$.
A partir de (3.12) se verifica para algún $\alpha \in \mathbb{R}$ que

$$\begin{aligned}
\mathscr{M}_{t-1} &= \overline{sp}\{X_s, -\infty < s \leq t-1\} \\
&= \overline{sp}\{X_s, -\infty < s < t-1\} + \alpha X_{t-1} \\
&= \overline{sp}\{X_s, -\infty < s < t-1\} + \alpha Z_{t-1} + \alpha P_{\mathscr{M}_{t-2}} X_{t-1} \\
&= \overline{sp}\{X_s, -\infty < s < t-1\} + \alpha Z_{t-1} \\
&= \overline{sp}\{X_s, -\infty < s < t-1, Z_{t-1}\}
\end{aligned}$$

en general

$$\mathscr{M}_{t-q} = \overline{sp}\{X_s, -\infty < s < t-q, Z_{t-q}, \cdots, Z_{t-1}\}$$

esto implica que $\mathscr{M}_{t-1}$ puede ser descompuesto en dos subespacios ortogonales $\mathscr{M}_{t-q-1}$ y $\overline{sp}\{Z_{t-q} \cdots Z_{t-1}\}$. Si $Y_t \in \mathscr{M}_{t-q-1}$ entonces

$$\begin{aligned}
E[X_t Y_t] &= E\left[X_t \sum_{i=1}^{\infty} \alpha_i X_{t-q-i}\right] \\
&= E\left[\sum_{i=1}^{\infty} \alpha_i X_t X_{t-q-i}\right] \\
&= \sum_{i=1}^{\infty} \alpha_i E[X_t X_{t-q-i}] \\
&= \sum_{i=1}^{\infty} \alpha_i E[X_t] E[X_{t-q-i}] \\
&= \sum_{i=1}^{\infty} \alpha_i (0)(0) \\
&= 0
\end{aligned}$$

ya que por hipótesis $\gamma_X(h) = 0$ para $|h| > q$ y $E[X_t] = 0$ para todo t, esto sigue que

$X_t \perp \mathscr{M}_{t-q-1}$ utilizando la proposición 2.6.1 $vi)$ y el teorema 2.6.2

$$\begin{aligned} P_{\mathscr{M}_{t-1}} X_t &= P_{\mathscr{M}_{t-q-1}} X_t + P_{\overline{sp}\{Z_{t-q}\cdots Z_{t-1}\}} X_t \\ &= 0 + \sum_{i=1}^{q} < X_t, Z_{t-i} > Z_{t-i} \\ &= \sum_{i=1}^{q} \sigma^{-2} E[X_t Z_{t-i}] Z_{t-i} \\ &= \sum_{i=1}^{q} \theta_i Z_{t-i} \end{aligned}$$

donde $\sigma^{-2}E[X_t Z_{t-i}] = \theta_i$, por la estacionariedad de X_t este es independiente de t para $i = 1, \cdots, q$. Sustituyendo por $P_{\mathscr{M}\,t-1} X_t$ (3.12) se obtiene (3.11). ▌

gracias a este resultado y con ayuda de la función de autocorrelación simple es posible determinar cuando un proceso tiene comportamiento $MA(q)$.

Ejemplo 3.2.7. Sea X_t el proceso $MA(5)$ con $\mu = 10$ y Z_t con $\sigma^2 = 2$

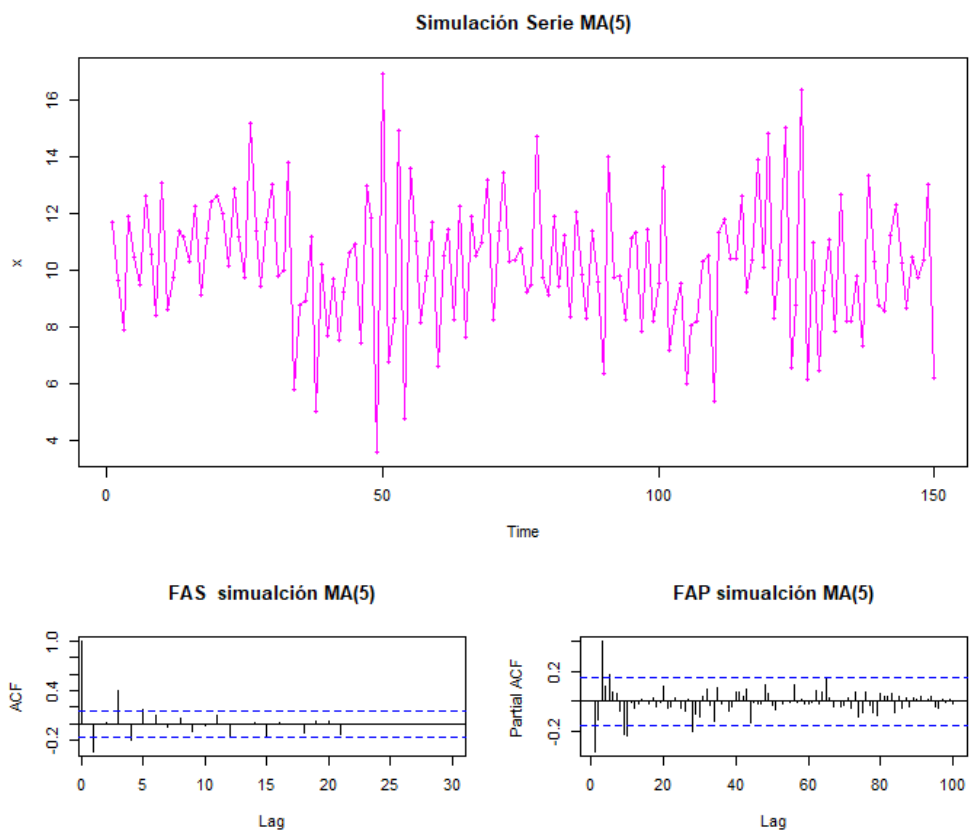


Figura 3.6: $X_t - 10 = Z_t - (0{,}5)Z_{t-1} + (0{,}2)Z_{t-2} + (0{,}3)Z_{t-3} - (0{,}2) + Z_{t-4} + (0{,}4)Z_{t-5}$.

