



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICA

Análisis de las variaciones porcentuales del Índice Nacional de Precios al Consumidor mediante el uso de Series Temporales

Trabajo Especial de Grado presentado ante la ilustre Universidad Central de Venezuela por la **Br. Rita Sleiman el Zakhem** para optar al título de Licenciado en Matemática.

Tutor: Dr(a). Mairene Colina.

Caracas, Venezuela
Julio, 2013.

Nosotros, los abajo firmantes, designados por la Universidad Central de Venezuela como integrantes del Jurado Examinador del Trabajo Especial de Grado titulado "**Análisis de las variaciones porcentuales del Índice Nacional de Precios al Consumidor mediante el uso de Series Temporales**", presentado por la **Br. Rita Sleiman el Zakhem**, titular de la Cédula de Identidad **20.413.120**, certificamos que este trabajo cumple con los requisitos exigidos por nuestra Magna Casa de Estudios para optar al título de **Licenciado en Matemática**.

Dra. Mairene Colina
Tutor

Dr. José Benito Hernández
Jurado

Msc. Angie Pineda
Jurado

Agradecimientos

Por sobre todas las cosas quiero darle gracias a Dios, por siempre haber guiado mi camino.

A mis padres, por su esfuerzo y constancia en ayudarme siempre a obtener lo que quiero.

A la Universidad Central de Venezuela por existir y permitirme crecer en sus pasillos, y aprender en sus aulas.

A la profesora Mairene Colina por su paciencia y dedicación.

A mi amigos, Greisi, Barbara, Yeffirson, y a todos esos angelitos que me acompañaron día a día en esta aventura.

Dedicatoria

A Dios por haberme guiado por el camino de la sabiduría para alcanzar esta meta.

A mi mamá por su amistad, la cual me acompaña cada día y ha sido la herramienta clave para alcanzar todos mis objetivos en la vida, pero más que nada por estar presente.

A mi papá que ha sido la roca sobre la cual no solo mi carrera sino mi vida se ha apoyado.

A mi hermana que aun desde la distancia logra día a día hacer sentir su amor.

A mi hermana del alma Greisi, que sólo con su presencia ha logrado que mi vida sea todo lo que he querido.

A mis amigos, Barbara y Yeffirson, por haber compartido durante los años de mi carrera muchos momentos que marcaron mi vida, gracias a su apoyo.

Y a mis amigos de WorldMUN, que se han convertido en mi segunda familia.

Introducción

Con las variaciones mensuales habituales presentadas por bienes y servicios en una familia promedio, era requerido un barómetro o medidor del comportamiento de la economía, que permitiese formular medidas de política social, ajustar salarios, prestaciones de seguridad y asistencia social, así nace el Índice de Precios al Consumidor (IPC), que es un indicador estadístico que mide el cambio promedio experimentado por los precios de una canasta de bienes y servicios representativa del consumo de los hogares en un período determinado, este desempeña una función importante en el proceso de ajuste de los ingresos.

Ahora bien, si hacemos referencia al IPC como objeto de estudio en cuanto al comportamiento de precios y servicios, es importante aclarar que la descripción de este indicador ha presentado una ardua labor para ciencias como la economía y la matemática, puesto que está conformado por una serie de variables independientes que inciden directamente en su caracterización. Para la matemática, el estudio de indicadores de esta naturaleza recibe el nombre de "Proceso Estocástico", un ejemplo claro es una serie de tiempo.

Es por esto, que una serie temporal empleada en datos recolectados del estudio del Índice de Precios al Consumidor, muestra información respecto al comportamiento de la información estimada por rubros de consumo básico como lo son los alimentos, el alquiler de viviendas y los servicios médicos. Dentro de este proyecto se analizarán con los datos recogidos por el Banco Central de Venezuela (BCV) en los últimos doce años, permitiendo, a través del análisis y el establecimiento de modelos, la predicción de las próximas variaciones porcentuales de los rubros estimados.

Así, este trabajo se divide en cuatro capítulos principales, los Capítulos 1 y 2 fundamentan los conceptos básicos de series temporales, su caracterización, propiedades y demás funciones relacionadas que permiten a través de esta herramienta la obtención de información. En el Capítulo 3, se presentan definiciones, antecedentes y cálculo del IPC, así mismo se presenta información pertinente con respecto al indicador específico de la República Bolivariana de Venezuela, es decir, el Índice Nacional de Precios al Consumidor.

Para culminar el Capítulo 4 presenta la aplicación directa del manejo de series temporales a los datos obtenidos del Banco Central de Venezuela, lo que permite a través de manipulación matemática, obtener predicciones para valores futuros, así como dilucidar características particulares de estas series específicas.

CONTENIDO

1 Series Temporales	9
1.1 Preliminares	9
1.2 Medida de Dependencia	12
1.3 Estacionaridad	15
1.4 Estimación de Correlación	20
1.5 Regresión de Series de Tiempo	22
2 Análisis exploratorio de datos	26
2.1 Modelos Autoregresivos y Promedio Móvil	29
2.1.1 Modelos Autoregresivos	29
2.1.2 Modelos de Promedio Móvil	31
2.1.3 Modelos Autoregresivos de Promedio Móvil	32
2.1.4 Otros modelos importantes	33
2.2 Funciones de Autocorrelación y Autocorrelación Parcial	34
2.2.1 Función de Autocorrelación (ACF)	35
2.2.2 Función de Autocorrelación Parcial (PACF)	36
2.3 Pronósticos	37
2.3.1 Pronósticos para un proceso ARMA	37
3 Índices Inflacionarios	40
3.1 Inflación	40
3.2 Índice de Precios al Consumidor	41
3.2.1 Año Base	42
3.3 Índice Nacional de Precios al Consumidor INPC	42
3.3.1 Antecedentes	42
3.3.2 Establecimiento del 2007 como año base en Venezuela	43
3.3.3 Cálculo del INPC	43
3.3.4 Proceso de recaudación de los datos	45

3.3.5	Problemas de Medición del INPC	45
4	Análisis de la Serie del INPC en Venezuela entre los años 2000 y 2012	47
4.1	Serie correspondiente a la variación acumulada del subgrupo "Alimentos" (Enero 2000 - Septiembre 2012)	47
4.2	Serie correspondiente a la variación mensual del subgrupo "Tabaco" (Enero 2008 - Septiembre 2012)	58
4.3	Serie correspondiente a la variación mensual del subgrupo "Alquiler de Viviendas" (Enero 2008 - Septiembre 2012)	66

Capítulo 1

Series Temporales

En este capítulo se estudiarán los conceptos básicos del análisis de Series Temporales, y se introducirán definiciones y propiedades que permitirán a lo largo de este trabajo la manipulación de información y datos para la obtención de predicciones.

1.1 Preliminares

Definición 1. Una terna (Ω, A, P) , formada por un espacio muestral Ω , una familia A de eventos y una probabilidad P se llama **espacio de probabilidad**.

El problema de cómo definir la función P , o sea, de cómo asignar una probabilidad a cada evento, debe ser resuelto de acuerdo a las condiciones concretas de cada experimento aleatorio en consideración. Considerando la definición anterior se introduce el siguiente concepto:

Definición 2. Un vector $X = (X_1, X_2, \dots, X_n)$ donde X_1, X_2, \dots, X_n son variables aleatorias definidas en un espacio de probabilidad (Ω, F, P) es llamado **vector aleatorio**.

Así como se define la distribución de una variable aleatoria también se define la distribución de un vector aleatorio.

Definición 3. (Función de distribución conjunta) La función de distribución $F_{X_1, \dots, X_n}(x)$ de un vector aleatorio (X_1, X_2, \dots, X_n) definida en un espacio de probabilidad (Ω, A, P) , está definida por:

$$\begin{aligned}
F_X(x) &= F_{X_1, \dots, X_n}(x) \\
&= P(w \in \Omega : X_1(w) \leq x_1, \dots, X_n(w) \leq x_n) \\
&= P(X_1 \leq x_1, \dots, X_n \leq x_n),
\end{aligned}$$

donde $x = (x_1, \dots, x_n) \in \mathbb{R}^n$

Si la distribución de un vector aleatorio X tiene *densidad* f_X , uno puede representar la función de distribución F_X de X como:

$$\begin{aligned}
F_X(x_1, \dots, x_n) &= \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_X(y_1, \dots, y_n) dy_1 \dots dy_n, \\
(x_1, \dots, x_n) &\in \mathbb{R}^n,
\end{aligned}$$

donde para todo $x \in \mathbb{R}^n$, la densidad es una función que satisface

$$f_X(x) \geq 0,$$

y se verifica que:

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_X(y_1, \dots, y_n) dy_1 \dots dy_n = 1.$$

Si un vector X tiene densidad f_X , todas sus componentes X_i , los vectores de pares (X_i, X_j) , triples (X_i, X_j, X_k) , etc., tienen una función de densidad. Estas son llamadas densidades marginales.