3.3 Estimadores

Al poseer un conjunto de datos de una serie temporal es importante determinar sus características, es por ello que es de vital importancia obtener estimadores de su media, función de covarianzas, función de autocorrelación y función de autocorrelación parcial.

- **Estimación de la media**
  Sea $\{X_t\}$ una serie de tiempo, la media muestral se define como

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{X_1 + X_2 + \cdots + X_n}{n} \tag{3.14}$$

  si cuando $n \to \infty$ se obtiene que $\gamma_X(n) = 0$, el error cuadrático de este estimador converge a 0. (Ver [3] pág 219)

- **Estimación de $\gamma_X(\cdot)$ y $\rho(\cdot)$**
  Los estimadores que se deben utilizar para $\gamma_X(\cdot)$ y $\rho(\cdot)$ son

$$\hat{\gamma}_X(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \overline{X}_n)(X_{t+h} - \overline{X}_n), \quad 0 \leq h \leq n-1, \tag{3.15}$$

  y

$$\hat{\rho}(h) = \frac{\hat{\gamma}_X(h)}{\hat{\gamma}_X(0)}. \tag{3.16}$$

- **Estimación de la función de autocorrelación parcial $\alpha(\cdot)$**

  Utilizando la estimación del de $\hat{\rho}(\cdot)$ y la solución al sistema (3.9)

$$\begin{bmatrix} \hat{\rho}(0) & \hat{\rho}(1) & \hat{\rho}(2) & \cdots & \hat{\rho}(k-1) \\ \hat{\rho}(1) & \hat{\rho}(0) & \hat{\rho}(1) & \cdots & \hat{\rho}(k-2) \\ \hat{\rho}(2) & \hat{\rho}(1) & \hat{\rho}(0) & \cdots & \hat{\rho}(k-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}(k-1) & \hat{\rho}(k-2) & \hat{\rho}(k-3) & \cdots & \hat{\rho}(0) \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \phi_{k3} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \hat{\rho}(3) \\ \vdots \\ \hat{\rho}(k) \end{bmatrix}, \qquad k \geq 1.$$

  asi estimador de $\hat{\alpha}(k)$ es ϕ_{kk}.

3.4 Predicción en Series de Tiempo $MA(q)$

Existen diferentes algoritmos y distintos métodos para la predicción de datos, los tratados aquí utilizan las propiedades que adquiere $\mathcal{L}^2(\Omega, \mathscr{F}, P)$ por ser un espacio de Hilbert. Se enfatiza que estos no son la única forma de obtener resultados óptimos a la hora de analizar series temporales, un buen ejemplo de esto es haciendo uso de la función de máxima verosimilitud (ver [9] pág 123).

Definición 3.4.1. *(Invertibilidad de un proceso $MA(q)$)*
Sea $\{X_t\}$ un proceso $MA(q)$ definido por la ecuación $X_t = \theta_q(B)Z_t$ se dice que es ***invertible*** *si existe una secuencia de constantes $\{\pi_j\}$ tales que*

$$\sum_{j=0}^{\infty} |\pi_j| < \infty \quad y \quad Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \qquad t = 0, \pm 1, \cdots$$

Teorema 3.4.1. *Sea $\{X_t\}$ un proceso $MA(q)$ entonces $\{X_t\}$ es invertible si y solo si $\theta(z) \neq 0$ para todo $z \in \mathbb{C}$ tal que $|z| \leq 1$. Los coeficientes $\{\pi_j\}$ están determinados por la relación*

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{1}{\theta(z)}, \qquad |z| \leq 1$$

(Demostración en [3], pág 87).
Esto quiere decir que un proceso es $MA(q)$ si es posible reescribirlo sólo como retardos de X_t.

Ejemplo 3.4.1. Si $\{X_t\}$ un proceso $MA(1)$, puede llevarse el siguiente esquema de inversión

$$X_t = Z_t - \theta Z_{t-1} \tag{3.17}$$

donde $Z_t = X_t + \theta Z_{t-1}$, además

$$\begin{aligned} Z_{t-1} &= X_{t-1} + \theta Z_{t-2} \\ Z_{t-2} &= X_{t-2} + \theta Z_{t-3} \\ Z_{t-3} &= X_{t-3} + \theta Z_{t-4} \\ &\vdots \end{aligned}$$

si sustituimos en (3.17)

$$\begin{aligned} Z_t &= X_t + \theta Z_{t-1} \\ &= X_t + \theta(X_{t-} + \theta Z_{t-2}) \\ &= X_t + \theta(X_{t-} + \theta(X_{t-2} + \theta Z_{t-3})) \\ &= X_t + \theta X_{t-1} + \theta^2 X_{t-2} + \theta^3 Z_{t-3} \end{aligned}$$

de manera general $Z_t = \sum_{n=1}^{\infty} \theta^n X_{t-n}$ para garantizar la invertibilidad del proceso se debe cumplir además que $\sum_{n=1}^{\infty} |\theta^n| < \infty$, lo cual tiene sentido sí $|\theta| < 1$.

3.4.1 La Mejor Predicción Lineal en $\mathcal{L}^2(\Omega, \mathscr{F}, P)$

Cuando es posible invertir un proceso de tipo $MA(q)$ se pueden utilizar los métodos de predicción en pasos, truncando la suma para un valor finito.

- **Ecuaciones de predicción de un-paso**
  Sea $\{X_t = 0 \pm 1 \cdots\}$ un proceso estacionario sobre $\mathcal{L}^2(\Omega\mathscr{F}, P)$ con media cero y función de autocovarianza $\gamma_X(\cdot)$, encontrar la combinación

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \phi_{nj} X_{n+1-j} \tag{3.18}$$

  la cual es la mejor aproximación de X_{n+1} en el sentido de que $\|X_{n+1} - \hat{X}_{n+1}\|^2$ es mínimo por la observación (2) es lo mismo que minimizar $E\left|X_{n+1} - \sum_{j=1}^{n} \phi_{nj} X_{n+1-j}\right|^2$.
  Si se define $\mathscr{M} = \{X_1, \cdots, X_n\}$ se denota $\hat{X}_{n+1} = P_{\mathscr{M}} X_{n+1}$, por el ejemplo (2.6.1)

$$< X_{n+1} - \sum_{j=1}^{n} \phi_{nj} X_{n+1-j}, X_{n+1-i} > = 0 \quad para \quad i = n, n-1 \cdots, 1$$

$$\sum_{j=1}^{n} \phi_{nj} < X_{n+1-j} X_{n+1-i} > = < X_{n+1}, X_{n+1-i} >$$

  definido el producto interior como en la proposición (2.5.1)

$$\sum_{j=1}^{n} \phi_{nj} E[X_{n+1-j} X_{n+1-i}] = E[X_{n+1}, X_{n+1-i}]$$

$$\sum_{j=1}^{n} \phi_{nj} \gamma_X(i-j) = \gamma(i)$$

  esta igualdad puede ser resuelta mediante un sistema de ecuaciones lineales

$$\Gamma_n \phi_n = \gamma_X(n) \tag{3.19}$$

  donde $\gamma_X(n) = (\gamma(1), \gamma(2), \cdots, \gamma(n))$, $\phi_n = (\phi_{n1}, \phi_{n2}, \cdots, \phi_{nn})'$ y $\Gamma_n = [\gamma_X(i-j)]_{i,j=1}^{n}$ el teorema de la proyección garantiza la existencia de al menos una solución para ϕ_n.

Observación 5. Γ_n es llamada la matriz de covarianzas sus entradas son de la forma

$$\Gamma_n = \begin{bmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(n-2) \\ \gamma(2) & \gamma(1) & \gamma(0) & \cdots & \gamma(n-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \gamma(n-3) & \cdots & \gamma(0) \end{bmatrix}$$

esta matriz es simétrica.

Proposición 3.4.1. Si $\gamma_X(0) > 0$ y $\lim_{h\to\infty} = 0$ entonces la matriz de covarianzas Γ_n, de $(X_1, X_2, \cdots X_n)$ es no singular para n.

(Demostración en [3], pág 167)
Un resultado importante de la proposición anterior garantiza que la mejor predicción $\hat{X}_{n+1}$ es X_{n+1} en términos de $\{X_1, \cdots, X_n\}$ es

$$\hat{X}_{n+1} = \sum_{i=1}^{n} \phi_{nj} X_{n+1-i}$$

donde $\phi_n = (\phi_{n1}, \phi_{n2}, \cdots, \phi_{nn})' = \Gamma_n^{-1}\gamma_X(n)$, siendo Γ_n^{-1} la matriz inversa de covarianzas y $\gamma_X(n) = (\gamma(1), \cdots, \gamma(n))$.

Corolario 3.4.1. El error cuadrático medio de la predicción lineal es $\nu_n = \gamma_X(0) - \gamma_X(n)'\Gamma^{-1}\gamma_X(n)$.

Demostración. Sean el vectores $X = (X_1, \cdots, X_n)$ y $\phi_n = (\phi_{n1}, \cdots \phi_{nn})$

$$\begin{aligned} \nu_n &= E[(X_{n+1} - \hat{X}_{n+1})^2] \\ &= E[X_{n+1}^2] + E[\hat{X}_{n+1}^2] - 2E[\hat{X}_{n+1}X_{n+1}] \\ &= \gamma_X(0) + \phi_n' E[XX']\phi_n - 2E[X_{n+1}X']\phi_n \end{aligned}$$

utilizando la ecuación (3.19)

$$\begin{aligned} \nu_n &= \gamma_X(0) + \phi_n'\Gamma_n\phi_n - 2\gamma_X(n)'\Gamma^{-1}\gamma_X(n) \\ &= \gamma_X(0) + \gamma_X(n)'\Gamma_n^{-1}\Gamma_n\Gamma_n^{-1}\gamma_X(n) - 2\gamma_X(n)'\Gamma^{-1}\gamma_X(n) \\ &= \gamma_X(0) + \gamma_X(n)'\Gamma_n^{-1}\gamma_X(n) - 2\gamma_X(n)'\Gamma^{-1}\gamma_X(n) \\ &= \gamma_X(0) - \gamma_X(n)'\Gamma^{-1}\gamma_X(n) \end{aligned}$$