Una de las definiciones que permite el cálculo de propiedades y la modelación de series temporales es la esperanza o media de un vector aleatorio.

Definición 4. La *esperanza o media* de un vector aleatorio X está dado por:

$$\mu_X = E[X] = (E[X_1], \dots, E[X_n]) = (\mu_{X_1}, \dots, \mu_{X_n}).$$

La matriz de **covarianza** de X está definida como:

$$\Sigma_X = (\text{cov}(X_i, X_j) : i, j = 1, \dots, n),$$

donde

$$\text{cov}(X_i, X_j) = E[(X_i - \mu_{X_i})(X_j - \mu_{X_j})] = E(X_i X_j) - \mu_{X_i} \mu_{X_j},$$

es la covarianza de X_i y X_j . Nótese que la $\text{cov}(X_i, X_i) = \text{Var}(X_i) := \sigma_{X_i}^2$.

Definición 5. (Independencia) Las variables aleatorias X_1, \dots, X_n son independientes si, para toda colección de índices $1 \leq i_1 < \dots < i_k \leq n$ y enteros $1 \leq k \leq n$ y todo subconjunto B_1, \dots, B_n de \mathbb{R} ,

$$P(X_{i_1} \in B_{i_1}, \dots, X_{i_k} \in B_{i_k}) = P(X_{i_1} \in B_{i_1}) \dots P(X_{i_k} \in B_{i_k}).$$

Esto significa que los eventos $X_1 \in B_1 \dots X_k \in B_k$ son independientes.

Así podemos considerar que las variables aleatorias $\{X_1\}, \dots, \{X_n\}$ son independientes si y solo si su función de distribución conjunta puede escribirse como sigue:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \dots F_{X_n}(x_n),$$

$(x_1, \dots, x_n) \in \mathbb{R}^n$.

Si el vector aleatorio $\mathbf{X} = \{X_1, \dots, X_n\}$ tiene densidad f_X , donde f_{X_i} representa la función de densidad de la variable aleatoria X_i , entonces X_1, \dots, X_n son independientes si y solo si:

$$f_X(x_1, \dots, x_n) = f_{X_1}(x_1) \dots f_{X_n}(x_n),$$

para $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Una consecuencia importante de la independencia de variables aleatorias es la siguiente propiedad:

Si X_1, \dots, X_n son independientes, entonces para funciones reales cualesquiera g_1, \dots, g_n ,

$$E[g_1(X_1) \dots g_n(X_n)] = E[g_1(X_1)] \dots E[g_n(X_n)],$$

siempre que las esperanzas consideradas estén bien definidas.

Definición 6. Un **proceso estocástico a tiempo continuo** es una sucesión $(X_t, t \in T) = (X_t(w), t \in T, w \in \Omega, T \subseteq \mathbb{R}^+)$ de variables aleatorias definidas sobre un espacio Ω .

- Para un instante de tiempo t fijo, $w \in \Omega$, $X_t(w)$ es una variable aleatoria.
- Para un w fijo es una función del tiempo $X_t(w), t \in T$. Esta función es llamada trayectoria del proceso.

Un ejemplo importante de un proceso estocástico es una serie de tiempo.

Definición 7. Una **serie de tiempo** es un conjunto de observaciones x_t , cada una registrada a un tiempo específico t .

Definición 8. Un **modelo de series de tiempo** para los datos observados x_t es una especificación de una distribución conjunta (o posiblemente solo de medias y covarianzas) de una sucesión de variables aleatorias $\{X_t\}$ de las cuales $\{x_t\}$ es una realización.

1.2 Medida de Dependencia

Una descripción completa de una serie de tiempo, observada como una colección de n variables aleatorias en puntos de tiempo enteros arbitrarios t_1, \dots, t_n para cada entero positivo n , es proporcionada por la función de distribución conjunta, evaluada como la probabilidad de que los valores de la serie sean conjuntamente menor que n constantes c_1, c_2, \dots, c_n , esto es:

$$F(c_1, \dots, c_n) = P(x_{t_1} \leq c_1, x_{t_2} \leq c_2, \dots, x_{t_n} \leq c_n).$$

Desafortunadamente, la función de distribución multidimensional usualmente no se puede escribir fácilmente a menos que las variables aleatorias tengan distribución normal conjunta, en cuyo caso, la ecuación anterior llega a ser la distribución normal multivariada usual.

Un caso particular en la cual la función de distribución multidimensional es fácil de escribir, será en el caso de variables aleatorias normal estándar independientes e idénticamente distribuidas, para lo cual la función

de distribución se puede expresar como el producto de las distribuciones marginales, es decir,

$$F(c_1, c_2, \dots, c_n) = \prod_{t_1}^n \Phi(c_t)$$

donde,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{z^2}{2}\right\} dz$$

es la función de distribución normal estándar acumulada.

Aunque la función de distribución multidimensional describa los datos completamente, esto es un instrumento poco manejable para mostrar y analizar datos de series de tiempo. La función de distribución debe ser evaluada como una función de n argumentos, entonces cualquier graficación de las correspondientes funciones de densidad multivariante es prácticamente imposible. La función de distribución unidimensional,

$$F_t(x) = P\{x_t \leq x\}$$

o la correspondiente función de densidad unidimensional

$$f_t(x) = \frac{\partial F_t(x)}{\partial x},$$

cuando existen, a menudo son más útiles para determinar si una coordenada en particular de la serie de tiempo tiene una función de densidad conocida, como la distribución normal (gaussiana).

Definición 9. La *función de media* es definida como

$$\mu_{xt} = \mathbb{E}(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx,$$

en caso de que exista, donde \mathbb{E} denota el operador usual de esperanza. Cuando no haya confusión sobre a que serie de tiempo nos referimos escribiremos μ_{X_t} como μ_t .

Lo importante de comprender sobre μ_t consiste en que es una media teórica para la serie de tiempo en un punto particular, donde la media se asume o calcula sobre todos los posibles eventos que podrían haber producido X_t .

Así mismo introducimos la definición de la función de autocovarianza.

Definición 10. *La función de autocovarianza es definida como producto del segundo momento*

$$\gamma_X(s, t) = \mathbb{E}[(X_s - \mu_s)(X_t - \mu_t)],$$

para todo t y s . Cuando no haya confusión en la existencia sobre a que serie nos referimos, escribiremos $\gamma_X(s, t) = \gamma(s, t)$.

Note que $\gamma_X(s, t) = \gamma_X(t, s)$ para todos los puntos s y t . La función de autocovarianza mide la dependencia lineal entre dos puntos de la misma serie en diferentes tiempos. La autocovarianza es el promedio de los productos cruzados relacionado con la desidad conjunta $F(X_s, X_t)$. Es claro que, para $s = t$, la autocovarianza se reduce a la varianza (en el caso infinito), dado que:

$$\gamma_s(t, t) = \mathbb{E}[(x_t - \mu_t)^2].$$

Las definiciones anteriores permiten introducir los siguientes conceptos:

Definición 11. *La función de autocorrelación (ACF) se define como:*

$$\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}}.$$

La ACF mide la predictibilidad lineal de una serie de tiempo t , digamos X_t usando solo el valor X_s . Es fácil de demostrar que $-1 \leq \rho(s, t) \leq 1$ usando la desigualdad de Cauchy-Schwarz. Si podemos predecir X_t exactamente de X_s a través de la relación lineal $X_t = \beta_0 + \beta_1 X_s$ entonces la correlación será 1 cuando $\beta_1 > 0$ y -1 cuando $\beta_1 < 0$.