∎

- **Ecuaciones de predicción de h-paso,** $h \geq 1$
La mejor predicción de X_{n+h} en términos de $X_1, \cdots, X_n$ para cualquier $h \geq 1$ puede ser encontrado de la misma forma que $\hat{X}_{n+1}$. Entonces

$$P_{\mathscr{M}_n} = \phi_{n1}^{(h)} X_n + \cdots + \phi_{nn}^{(h)} X_1, \quad n, h \geq 1, \tag{3.20}$$

donde $\phi_n^{(h)} = (\phi_{n1}^{(h)}, \cdots, \phi_{nn}^{(h)})'$ es cualquier solución (única sí Γ_n es no singular) de

$$\Gamma_n \phi_n^{(h)} = \gamma_n^{(h)}, \tag{3.21}$$

luego

$$\begin{aligned}\Gamma_n =& Cov(X, X)\\ \gamma_n^{(h)} =& Cov(X_{n+h}, X)\\ =& (\gamma(h), \gamma(h+1), \cdots, \gamma(n+h-1))'\end{aligned}$$

además, $E[(X_{n+h} - \hat{X}_{n+h})^2] = \gamma_X(0) - \phi^{(h)\prime}\gamma_n^{(h)}$.

Definición 3.4.2. *(El operador predicción)*
Para las variables aleatorias $Y, Z_1, \cdots, Z_n$, define la mejor predicción de Y dado $Z = (Z_1, \cdots, Z_n)'$ como el operador $P(\cdot|Z)$ aplicado a Y:

$$P(Y|Z) = \mu_Y + \phi'(Z - \mu_Z) \qquad \Gamma\phi = \gamma,$$

donde $\gamma = Cov(Y, Z)$, $\Gamma = Cov(Z, Z)$.

las principales propiedades de este operador son:

1. $E[Y - P(Y|Z)] = 0$, $E[(Y - P(Y|Z))Z] = 0$.
2. $E[(Y - P(Y|Z))^2] = \sigma_Y^2 - \phi'\gamma$.
3. $P(\alpha_0 + \sum_{i=1}^{n} Y_n|Z) = \alpha_0 + \sum_{i=1}^{n} P(Y_n|Z)$.
4. $P(Z_i|Z) = Z_i$.
5. $P(Y|Z) = E[Y]$ si $\gamma = 0$.

Ejemplo 3.4.2. (Predicción de un proceso invertible $MA(q)$)
Para $X_t = \sum_{i=1}^{q} Z_{t-i} + Z_t$, invertible existen $\{\pi_i\}$ tales que $X_t = -\sum_{i=1}^{\infty} \pi_i X_{t-i} + Z_t$

$$\begin{aligned}\hat{X}_{n+1} &= P(X_{n+1}|X_1, \cdots, X_n)\\ &= P\Big(-\sum_{i=1}^{\infty} \pi_i X_{n+1-i} + Z_{n+1}|X_1, \cdots, X_n\Big)\\ &= -\sum_{i=1}^{\infty} \pi_i P(X_{n+1-i}|X_1, \cdots, X_n)\\ &= -\sum_{i=1}^{n} \pi_i X_{n+1-i} - \sum_{i=n+1}^{\infty} \pi_i P(X_{n+1-i}|X_1, \cdots, X_n)\end{aligned}$$

3.4.2 Estimación de Parámetros con el Algoritmo de Innovaciones

Para un proceso que pueda mostrar comportamiento $MA(q)$ la estimación de los parámetros del polinomio $\theta_q(B)$ pueden ser hallados con el Algoritmo de Innovaciones, el cual utiliza la recursión.

El algoritmo consiste en encontrar la proyección $P_{\mathscr{M}_n} X_{n+1} = \hat{X}_{n+1}$ donde $\mathscr{M}_n = \overline{sp}\{X_1 - \hat{X}_1, X_2 - \hat{X}_2, \cdots, X_n - \hat{X}_n\}$ así que

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}) \tag{3.22}$$

Proposición 3.4.2. (El Algoritmo de Innovación)
Sí $\{X_t\}$ tiene media cero y $E[X_i X_j] = \kappa(i,j)$, donde la matriz $[(i,j)]_{i,j=1}^{n}$ es no singular para cada $n = 1, 2 \cdots$, entonces la predicción en un paso de $\hat{X}_{n+1}$, $n \geq 0$, y su error medio cuadrático ν_n, $n \geq$, están dados por

$$\hat{X}_{n+1} = \begin{cases} 0 & si \quad n = 0 \\ \displaystyle\sum_{j=1}^{n} \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}) & si \quad n \geq 1 \end{cases} \tag{3.23}$$

y

$$\hat{X}_{n+1} = \begin{cases} \nu_0 = \kappa(1,1), \\ \theta_{n,n-k} = \nu_k^{-1}\left(\kappa(n+1,k+1) - \displaystyle\sum_{j=0}^{k-1} \theta_{k,k-j}\theta_{n,n-j}\nu_j\right), \quad k = 0, 1 \cdots, n-1 \\ \nu_n = \kappa(n+1,n+1) - \displaystyle\sum_{j=0}^{n-1} \theta_{n,n-j}^2 \nu_j. \end{cases} \tag{3.24}$$

(se obtienen de forma recursiva y en orden los parámetros $\nu_0; \theta_{11}, \nu_1; \theta_{22}, \theta_{21}, \nu_2; \theta_{33}, \theta_{32}\theta_{31}; \cdots$)

(Demostración [3] pág 172)