Definición 12. *La función de covarianza cruzada entre dos series X_t y Y_t se define como*

$$\gamma_{XY}(s, t) = \mathbb{E}[(X_s - \mu_{X_s})(Y_t - \mu_{Y_t})]$$

Definición 13. La **función de correlación cruzada (CCF)** es definida como

$$\rho_{XY}(s, t) = \frac{\gamma_{XY}(s, t)}{\sqrt{\gamma_X(s, s)\gamma_Y(t, t)}}.$$

1.3 Estacionaridad

Definición 14. Una serie de tiempo $\{X_t\}_{t \in T}$ **estrictamente estacionaria** es una serie para la cual el comportamiento probabilístico de cada sucesión de valores $\{X_{t_1}, X_{t_2}, \dots, X_{t_k}\}$ es idéntico a la serie trasladada en el tiempo $\{X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h}\}$. Esto es,

$$P\{X_{t_1} \leq c_1, \dots, X_{t_k} \leq c_k\} = P\{X_{t_1+h} \leq c_1, \dots, X_{t_k+h} \leq c_k\}$$

para todo $k = 1, 2, \dots$, todos los puntos de tiempos t_1, t_2, \dots, t_k y números c_1, c_2, \dots, c_k y todo salto $h = 0, \pm 1, \pm 2, \dots$,

Observaciones

1. Si una serie de tiempo es estrictamente estacionaria, entonces todas las funciones de distribución multivariadas para subconjuntos de variables deben coincidir con sus contra partes en el conjunto trasladado, para todos los valores del parámetro h . Por ejemplo para $k = 1$, la ecuación anterior implica que:

$$P\{x_s \leq c\} = P\{x_t \leq c\}$$

para cada par de puntos s y $t \in T$.

Esta declaración implica, por ejemplo, que si la probabilidad de un valor de una serie de tiempo muestreada cada hora es negativa a la 1 : 00 a.m, la probabilidad a la 10 : 00 a.m. es la misma. Además, si la función de media, μ_t de la serie X_t existe, esta propiedad implica que $\mu_s = \mu_t$ para todo s y t , y por consiguiente μ_t debe ser constante. Note, que un proceso de camino aleatorio con tendencia no es estrictamente estacionario porque la función de media cambia con el tiempo.

2. Cuando $k = 2$, podemos escribir la ecuación anterior como:

$$P\{X_s \leq c_1, X_t \leq c_2\} = P\{X_{s+h} \leq c_1, X_{t+h} \leq c_2\}$$

para cada par de puntos s y t y salto h .

Entonces, si la función de varianza del proceso existe, la ecuación anterior implica que la función de autocovarianza de la serie X_t satisface $\gamma(s, t) = \gamma(s + h, t + h)$ para todos s y t y salto h .

Podemos interpretar este resultado diciendo que la función de autocovarianza del proceso depende sólo de las diferencias de tiempo entre s y t , y no del tiempo actual.

La Propiedad de estricta estacionariedad es muy fuerte para la mayoría de las aplicaciones. En vez de imponer condiciones sobre todas las posibles distribuciones de una serie de tiempo, usaremos una versión más suave que imponga condiciones solo sobre los dos primeros momentos de la serie. Tenemos por lo tanto la siguiente definición:

Definición 15. Una serie de tiempo $\{X_t\}_{t \in T}$ **débilmente estacionaria**, es un proceso de varianza finita tal que:

1. la función de media μ_t , es constante y no depende del tiempo t , y
2. la función de covarianza, $\gamma(s, t)$, depende solo de las diferencias de s y t , $|(s - t)|$.

Por consiguiente, usaremos el término **estacionariedad** para referirnos a estacionariedad débil; si un proceso es estacionario en el sentido estricto usaremos el término estrictamente estacionario.

Es claro de la definición anterior de "estricamente estacionario" que una serie de tiempo estrictamente estacionaria con varianza finita, también es una serie estacionaria. El recíproco no es cierto a menos que impongamos condiciones adicionales. Un importante caso donde estacionariedad implica estricta estacionariedad es el caso de series de tiempo gaussianas.

Ya que la función de media $\mathbb{E}(X_t) = \mu_t$ de una serie de tiempo estacionaria es independiente del tiempo t , escribimos:

$$\mu_t = \mu.$$

Debido a que la función de covarianza de una serie de tiempo estacionaria, $\gamma(s, t)$ en tiempos s y t depende sólo de la diferencia $|(s - t)|$, podemos simplificar la notación. Sea $s = t + h$, donde h representa el tiempo de traslación o salto, entonces:

$$\begin{aligned}\gamma(s, t) &= \mathbb{E}[(X_{t+h} - \mu)(X_t - \mu)] \\ &= \mathbb{E}[(X_h - \mu)(X_0 - \mu)] \\ &= \gamma(h, 0)\end{aligned}$$

no depende del argumento de tiempo t ; asumiendo que $\text{var}(X_t) = \gamma(0, 0) < \infty$. De ahora en adelante, por conveniencia, prescindiremos del segundo argumento de $\gamma(h, 0)$, es decir, la función de covarianza se denotará $\gamma(h)$.

Definición 16. *La función de autocovarianza de una serie de tiempo estacionaria se escribirá como:*

$$\gamma(h) = \mathbb{E}[(X_{t+h} - \mu)(X_t - \mu)].$$

Definición 17. *La función de autocorrelación (ACF) de una serie de tiempo estacionaria será escrita:*

$$\rho(h) = \frac{\gamma(t+h, t)}{\sqrt{\gamma(t+h, t+h)\gamma(t, t)}} = \frac{\gamma(h)}{\gamma(0)}.$$

Propiedades

1. Para el valor en $h = 0$, la función de autocovarianza

$$\gamma(0) = \mathbb{E}[(X_t - \mu)^2]$$

es la varianza de la serie de tiempo; note que la desigualdad de Cauchy-Schwarz implica que $|\gamma(h)| \leq \gamma(0)$.

2. La autocovarianza de una serie estacionaria es simétrica respecto al origen, esto es $\gamma(h) = \gamma(-h)$ para todo h . Esta propiedad se debe a que trasladar la serie por h significa que:

$$\begin{aligned}
 \gamma(h) &= \gamma(t+h, t) \\
 &= \mathbb{E}[(x_{t+h} - \mu)(x_t - \mu)] \\
 &= \mathbb{E}[(x_t - \mu)(x_{t+h} - \mu)] \\
 &= \gamma(t - (t+h)) \\
 &= \gamma(-h).
 \end{aligned}$$

Lo cual muestra como usar la notación para demostrar el resultado.

Definición 18. *Dos series de tiempo $\{X_t\}$ y $\{X_s\}$ se dice que son **conjuntamente estacionarias** si cada una de ellas es estacionaria y la función de correlación cruzada*

$$\gamma_{XY}(h) = \mathbb{E}[(X_{t+h} - \mu_X)(Y_t - \mu_Y)]$$

es una función sólo del salto h .

Definición 19. *La **función de correlación cruzada (CCF)** de dos series conjuntamente estacionarias $\{X_t\}$ y $\{Y_t\}$ se define como*

$$\rho_{XY}(h) = \frac{\gamma_{XY}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}}$$

De nuevo, tenemos el resultado $-1 \leq \rho_{XY}(h) \leq 1$ lo cual permite comparar los valores extremos -1 y 1 cuando vemos la relación entre X_{t+h} y Y_t . La función de correlación cruzada satisface:

$$\rho_{XY}(h) = \rho_{YX}(-h)$$

A continuación se presentarán los conceptos de proceso lineal y proceso gaussiano, con los cuales culmina la sección referente a las Medidas de Dependencia.

Definición 20. Un **proceso lineal** $\{X_t\}$ se define como una combinación lineal de variables aleatorias de ruido blanco $\{w_t\}$, donde $t \in T \subseteq \mathbb{R}^+$, y está dado por

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j w_{t-j},$$

donde los coeficientes $\psi_j \in \mathbb{R}$ satisfacen

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

Para un proceso lineal, podemos demostrar que la función de autocovarianza está dada por:

$$\gamma(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_{j+h} \psi_j$$

para todo $0 \leq h$ y donde σ_w^2 es la varianza del ruido blanco; recuerde que $\gamma(-h) = \gamma(h)$, [1] (Brockwell and Davis, 1996). Finalmente como mencionamos anteriormente, un caso importante en el cual una serie débilmente estacionaria es también estrictamente estacionaria es la serie normal o gaussiana.

Definición 21. Un proceso $\{X_t\}$, se dice que es un **proceso gaussiano** si el k -ésimo vector dimensional $\hat{x} = (x_{t_1}, x_{t_2}, \dots, x_{t_k})'$, para cada conjunto de puntos t_1, t_2, \dots, t_k y cada entero positivo k tiene distribución normal multivariada.

Definiendo $k \times 1$ vector de medias $\hat{\mu} = (u_{t_1}, \dots, u_{t_k})'$ y la $k \times k$ matriz de covarianza positiva como $\Gamma = \{\gamma(t_i, t_j); i, j = 1, \dots, k\}$, la función de densidad normal multivariada se puede escribir como

$$f(\hat{x}) = (2\pi)^{(k/2)} |\Gamma|^{(-1/2)} \exp\left\{-\frac{1}{2}(\hat{x} - \hat{\mu})' \Gamma^{-1} (\hat{x} - \hat{\mu})\right\}$$

donde $|\cdot|$ denota el determinante. Esta distribución forma la base para resolver problemas que envuelven inferencia estadística para series de tiempo. Si una serie de tiempo gaussiana $\{x_t\}$ es débilmente estacionaria, entonces

$\mu_t = \mu$ y $\gamma(t_i, t_j) = \gamma(|t_i - t_j|)$, de modo que el vector $\hat{\mu}$ y la matriz Γ son independientes del tiempo. Este hecho implica que todas las distribuciones finitas, de la serie $\{x_t\}$ dependen sólo del salto de tiempo y no del tiempo actual, y por consiguiente la serie debe ser estrictamente estacionaria.