Ejemplo 3.4.3. Sea X_t una serie con comportamiento $MA(q)$

$$X_t = Z_t + \theta_{q1}Z_{t-1} + \theta_{q2}Z_{t-2} + \cdots + \theta_{qq}Z_{t-q}$$

por la ecuación (3.12)

$$\begin{aligned} X_t &= X_t - \hat{X}_t + \theta_{q1}(X_{t-1} - \hat{X}_{t-1}) + \theta_{q2}(X_{t-2} - \hat{X}_{t-2}) + \cdots + \theta_{qq}(X_{t-q} - \hat{X}_{t-q}) \\ \hat{X}_t &= \theta_{q1}(X_{t-1} - \hat{X}_{t-1}) + \theta_{q2}(X_{t-2} - \hat{X}_{t-2}) + \cdots + \theta_{qq}(X_{t-q} - \hat{X}_{t-q}) \\ \hat{X}_t &= \sum_{j=1}^{q} \theta_{qj}(X_{q+1-j} - \hat{X}_{q+1-j}) \end{aligned}$$

para estimar los parámetros $\theta_{qq} = (\theta_{q1}, \theta_{q2}, \cdots, \theta_{qq})$ se hace uso del Algoritmo de Innovaciones.

El algoritmo de innovaciones además de brindar un método para obtener parámetros, también puede ser utilizado para hacer pronósticos.

Ejemplo 3.4.4. *(Algoritmo de Innovaciones para un proceso $MA(1)$)*
Sea $\{X_t\}$ un proceso $MA(1)$, al ser estacionario su media es cero y además

$$\gamma_X(0) = Cov(X_1, X_1) = E[X_1^2] - \mu_X^2 = \kappa(1,1) = \kappa(n+1, n+1)$$

también se tiene que

$$\gamma_X(n-k) = Cov(X_{n+1}, X_{k+1}) = E[X_{n+1}X_{k+1}] = \kappa(n+1, k+1)$$

con $k = 0, 1 \cdots, n-1$. La ecuación (3.24) queda determinada así

$$\hat{X}_{n+1} = \begin{cases} \nu_0 = \gamma_X(0), \\ \theta_{n,n-k} = \nu_k^{-1}\left(\gamma_X(n-k) - \sum_{j=0}^{k-1}\theta_{k,k-j}\theta_{n,n-j}\nu_j\right), \\ \nu_n = \gamma_X(0) - \sum_{j=0}^{n-1}\theta_{n,n-j}^2\nu_j. \end{cases}$$

Para un proceso $MA(1)$, por el ejemplo (3.2.4), $\gamma_X(0) = \sigma^2(1+\theta^2), \gamma_X(1) = \theta\sigma^2$.

$$\begin{aligned} &\theta_{11} = \frac{\gamma_X(1)}{\nu_0} = \frac{\theta}{1+\theta^2} \\ &\theta_{22} = 0, && \theta_{21} = \frac{\gamma_X(1)}{\nu_1}; \\ &\theta_{33} = 0, && \theta_{32} = 0, && \theta_{31} = \frac{\gamma_X(1)}{\nu_2}; \\ &\cdots \end{aligned}$$

ya que $\gamma_X(n-k) \neq 0$ sólo para $k = n-1$, únicamente $\theta_{n1} \neq 0$. Así para un proceso $MA(1)$

$$X_t = Z_t + \theta Z_{t-1}$$

el mejor pronóstico lineal en representación del algoritmo de innovaciones es

$$\hat{X}_{n+1} = \theta_{n1}(X_n - \hat{X}_n).$$

de manera más general para un proceso $MA(q)$, se tiene que $\theta_{ni} = 0$ para $i > q$.

4 Aplicación

Con ayuda del software R y el uso de algunas funciones de este se mostrará el comportamiento de una serie con datos correspondiente a los precios históricos diarios de las acciones de ecopetrol comprendidas desde el 1 de Enero de 2017 hasta el 30 de junio de 2017. (Datos tomados de [11])

1370	1400	1385	1340	1320	1340	1380	1365	1430	1375	1365
1390	1405	1370	1335	1325	1350	1375	1385	1425	1395	1385
1400	1400	1360	1330	1310	1350	1380	1385	1420	1370	
1415	1410	1365	1335	1315	1350	1375	1390	1405	1375	
1415	1400	1380	1330	1320	1355	1375	1380	1390	1350	
1390	1395	1375	1325	1325	1375	1365	1450	1370	1365	
1415	1395	1375	1315	1325	1365	1385	1485	1355	1350	
1425	1390	1370	1335	1325	1375	1360	1485	1340	1345	
1405	1385	1380	1335	1330	1390	1360	1455	1375	1350	
1400	1395	1360	1330	1325	1395	1355	1420	1400	1355	
1405	1395	1355	1330	1320	1400	1355	1440	1400	1355	
1400	1390	1360	1310	1325	1380	1355	1440	1365	1355	

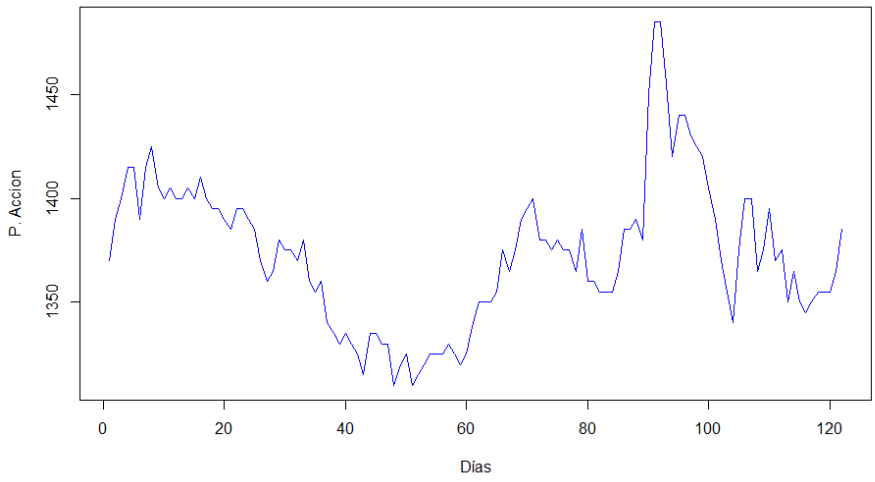


Figura 4.1: Primeros 123 días, precio de acciones Ecopetrol 2017

```
x=c(1370,1390,1400,1415,1415,1390,1415,1425,1405,1400,
1405,1400,1400,1405,1400,1410,1400,1395,1395,1390,1385,
1395,1395,1390,1385,1370,1360,1365,1380,1375,1375,1370,
1380,1360,1355,1360,1340,1335,1330,1335,1330,1325,1315,
1335,1335,1330,1330,1310,1320,1325,1310,1315,1320,1325,
1325,1325,1330,1325,1320,1325,1340,1350,1350,1350,1355,
1375,1365,1375,1390,1395,1400,1380,1380,1375,1380,1375,
1375,1365,1385,1360,1360,1355,1355,1355,1365,1385,1385,
1390,1380,1450,1485,1485,1455,1420,1440,1440,1430,1425,
1420,1405,1390,1370,1355,1340,1375,1400,1400,1365,1375,
1395,1370,1375,1350,1365,1350,1345,1350,1355,1355,1355,
1365,1385)
app=ts(x,start=1)
```

esta serie no es estacionaria así que se le aplicará el operador diferencia dos veces y se graficará su función de autocorrrelación simple y parcial.

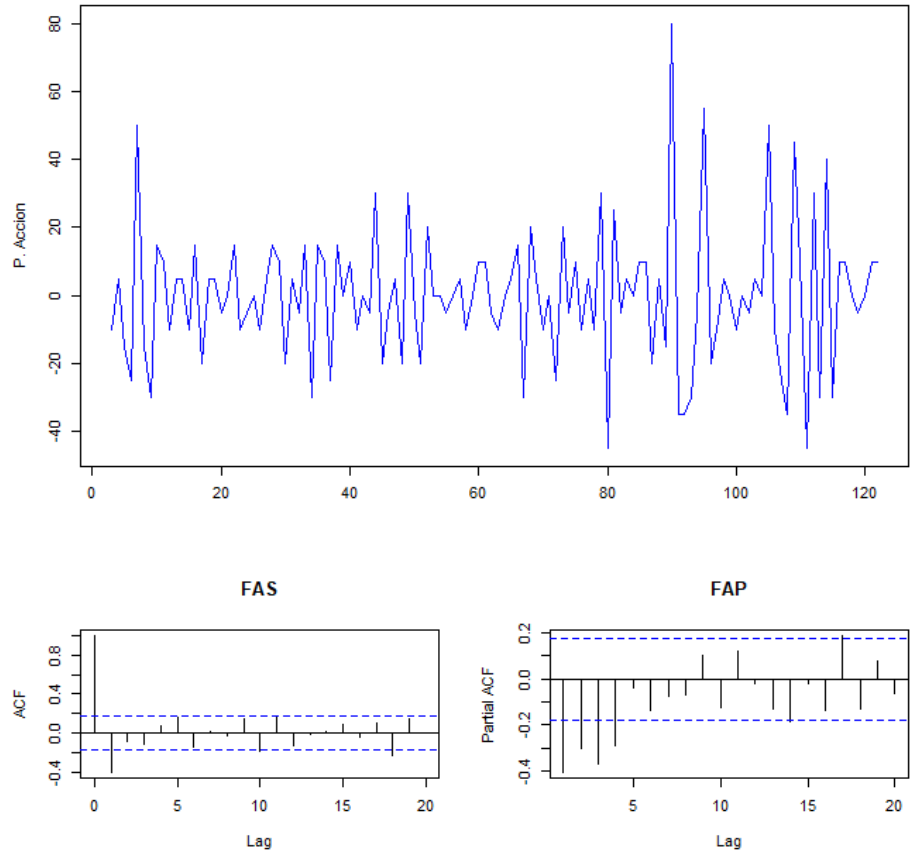


Figura 4.2

```
plot(app,col="4",ylab='P. Accion', xlab='D as ')
d=diff(diff(app))
layout(matrix(c(1, 1,
                1, 1,
                2, 3 ), nrow=3, byrow=TRUE))
layout.show(n=3)
plot(d, ylab='P. Accion', xlab='',type='l',col="4")
acf(d, main="FAS",20)
pacf(d, main="FAP",20)
```

Según el correlograma de la función de autocorrelación simple, es posible que este sea un modelo $MA(1)$, haciendo uso de la función $ia(x,q,m)$ de la librería *itsmr*, se establece el parámetro θ con el algoritmo de innovaciones.

```
library(itsmr)
ia(d, 1,1)
phi
[1] 0
theta
[1] -0.4050104
sigma2
[1] 302.5718
aicc
[1] 1030.305
se.phi
[1] 0
se.theta
[1] 0.09128709
```