1.4 Estimación de Correlación

Aunque las funciones teóricas de autocorrelación y correlación cruzada son muy útiles para describir las propiedades de ciertos modelos hipotéticos, la mayoría de los análisis se realizan usando datos muestrales. Esta limitación significa que los puntos muestrales x_1, x_2, \dots, x_n , sólo permiten estimar las funciones de media, autocovarianza y correlación. Desde el punto de vista de la estadística clásica esto plantea un problema porque no se tienen copias distribuidas de X_t que sean o estén disponibles para estimar las funciones de covarianza y correlación.

Definición 22. Sea x_1, \dots, x_n una muestra de una serie de tiempo. La **media muestral** de x_1, \dots, x_n es

$$\bar{x} = \frac{1}{n} \sum_{t=1}^n x_t.$$

La **función de autocovarianza muestral** se define como

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(x_t - \bar{x}),$$

con $\hat{\gamma}(-h) = \hat{\gamma}(h)$ para $h = 0, 1, \dots, n - 1$.

La **función de autocorrelación muestral** se define como:

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

La suma de la función de autocovarianza muestral está restringida hasta $n - h$ porque X_{t+h} no toma valores para $t+h > n$. En el estimador a veces se prefiere dividir por $n - h$ porque la autocovarianza muestral es una función no-negativa definida. La no-negatividad de la función de autocovarianza $\gamma(h)$

es una propiedad importante porque nos asegura que la varianza de combinaciones lineales de valores de la serie de tiempo nunca será negativa y la estimación conserva las propiedades.

La función de Autocorrelación presenta propiedades interesantes cuando las muestras son grandes, y permiten modificar los cálculos realizados anteriormente.

Propiedad 1. Distribución de la ACF para muestras grandes

Bajo ciertas condiciones generales, como que X_t sea idénticamente distribuida con cuarto momento finito, o que X_t sea un ruido blanco gaussiano, entonces para n grande, la ACF muestral $\hat{\rho}_X(h)$ para $h = 1, 2, \dots, H$ donde H es un valor fijo arbitrario, es aproximadamente normal distribuida con media cero y desviación estándar dada por

$$\sigma_{\hat{\rho}(h)} = \frac{1}{\sqrt{n}}.$$

Basándonos en el resultado anterior, podemos obtener un método basto de evaluación de si los picos en $\hat{\rho}(h)$ son significativos, por medio de determinar si los picos observados están fuera del intervalo $\pm 2/\sqrt{n}$ (esto es, mas o menos dos veces el error estándar); para una sucesión de ruido blanco aproximadamente el 95 por ciento de la ACF muestral debe estar entre estos límites.

Definición 23. *El estimador para la función de covarianza cruzada $\gamma_{XY}(h)$ es la **función de covarianza cruzada muestral** definida como:*

$$\hat{\gamma}_{XY}(h) = n^{-1} \sum_{t=1}^{n-h} (X_{t+h} - \bar{X})(Y_t - \bar{Y}),$$

donde $\hat{\gamma}_{XY}(-h) = \hat{\gamma}_{YX}(h)$ determina la función para saltos negativos, y la **función de correlación cruzada muestral** es

$$\hat{\rho}_{XY}(h) = \frac{\hat{\gamma}_{XY}(h)}{\sqrt{\hat{\gamma}_X(0)\hat{\gamma}_Y(0)}}.$$

Propiedad 2. Distribución de la correlación cruzada para muestras grandes con independencia

La distribución de $\hat{\rho}_{XY}(h)$ para muestras grandes es normal con media cero y

$$\sigma_{\hat{\rho}_{XY}} = \frac{1}{\sqrt{n}}$$

si al menos uno de los procesos es un ruido blanco independiente.

1.5 Regresión de Series de Tiempo

Los modelos de regresión son importantes para modelos en el dominio de tiempo y de frecuencia. La idea principal depende de poder expresar una serie respuesta $\{X_t\}$ como una combinación lineal de entradas $z_{t_1}, z_{t_2}, \dots, z_{t_q}$. La estimación de los coeficientes $\beta_1, \beta_2, \dots, \beta_q$ de la combinación por mínimos cuadrados proporciona un método para modelar $\{X_t\}$ en términos de las entradas.

Supongamos que tenemos x_t , para $t = 1, 2, \dots, n$ formada por una colección de series temporales independientes $z_{t_1}, z_{t_2}, \dots, z_{t_q}$, donde se consideran primero que las entradas son fijas y conocidas. Esta relación se puede expresar como:

$$x_t = \beta_1 z_{t_1} + \beta_2 z_{t_2} + \dots + \beta_q z_{t_q} + w_t, \quad (1.1)$$

donde β_1, \dots, β_q son los coeficientes de regresión fijos y desconocidos, $\{w_t\}_t$ es un error aleatorio o un proceso de ruido blanco consistente de variables normales iguales e idénticamente distribuidas con media cero y varianza σ_w^2 .

El modelo lineal descrito en (1.1) se puede escribir de forma más general con los vectores columna $z_t = (z_{t_1}, \dots, z_{t_q})'$ y $\beta = (\beta_1, \dots, \beta_q)'$, donde $'$ denota la traspuesta, usando esta representación vectorial (1.1) se escribe como,

$$x_t = \beta' z_t + w_t \quad (1.2)$$

donde w_t es normal, igual e idénticamente distribuida con media cero y varianza σ_w^2 .

Es natural considerar la estimación de los coeficientes del vector β minimizando la suma residual de cuadrados

$$RSS = \sum_{t=1}^n (x_t - \beta' z_t)^2 \quad (1.3)$$

con respecto a β_1, \dots, β_q .

Minimizando RSS nos da el estimador común de mínimos cuadrados. Esta minimización se puede hacer por diferenciación de (1.3) con respecto al vector β o usando las propiedades de proyección. En la notación anterior, y tomando $\hat{\beta}$ el estimador, se tiene:

$$\left(\sum_{t=1}^n z_t z_t' \right) \hat{\beta} = \sum_{t=1}^n z_t x_t. \quad (1.4)$$

Definiendo la matriz $Z = (z_1, z_2, \dots, z_q)'$ como una matriz $n \times q$ compuesta de n muestras de las variables de entradas y el vector observado $x = (x_1, x_2, \dots, x_n)'$ se puede hacer una simplificación de la notación. Esta identificación nos lleva a

$$(Z'Z)\hat{\beta} = Z'x \quad (1.5)$$

y la solución es

$$\hat{\beta} = (Z'Z)^{-1}Z'x \quad (1.6)$$

cuando la matriz $Z'Z$ es de rango q . El residual minimizado de suma de cuadrados (1.3) tiene la forma matricial equivalente

$$\begin{aligned} RSS &= (x - Z\hat{\beta})'(x - Z\hat{\beta}) \\ &= x'x - \hat{\beta}'Z'x \\ &= x'x - x'Z(Z'Z)^{-1}Z'x. \end{aligned}$$

El estimador común de mínimos cuadrados es insesgado, esto es, $\mathbb{E}(\hat{\beta}) = \beta$, y tiene la menor varianza de todos los estimadores insesgados lineales.

Si los errores $\{W_t\}$ son normalmente distribuidos (Gaussianos), $\hat{\beta}$ es también el estimador de máxima verosimilitud para β y es normalmente distribuido con

$$\begin{aligned}
\text{cov}(\widehat{\beta}) &= \sigma_w^2 \left(\sum_{t=1}^n z_t z_t' \right)^{-1} \\
&= \sigma_w^2 (Z'Z)^{-1} \\
&= \sigma_w^2 C
\end{aligned}$$

donde

$$C = (Z'Z)^{-1}. \quad (1.7)$$

Un estimador insesgado para la varianza σ_w^2 es

$$s_w^2 = \frac{RSS}{n - q} \quad (1.8)$$

contrastado con el estimador de máxima verosimilitud $\widehat{\sigma}_w^2 = RSS/n$ el cual tiene divisor n . Bajo la suposición de que s_w^2 tiene distribución proporcional a una variable aleatoria chi-cuadrado con $n - q$ grados de libertad, χ_{n-q}^2 , e independiente de $\widehat{\beta}$, se sigue que

$$t_{n-q} = \frac{(\widehat{\beta}_i - \beta_i)}{s_w \sqrt{c_{ii}}}$$

tiene una distribución t-de Student con $n - q$ grados de libertad; c_{ii} denota el i -ésimo elemento de la diagonal de C , como se definió en (1.7).