El algoritmo nos apunta que los parámetros son $\theta = -0{,}4050104$, $\sigma^2 = 302{,}5718$ la cual es la varianza del ruido blanco Z_t. Hallados los parámetros podemos simular el proceso $MA(1)$ $X_t = Z_t - 0{,}4050104 Z_{t-1}$ con el siguiente código

```
RB = arima.sim(list(order=c(0,0,1),
ma= -0.4050104), n=123, sd=sqrt( 302.5718))
layout(matrix(c(1, 1,
                1, 1,
                2, 3 ), nrow=3, byrow=TRUE))
layout.show(n=3)
plot(RB, ylab='MA(1)', xlab='',type='l',col="3")
acf(RB, main="FAS",20)
pacf(RB, main="FAP",20)
```

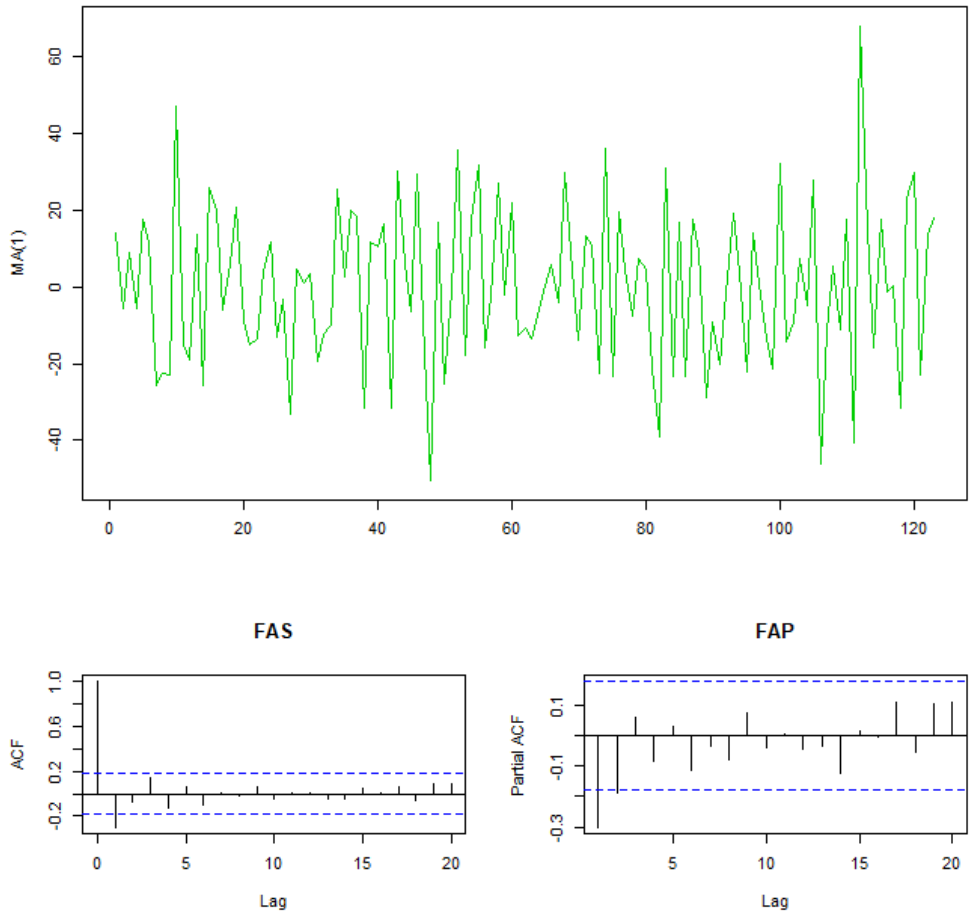


Figura 4.3

es evidente la similitud en el comportamiento de las series de las figuras (4.2) y (4.4), además de sus funciones de autocorrelación simple y parcial.

```
y=c(1370,1390,1400,1415,1415,1390,1415,1425,1405,1400,1405,
1400,1400,1405,1400,1410,1400,1395,1395,1390,1385,1395,1395,
1390,1385,1370,1360,1365,1380,1375,1375,1370,1380,1360,1355,
1360,1340,1335,1330,1335,1330,1325,1315,1335,1335,1330,1330,
1310,1320,1325,1310,1315,1320,1325,1325,1325,1330,1325,1320,
1325,1340,1350,1350,1350,1355,1375,1365,1375,1390,1395,1400,
1380,1380,1375,1380,1375,1375,1365,1385,1360,1360,1355,1355,
1355,1365,1385,1385,1390,1380,1450,1485,1485,1455,1420,1440,
1440,1430,1425,1420,1405,1390,1370,1355,1340,1375,1400,1400,
1365,1375,1395,1370,1375,1350,1365,1350,1345,1350,1355,1355,
1355,1365,1385,1395,1365,1380,1355,1355,1375,1390,1395,1400)
```

Se utiliza la función $forecast(object)$ de la librería $'forecast'$ la cual nos muestra

un intervalo dónde se concentra el pronóstico para los próximos 9 días, estos son contrastados con los datos reales.

```
appy=ts(y,start=1)
plot(app,ylab='P. Accion', xlab='Dias',
ylim=c(1300, 1490),xlim=c(0, 132))
plot(appy,col="2",ylab='P. Accion', xlab='Dias',
ylim=c(1300, 1490),xlim=c(0, 132))
plot(forecast(app, h=9),main="Pronostico",
ylim=c(1300, 1490), xlim=c(0, 132))
```

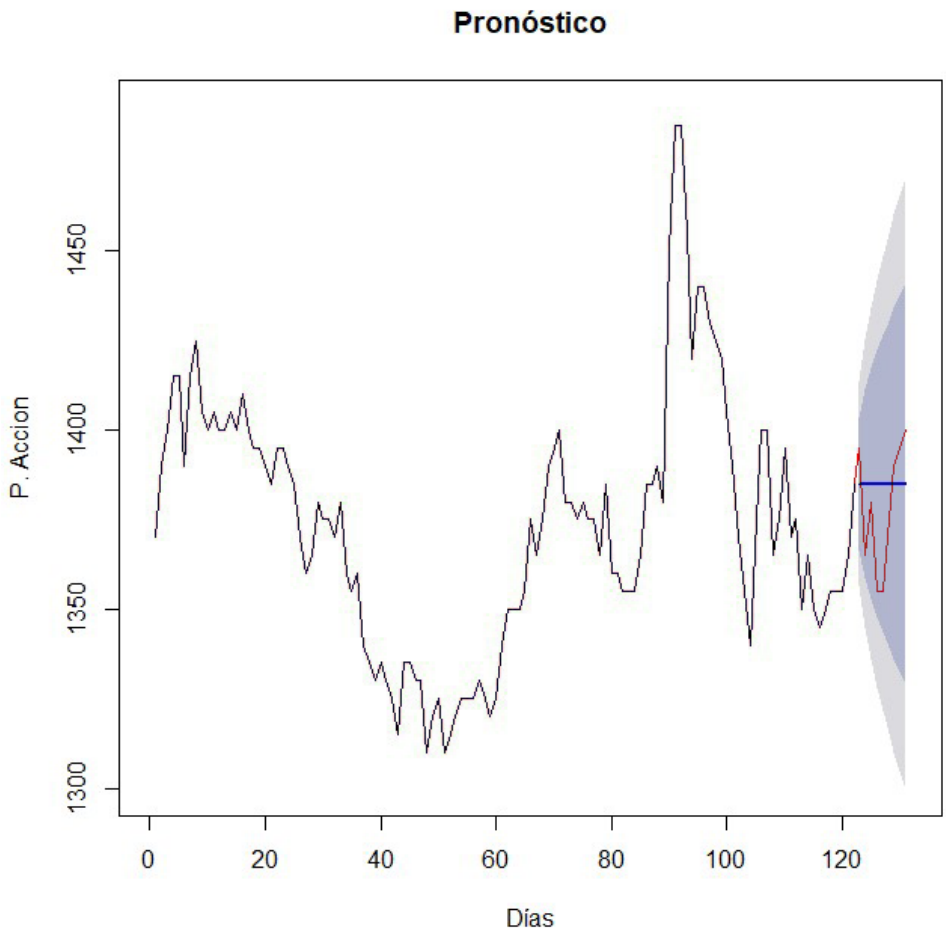


Figura 4.4

esto comprueba la efectividad de la función $forecast(object)$ ya que los datos reales se encuentran en el intervalo.

5 Conclusiones

- Debido a que $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ es un espacio de Hilbert, la ortogonalidad y el teorema de la proyección juegan un papel fundamental en los pronósticos, ya que son los conceptos matemáticos que dan sustento a la dependencia que pueda existir entre las variables aleatorias de una serie temporal.
- Al momento de determinar sí una serie es de tipo $MA(q)$ su función de autocovarianzas ofrece más información que su función de autocorrelación parcial, puesto que al estar representada en una serie Z_t de tipo ruido blanco, se puede presentar como una suma de innovaciones, estas son restas de una variable aleatoria con su predicción respecto a todos las variables aleatorias anteriores.
- Es indispensable que una serie para ser pronosticada deba ser estacionaria, esto garantiza dependencia ya que si no posee esta característica es posible estar trabajando en una serie completamente aleatoria, los operadores rezago y diferencia logran eliminar la aleatoriedad que pueda existir, puesto que se encargan de relacionar los datos.
- Se pueden obtener diferentes pronósticos para una misma serie dependiendo el algoritmo que se trabaje al momento de determinar sus parámetros, aunque puedan variar los datos estos están comprendidos por un intervalo de confianza dados por los errores de estimación de los parámetros.

Bibliografía

[1] BLANCO L., *Probabilidad*. Bogotá, UNILIBLOS, 2004.

[2] BROCKWELL, P.J., y DAVIS, R.A., *Time Series: Theory and Methods*. NewYork, Springer, 1991.

[3] BROCKWELL, P.J., y DAVIS, R.A., *Introduction to Time Series and Forecasting*. NewYork, Springer, 2002.

[4] CHUNG K.L., *A Course In Probabiility Theory*. New York, Academic Press Inc, 1995.

[5] GIRALDO, N.D.,*NOTAS DE CLASE Series de Tiempo con R* . Medellín, Universidad Nacional De Colombia, 2006.

[6] GRIGORIU M., *Stochastic Calculus: Applications in Science and Engineering*. NewYork, Springer+Business Media, LLC, 2002.

[7] GRISEL M. BRITOS, *Tópicos en Series De Tiempo*. Córdoba, 2012.

[8] RESNICK S.I., *A Probability Path*. Boston, Birkhauser, 1999.

[9] ROBERT, H.S., y DAVID, S.S., *Time Series Analysis y Its Applications*. California, Springer, 2011.

[10] http://datos.bancomundial.org/indicador/EN.ATM.CO2E.PC

[11] http://www.ecopetrol.com.co/wps/portal/es/ecopetrol-web/historico-de-la-accion