Hay varios modelos que se pueden utilizar de manera de seleccionar el mejor subconjunto de variables independientes. Suponiendo que se tiene un modelo que sólo considera un subconjunto $q_1 < q$ de variables independientes $z_{1t} = (z_{t_1}, \dots, z_{t_{q_1}})'$ que influncian a la variable x_t , así el modelo es

$$x_t = \beta_1' Z_{1t} + w_t \quad (1.9)$$

llega a ser la hipótesis nula, donde $\beta_1 = (\beta_1, \dots, \beta_{q_1})'$ es un subconjunto de los coeficientes de las q variables originales. Podemos contrastar el modelo reducido (1.9) contra el modelo completo (1.2) comparando el residual de la suma de cuadrados bajo los dos modelos usando el estadístico F .

La información envuelta en la prueba se resume en una tabla de Análisis de Varianza (ANOVA). La diferencia en el numerador es llamada regresión

de la suma de cuadrados.

Las técnicas discutidas se pueden usar para hacer comparación entre varios modelos. Estas pruebas han sido usadas en el pasado en una manera gradual, donde las variables son añadidas o suprimidas cuando los valores de la prueba F exceden algunos niveles predeterminados.

El procedimiento, llamado regresión múltiple por pasos, es útil para conseguir un conjunto de variables que sea de utilidad. Una manera alternativa de realizarlo es enfocarse en un procedimiento para selección del modelo que no sea secuencial, sino simplemente evaluar cada modelo en base a sus propios méritos. Suponiendo que se considera un modelo de regresión con k coeficientes y se denota el estimador de máxima verosimilitud para la varianza como

$$\hat{\sigma}_k^2 = \frac{RSS_k}{n} \quad (1.10)$$

donde RSS_k denota la suma residual de cuadrados bajo el modelo con k coeficientes de regresión.

Se considerará a continuación otro criterio para la selección del mejor modelo de regresión de series temporales, introducido por Hirotugu Akaike en 1974.

Definición 24. Criterio de Información de Akaike (AIC)

$$AIC = \ln \hat{\sigma}_k^2 + \frac{n + 2k}{n} \quad (1.11)$$

donde $\hat{\sigma}_k^2$ está dado por (1.10) y k es el número de parámetros en el modelo.

El valor de k que minimiza AIC especifica el mejor modelo. La idea es que la minimización de $\hat{\sigma}_k^2$ sea razonablemente objetiva, excepto que decrezca monotonamente cuando k crece. Por lo tanto, debemos penalizar la variación del error por un término proporcional al número de parámetros. La elección del término de penalización dado por la ecuación anterior no es único.

Capítulo 2

Análisis exploratorio de datos

En general, para que datos de series de tiempo sean estacionarios, es necesario hacer un promedio de productos en el tiempo, como en la sección anterior. Como para datos de serie de tiempo, es importante medir la dependencia entre los datos de la serie; al menos, debemos ser capaces de estimar la autocorrelación con precisión. Será difícil medir la dependencia de estos valores si la estructura de dependencia no es regular o si cambia en el tiempo. De ahí, que para realizar cualquier análisis estadístico significativo de datos de series de tiempo, será crucial que las funciones de media y autocovarianza satisfagan las condiciones de estacionaridad dadas en la Definición 15. A menudo, este no es el caso, y en esta sección daremos algunos métodos para lidiar con los efectos de no-estacionaridad sobre las propiedades estacionarias de las series a estudiar.

Quizás la forma más fácil de trabajar con series no-estacionarias es el modelo de tendencia estacionaria donde el proceso tiene comportamiento estacionario alrededor de una tendencia.

Consideremos el modelo de tendencia estacionaria donde el proceso tiene comportamiento estacionario alrededor de una tendencia. Podemos escribir este tipo de modelos como:

$$x_t = \mu_t + y_t,$$

donde x_t son las observaciones, μ_t denota la tendencia y y_t es un proceso estacionario.

Por lo general, una tendencia fuerte μ_t puede oscurecer el comportamiento del proceso estacionario y_t . De aquí, será una ventaja el que podamos remover la tendencia como un primer paso para un análisis exploratorio de los datos. Los pasos permiten obtener un estimador razonable del componente de tendencia, el cual puede ser llamado $\hat{\mu}_t$ y entonces trabajar con el residual

$$\hat{y}_t = x_t - \hat{\mu}_t.$$

Definición 25. *Definimos el **operador de cambio** por*

$$Bx_t = x_{t-1}$$

y extendemos a la potencia $B^2x_t = B(Bx_t) = Bx_{t-1} = x_{t-2}$ y así sucesivamente. Entonces

$$B^k x_t = x_{t-k}.$$

Considerando la primera diferenciación de series de tiempo como $\nabla x_t = x_t - x_{t-1}$ y la ecuación anterior, se puede reescribir la diferencia como:

$$\nabla x_t = (1 - B)x_t$$

y podemos extender la noción anterior. Por ejemplo, la segunda diferenciación será

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

por linealidad del operador. Para verificar esta igualdad, tomamos la diferencia de la primera diferencia $\nabla(\nabla x_t) = \nabla(x_t - x_{t-1}) = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2})$.

Definición 26. *Las **diferencias de orden d** son definidas como*

$$\nabla^d = (1 - B)^d$$

donde podemos desarrollar el operador $(1 - B)^d$ algebraicamente para evaluar valores superiores enteros de d . Cuando $d = 1$, podemos prescindir de él en la notación.

Una alternativa a la diferenciación es una operación menos severa que sigue asumiendo estacionaridad de una serie de tiempo. Este método es llamado diferenciación fraccionaria, y extiende la noción del operador diferencia a potencias fraccionarias $-0.5 < d < 0.5$ el cual sigue definiendo series estacionarias.

Un caso particular muy usado es la transformación

$$y_t = \ln(x_t)$$

el cual tiende a suprimir grandes fluctuaciones que ocurren sobre porciones de la serie donde los valores subyacentes son más grandes. Otra posibilidad es la transformación de potencias de la familia Box-Cox de la forma

$$y_t = \begin{cases} (x_t^\lambda - 1)/\lambda & \text{si } \lambda \neq 0 \\ \ln(x_t) & \text{si } \lambda = 0 \end{cases} .$$

Suavizado de Series de Tiempo

Realizar suavizado de series de tiempo es útil para descubrir ciertos rasgos, tales como tendencias de grandes términos o componentes estacionales. En particular si x_t representa las observaciones

$$m_t = \sum_{j=-k}^k a_j x_{t-j}$$

donde $a_j = a_{-j} \leq 0$ y $\sum_{j=-k}^k a_j = 1$ es un promedio móvil simétrico de los datos, el cual directamente suaviza la serie temporal.

Hay muchas otras técnicas para el suavizado de series de tiempo basándose en métodos de suavizado de gráficos de dispersión. La configuración general de una serie de tiempo está dada por:

$$x_t = f_t + y_t, \tag{2.1}$$

donde f_t es alguna función suave de tiempo, y y_t es un proceso estacionario. Se puede pensar del suavizado de promedio móvil planteado anteriormente como un estimador de f_t .

2.1 Modelos Autoregresivos y Promedio Móvil

Para la modelación de cualquier serie temporal, el uso de modelos autorregresivos y de media móvil siempre serán la primera herramienta, es por esto, que en esta sección se introducirán las definiciones de Modelo Autorregresivo de orden p ó $AR(p)$, Modelo de Promedio Móvil de orden q ($MA(q)$), y para concluir los Modelos Autorregresivos de Promedio Móvil o $ARMA(p,q)$.

2.1.1 Modelos Autoregresivos

Los modelos autoregresivos están basados en la idea de que el valor actual de la serie x_t se puede explicar como una función de p valores pasados x_1, x_2, \dots, x_{t-p} donde p determina el número de pasos necesarios para predecir el valor actual. Asumiremos que el valor actual es una función lineal particular de los valores pasados.

Definición 27. *Un modelo autotregresivo de orden p , abreviado $AR(p)$, es de la forma*

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t$$

donde x_t es estacionario, $t \in T \subseteq \mathbb{R}^+$, ϕ_1, \dots, ϕ_p son constantes ($\phi_p \neq 0$). A menos que se declare lo contrario, se asume que w_t es un ruido blanco gaussiano de media cero y varianza σ_w^2 . La media de x_t es cero. Si la media μ de x_t no es cero, reemplazamos x_t por $x_t - \mu$, es decir,

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \dots + \phi_p(x_{t-p} - \mu) + w_t$$

o escribimos

$$x_t = \alpha + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t,$$

donde $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$.

Note que la ecuación anterior es similar al modelo de regresión dado anteriormente y por consiguiente el término autoregresión. Sin embargo, se presentan algunas dificultades técnicas para la aplicación de ese modelo, porque los regresores x_{t-1}, \dots, x_{t-p} son aleatorios, mientras que z_t se asume fijo. Una forma más útil se deriva de usar el operador de cambios para escribir el modelo $AR(p)$ como

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)x_t = w_t,$$

o más conciso como

$$\phi(B)x_t = w_t.$$

Las propiedades de $\phi(B)$ son importantes para resolver la ecuación anterior. Se verán algunas a continuación.

Definición 28. *El operador autoregresivo se define como*

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p.$$

Consideremos ahora el modelo de primer orden AR(1), dado por $x_t = \phi x_{t-1} + w_t$. Iterando el operador de cambio k veces, obtenemos:

$$\begin{aligned} x_t &= \phi x_{t-1} + w_t = \phi(\phi x_{t-2} + w_{t-1}) + w_t \\ &= \phi^2 x_{t-2} + \phi w_{t-1} + w_t \\ &\vdots \\ &= \phi^k x_{t-k} + \sum_{j=0}^{\infty} \phi^j w_{t-j}. \end{aligned}$$

Este método sugiere por iteración continua del operador de cambio, siempre que $|\phi| < 1$ y x_t sea estacionario, se podrá representar un modelo AR(1) como un proceso lineal dado por:

$$x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j}.$$

El proceso AR(1) definido anteriormente es estacionario con media

$$\mathbb{E}(x_t) = \sum_{j=0}^{\infty} \phi^j \mathbb{E}(w_{t-j})$$

y la función de autocovarianza

$$\begin{aligned} \gamma(h) &= \text{cov}(x_{t+h}, x_t) = \mathbb{E}\left[\left(\sum_{j=0}^{\infty} \phi^j w_{t+h-j}\right)\left(\sum_{k=0}^{\infty} \phi^k w_{t-k}\right)\right] \\ &= \sigma_k^2 \sum_{j=0}^{\infty} \phi^j \phi^{j+h} = \sigma_k^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} = \frac{\sigma_k^2 \phi^h}{1 - \phi^2}, \quad h \geq 0. \end{aligned}$$

Recordando que $\gamma(h) = \gamma(-h)$ de modo que solo presentamos la función de autocovarianza para $h \geq 0$. De la ecuación anterior vemos que la ACF de un modelo AR(1) es

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \quad h \geq 0,$$

y $\rho(h)$ satisface la recursión

$$\rho(h) = \phi\rho(h-1), \text{ con } h = 1, 2, \dots$$

2.1.2 Modelos de Promedio Móvil

Como una alternativa a la representación autoregresiva en la cual x_t del lado izquierdo de la ecuación se asume como una combinación lineal, en los modelos de promedio móvil de orden q abreviados MA(q) asumimos el ruido blanco w_t del lado derecho de la ecuación que los define como una combinación lineal de los datos observados.

Definición 29. *El modelo de promedio móvil de orden q o modelo MA(q), se define como*

$$x_t = w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q},$$

donde hay q pasos en el promedio móvil y $\theta_1, \dots, \theta_q (\theta_q \neq 0)$ son parámetros. El ruido w_t se asume como un ruido blanco gaussiano.

Podemos también escribir el proceso MA(q) en la forma equivalente

$$x_t = \theta(B)w_t$$

usando la siguiente definición:

Definición 30. *El operador de promedio móvil es*

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q.$$

A diferencia del proceso autoregresivo, el proceso de promedio móvil es estacionario para cada valor de los parámetros $\theta_1, \dots, \theta_q$.

2.1.3 Modelos Autoregresivos de Promedio Móvil

Definición 31. Una serie de tiempo $\{x_t; t = 0, \pm 1, \pm 2, \dots\}$ es **ARMA** (p, q) si es estacionario y

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}$$

con $t \in T \subseteq \mathbb{R}^+$, $\phi_p \neq 0, \theta_q \neq 0$, y $\sigma_w^2 > 0$, donde w_t es un ruido blanco. Los parámetros p y q son llamados ordenes autoregresivos y de promedio móvil respectivamente. Si x_t tiene media μ distinto de cero, hacemos $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$ y escribimos el modelo como

$$x_t = \alpha + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}.$$

Como se observó previamente, cuando $q = 0$, el modelo es llamado autoregresivo de orden p , AR(p), y cuando $p = 0$ el modelo es llamado modelo de promedio móvil de orden q , MA(q).

Definición 32. Los **Polinomios AR y MA** se definen como

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p, \phi_p \neq 0$$

y

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q, \theta_q \neq 0$$

respectivamente, donde z es un número complejo.

Para abordar el primer problema, de ahora en adelante se hará referencia a un modelo ARMA(p, q) en el sentido de su forma más simple. Esto es, además de la definición original dada en la siguiente definición, requeriremos también que $\phi(z)$ y $\theta(z)$ no tengan factores comunes.

Definición 33. Un modelo ARMA(p, q), $\phi(B)x_t = \theta(B)w_t$, se dice que es **causal** si la serie de tiempo $\{x_t; t = 0, \pm 1, \pm 2, \dots\}$ se puede escribir como un proceso lineal de un lado

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \psi(B)w_t,$$

donde $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ y $\sum_{j=0}^{\infty} |\psi_j| < \infty$; haciendo $\psi_0 = 1$.

Definición 34. Un modelo ARMA(p, q), $\phi(B)x_t = \theta(B)w_t$ se dice **invertible** si la serie de tiempo $\{x_t; t = 0, \pm 1, \pm 2, \dots\}$ se puede escribir como

$$\pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = w_t,$$

donde $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ y $\sum_{j=0}^{\infty} |\pi_j| < \infty$; hacemos $\pi_0 = 1$.

Propiedades de los modelos ARMA(p, q):

1. **Causalidad de un proceso ARMA(p, q).** Un modelo ARMA(p, q) es causal si y solo si $\phi(z) \neq 0$ para $|z| \leq 1$. El coeficiente del proceso lineal se puede determinar resolviendo

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, |z| > 1.$$

Otra manera de ver la propiedad 1, es que un modelo ARMA es causal sólo cuando las raíces de $\phi(z)$ están fuera del círculo unitario, esto es $\phi(z) = 0$ sólo cuando $|z| > 1$.

2. **Invertibilidad de un proceso ARMA(p, q).** Un modelo ARMA(p, q) es invertible si y sólo si $\theta(z) \neq 0$ para $|z| \leq 1$. El coeficiente π_j de $\pi(B)$ se puede determinar al resolver

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

Otra manera de escribir la propiedad 2, es que un proceso ARMA es invertible solo cuando las raíces de $\theta(z)$ están fuera del círculo unitario, esto es, $\theta(z) = 0$ sólo cuando $|z| > 1$.

2.1.4 Otros modelos importantes

Modelos ARMA Integrado o ARIMA

Definición 35. Un proceso x_t se dice que es **ARIMA** (p, d, q) si:

$$\nabla^d x_t = (1 - B)^d x_t$$

es ARMA (p, q) . En general, se puede escribir el modelo como:

$$\phi(B)(1 - B)^d x_t = \theta(B)w_t.$$

Si $\mathbb{E}(\lambda^d x_t) = \mu$, el modelo se puede escribir como:

$$\phi(B)(1 - B)^d x_t = \alpha + \theta(B)w_t$$

donde $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$.

Los modelos ARIMA fueron popularizados en los años 70 por George Box y Gwilym Jenkins, y sus nombres se utilizan, frecuentemente, como sinónimos de la metodología ARIMA aplicada a análisis y predicción de series. Esta familia de modelos ha sido utilizada ampliamente a partir de los 80, debido a los avances de recursos de cálculo y de optimización. La principal ventaja de esta metodología es que proporciona predicciones óptimas en el plazo inmediato y en el corto plazo.

Modelo Autorregresivo de Promedio Móvil Estacional Multiplicativo

Definición 36. El modelo *autorregresivo de promedio móvil estacional multiplicativo* o modelo **SARIMA** de Box y Jenkins (1970) está dado por:

$$\Phi_p(B^s)\phi(B)\lambda_s^D \lambda^d x_t = \alpha + \Theta_Q(B^s)\theta(B)w_t$$

donde w_t es el proceso de ruido blanco gaussiano usual. El modelo general e denota como **ARIMA** $(p, d, q)x(P, D, Q)s$. Las componentes autorregresiva y de promedio móvil ordinarias están representadas por los polinomios $\phi(B)$ y $\theta(B)$ de ordenes p y q respectivamente. Por otro lado, las componentes autorregresivas y de promedio móvil estacional están representadas por $\Phi_p(B^s)$ y $\Theta_Q(B^s)$ de ordenes P y Q respectivamente. Las componentes de diferencia ordinaria y estacional por su parte se representan con $\lambda^d = (1 - B)^d$ y $\lambda_s^D = (1 - B^s)^D$.

2.2 Funciones de Autocorrelación y Autocorrelación Parcial

Para la obtención del orden (p, q) de un modelo ARMA se realiza una comparación entre las características que dos importantes funciones estadísticas

presentan para los distintos modelos teóricos y las características que tales funciones presentan en la serie objeto de estudio.

Tales funciones estadísticas son la función de autocorrelación (ACF), y la función de autocorrelación parcial (PACF), que son los dos instrumentos básicos en la fase de identificación del modelo, al permitirnos inferir el verdadero mecanismo subyacente que ha generado los datos.

2.2.1 Función de Autocorrelación (ACF)

Para un modelo ARMA (p,q) causal $\phi(B)x_t = \theta(B)w_t$, donde los ceros de $\phi(z)$ están fuera del círculo unitario, podemos escribir

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}.$$

De esta representación se sigue inmediatamente que $\mathbb{E}(x_t) = 0$. También, la función autocovarianza de x_t se puede escribir como,

$$\gamma(h) = cov(x_{t-h}, x_t) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}, h \geq 0.$$

Para obtener una ecuación en diferencias homogénea directamente en términos de $\gamma(h)$, se escribe:

$$\begin{aligned} \gamma(h) &= cov(x_{t+h}, x_t) = \mathbb{E} \left[\left(\sum_{j=1}^p \phi_j x_{t+h-j} + \sum_{j=0}^q \theta_j w_{t+h-j} \right) x_t \right] \\ &= \sum_{j=1}^p \phi_j \gamma(h-j) + \sigma_w^2 \sum_{j=h}^q \theta_j \psi_{j-h}, \quad h \geq 0, \end{aligned}$$

donde hemos usado el hecho de que $x_t = \sum_{k=0}^{\infty} \psi_k w_{t-k}$ y para $h \geq 0$,

$$\mathbb{E}(w_{t+h-j} x_t) = \mathbb{E} \left[w_{t+h-j} \left(\sum_{k=0}^{\infty} \psi_k w_{t-k} \right) \right] = \psi_{j-h} \sigma_w^2.$$

De las ecuaciones anteriores podemos escribir una ecuación general homogénea para la ACF de un proceso ARMA causal:

$$\gamma(h) - \phi_q \gamma(h-1) - \dots - \phi_p \gamma(h-p) = 0, \quad h \geq \max(p, q+1)$$

con condiciones iniciales

$$\gamma(h) - \sum_{j=1}^p \phi_j \gamma(h-j) = \sigma_w^2 \sum_{j=h}^q \theta_j \psi_{j-h}, \quad 0 \leq h < \max(p, q+1)$$

Dividiendo las dos ecuaciones anteriores por $\gamma(0)$ se obtiene que la ACF viene dada por $\rho(h) = \gamma(h)/\gamma(0)$.

2.2.2 Función de Autocorrelación Parcial (PACF)

Sabemos que para un modelo MA(q) la ACF proporciona una cantidad considerable de información sobre el orden de dependencia, cuando el proceso es proceso de promedio móvil. Si el proceso, sin embargo, es ARMA o AR, la ACF dice poco sobre los procesos de dependencia. Por lo tanto vale la pena buscar una función que se comporte como la ACF de los modelos MA, pero para los modelos AR, esta será la función de autocorrelación parcial.

Definición 37. La *función de autocorrelación parcial (PACF)* de un proceso estacionario x_t , denotada ϕ_{hh} , para $h = 1, 2, \dots$, es

$$\phi_{11} = \text{corr}(x_1, x_0) = \rho(1)$$

y

$$\phi_{hh} = \text{corr}(x_h - x_h^{h-1}, x_0 - x_0^{h-1}) \text{ para } h \geq 2.$$

Tanto $(x_h - x_h^{h-1})$ como $(x_0 - x_0^{h-1})$ son no-correlacionados con $\{x_1, x_2, \dots, x_{h-1}\}$.

Por estacionaridad, la PACF ϕ_{hh} es la correlación entre x_t y x_{t-h} con la dependencia lineal $\{x_{t-1}, \dots, x_{t-(h-1)}\}$ removida en cada uno.

Si el proceso x_t es gaussiano, entonces $\phi_{hh} = \text{corr}(x; t, x_{t-h} | x_{t-1}, \dots, x_{t-(h-1)})$. Esto es, ϕ_{hh} es el coeficiente de correlación entre x_t y x_{t-h} en la distribución bivariada de (x_t, x_{t-h}) condicionada por $\{x_{t-1}, \dots, x_{t-(h-1)}\}$.

2.3 Pronósticos

El objetivo en el pronóstico, es predecir los valores futuros de una serie de tiempo x_{n+m} , $m = 1, 2, \dots$ basado en los valores de la serie observados hasta el tiempo actual $x = \{x_n, x_{n-1}, \dots, x_1\}$. Asumiremos que x_t es estacionario y que los parámetros del modelo son conocidos.

El mínimo del error cuadrático medio del predictor x_{n+m} es:

$$x_{n+m}^n = \mathbb{E}(x_{n+m} | x_n, \dots, x_1)$$

porque la esperanza condicional minimiza el error cuadrático medio

$$\mathbb{E}[x_{n+m} - g(x)]^2$$

donde $g(x)$ es una función de las observaciones x .

Primero, nos restringiremos a los predictores que son función lineal de los datos, esto es, predictores de la forma

$$x_{n+m}^n = \alpha_0 + \sum_{k=1}^n \alpha_k x_k$$

donde $\alpha_0, \alpha_1, \dots, \alpha_n$ son números reales. Los predictores lineales de la forma anterior que minimizan el error cuadrático medio del predictor son llamados el mejor predictor lineal (**BLP's**).

2.3.1 Pronósticos para un proceso ARMA

Las ecuaciones de predicción general anteriores pueden dar una pequeña intuición en el pronóstico de los modelos ARMA en general. Asumiremos que x_t es un proceso ARMA(p,q) causal e invertible $\phi(B)x_t = \theta(B)w_t$ donde $w_t \sim iidN(0, \sigma_w^2)$.

Escribiremos el mínimo del error cuadrático medio del predictor x_{n+m} como x_{n+m}^n basado en los datos $\{x_n, x_{n-1}, \dots, x_1\}$, esto es:

$$x_{n+m}^n = \mathbb{E}(x_{n+m} | x_n, \dots, x_1).$$

Para un modelo ARMA, es fácil calcular el predictor de x_{n+m} asumiendo que tenemos el historial de los datos hasta n $\{x_n, x_{n-1}, \dots, x_1\}$. Denotaremos el predictor de x_{n+m} como:

$$\hat{x}_{n+m} = \mathbb{E}(x_{n+m}|x_n, \dots).$$

Ahora, escribamos el predictor en sus formas causal e invertible

$$x_{n+m} = \sum_{j=0}^{\infty} \psi_j w_{n+m-j}, \psi_0 = 1.$$

$$w_{n+m} = \sum_{j=0}^{\infty} \pi_j x_{n+m-j}, \pi_0 = 1.$$

Entonces, tomando esperanza condicional en el predictor de forma causal, tenemos:

$$\hat{x}_{n+m} = \sum_{j=0}^{\infty} \psi_j \hat{w}_{n+m-j} = \sum_{j=0}^{\infty} \psi_j w_{n+m-j}$$

ya que por la definición del predictor de forma invertible tenemos:

$$\hat{w}_t \equiv \mathbb{E}(w_t|x_n, x_{n-1}, \dots) = \begin{cases} 0, & \text{si } t > n \\ w_t, & \text{si } t \leq n \end{cases}.$$

De forma similar, tomando esperanza condicional en el predictor invertible, se tiene

$$0 = \hat{x}_{n+m} + \sum_{j=1}^{\infty} \pi_j \hat{x}_{n+m-j}$$

ó

$$\hat{x}_{n+m} = - \sum_{j=1}^{m-1} \pi_j \hat{x}_{n+m-j} - \sum_{j=m}^{\infty} \pi_j x_{n+m-j}$$

usando el hecho de que $\mathbb{E}(x_t|x_n, x_{n-1}, \dots) = x_t$, para $t \leq n$.

La predicción se consigue recursivamente usando la ecuación anterior iniciando con un predictor de un paso $m = 1$ y continuando para $m = 2, 3, \dots$. Usando la ecuación podemos escribir

$$x_{n+m} - \hat{x}_{n+m} = \sum_{j=0}^{m-1} \psi_j w_{n+m-j}$$

de modo que el error cuadrático medio de predicción se puede escribir como

$$P_{n+m}^n = \mathbb{E}(x_{n+m} - \hat{x}_{n+m})^2 = \sigma_w^2 \sum_{j=0}^{m-1} \psi_j^2.$$

Capítulo 3

Índices Inflacionarios

En este capítulo se introducirán las nociones y conceptos básicos con respecto a los indicadores macroeconómicos utilizados en el trabajo. Se analizará la importancia del Índice Nacional de Precios al Consumidor, su origen y antecedentes, su análisis y cálculo en la República Bolivariana de Venezuela.

Lo aquí presentado será la base para el análisis de los datos tomados de la sección estadística del Banco Central de Venezuela, los cuales son manipulados en el capítulo siguiente.

3.1 Inflación

En una economía de mercado los precios de los bienes y de los servicios están sujetos a cambios. Algunos aumentan y otros disminuyen. Se habla de inflación cuando se produce un aumento generalizado de los precios que no se limita a determinados artículos. Como resultado, pueden adquirirse menos bienes y servicios por cierta cantidad fija de dinero.

Por ende, podemos decir que la inflación es una medida económica que indica el crecimiento generalizado de los precios de bienes, servicios y factores productivos dentro de una economía en un periodo determinado. Para su cuantificación se usa el "índice de precios al consumidor".

El fenómeno inverso a la inflación es la deflación y esto sucede cuando los precios de los bienes, servicios y factores productivos decrecen de forma

continua y generalizada.

3.2 Índice de Precios al Consumidor

Es un indicador estadístico que mide, en un período determinado, los cambios ocurridos en los precios de una canasta de bienes y servicios representativa del consumo en hogares de una área geográfica específica. Está compuesto principalmente por las Encuestas de Presupuestos Familiares (EPF) y las Encuestas de Precios (EP).

La Encuesta de Presupuestos Familiares es una investigación por muestreo dirigida a los hogares, que tiene por objeto obtener información sobre sus ingresos, egresos, características de las viviendas que habitan, composición y otras variables económicas y sociales de sus miembros. Su principal objetivo es obtener la canasta de bienes y servicios y las ponderaciones del IPC.

Por otro lado, las Encuestas de Precios, tienen por objeto recolectar los precios de los productos contenidos en la canasta de bienes y servicios que conforman el IPC, en una muestra de los establecimientos que los comercializan.

Dentro de los principales usos del IPC, se encuentra su utilización como indicador de inflación (sabiendo que el IPC no incluye los precios de los consumos intermedios de las empresas ni de los bienes exportados); deflactor de las Cuentas Nacionales o contabilidad nacional y de otras estadísticas; actualización de deudas o de montos judiciales y como cláusula de revisión salarial.

En el mercado existe tal variedad de productos y servicios a los que podemos acceder que resulta prácticamente imposible registrarlos todos. Precisamente por eso el IPC se elabora en función de una cesta seleccionada de productos y servicios que se divide en grupos y subgrupos. Se trata de los bienes más habituales y más relacionados con la economía real y familiar.

3.2.1 Año Base

En las Cuentas Nacionales el año base se utiliza para las estimaciones a precios constantes de importantes indicadores como la producción y la demanda de bienes y servicios, a fin de determinar la correspondiente evolución en términos reales o físicos. En cuanto al IPC, en el año base se establecen las estructuras de ponderaciones, conforme a los gastos de consumo de las familias a partir del cual se define la canasta representativa de bienes y servicios para el cálculo del índice.

3.3 Índice Nacional de Precios al Consumidor INPC

El Índice Nacional de Precios al Consumidor INPC, es un indicador estadístico que mide en un período determinado, los cambios ocurridos en los precios de una canasta de bienes y servicios representativa del consumo de los hogares que habitan un área geográfica específica dentro del territorio nacional.

3.3.1 Antecedentes

El Banco Central de Venezuela (BCV) ha producido ininterrumpidamente, cada mes, desde 1950, el índice de Precios al Consumidor del área metropolitana de Caracas (IPC-AMC). Este indicador ha sido, desde entonces, una de las cifras estadísticas de mayor demanda por parte de los usuarios nacionales e internacionales.

En la década de 1980, el BCV producía los resultados del Índice Nacional de Precios del Área Metropolitana de Caracas (IPC-AMC) y de otras cinco ciudades: Maracay, Valencia, Puerto La Cruz, Barquisimeto y Maracaibo. Esta iniciativa se mantuvo hasta el año 1990, cuando la entonces Oficina Central de Estadísticas e Informática (OCEI) y el BCV, comenzaron a elaborar un índice de precios al consumidor con cobertura nacional, el cual se discontinuó en el año 2000.

Sin embargo, la necesidad de contar con un índice de mayor cobertura geográfica, impulsó la decisión de que el BCV y el Instituto Nacional de Es-

tadística (INE) unieran sus esfuerzos en procura de definir un nuevo indicador de precios al consumidor con mayor alcance geográfico. Es así como nace, en el año 2004, la idea del Índice Nacional de Precios al Consumidor (INPC).

En el año 2006, el BCV culminó los trabajos para producir un IPC representativo del área metropolitana de Maracaibo, el cual se comenzó a publicar en febrero de 2007.

Más recientemente, el BCV introdujo un conjunto de mejoras metodológicas en los índices de Caracas y Maracaibo, con base en los resultados de la III Encuesta Nacional de Presupuestos Familiares (III ENPF), levantada durante el año 2005.

3.3.2 Establecimiento del 2007 como año base en Venezuela

En el año 1996, el Banco Central de Venezuela inició la ejecución del Programa de Actualización de las Estimaciones Macroeconómicas (PRACEM) para la adopción de los nuevos manuales internacionales de Cuentas Nacionales y de Balanza de Pagos, los cuales fueron editados por la Organización de las Naciones Unidas (ONU) y el Fondo Monetario Internacional (FMI), respectivamente, con el fin de actualizar y mejorar los aspectos conceptuales y metodológicos de esos sistemas contables.

Como parte de los objetivos del Pracem se contempló la actualización del año base de referencia de las Cuentas Nacionales y del Índice de Precios al Consumidor (IPC), en concordancia con la adopción de los nuevos manuales y ante los importantes cambios estructurales e institucionales ocurridos en la producción, demanda y precios, entre otras variables macroeconómicas. En ese sentido, se definió el año 1997 como nueva base en sustitución de 1984.

3.3.3 Cálculo del INPC

Actualmente la canasta de gasto familiar está compuesta de 287 rubros, los cuales representan casi la totalidad del gasto de consumo total que realizan los hogares del área metropolitana de Caracas. Esta canasta está subdividida en 13 grupos, cuya estructura de ponderación para el cálculo del INPC es la siguiente:

Grupos de Bienes y Servicios	Ponderaciones
1. Alimentos y Bebidas no Alcohólicas	22.9
2. Bebidas alcohólicas y tabaco	1.5
3. Restaurantes y hoteles	6.1
4. Vestido y calzado	6.5
5. Alquiler de vivienda	15.1
6. Servicios de la vivienda	3.5
7. Mobiliarios, equipos del hogar y mantenimiento	5.6
8. Salud	4.2
9. Transporte	13.3
10. Comunicaciones	5.2
11. Esparcimiento y Cultura	5.1
12. Servicios de Educación	4.8
13. Bienes y Servicios diversos	6.2
Total General	100.0

El valor del IPC es afectado por la ponderación o grado de importancia que tiene cada artículo en el gasto familiar. Es decir, mientras mayor sea el gasto destinado a un grupo de bienes o servicios, mayor será su influencia en el resultado del IPC, ya que éste expresa las variaciones promedio de precios, integrados por la importancia o peso relativo que tiene cada rubro en el gasto familiar. Por ejemplo, en 1997 una familia del área metropolitana de Caracas destinaba 22,9 por ciento de su gasto total a alimentos y bebidas no alcohólicas, mientras que a bebidas alcohólicas y tabaco asignaba 1,5 por ciento. Por esta estructura de gastos, en el resultado del IPC influyen más los precios del primer grupo que los del segundo.

En el cálculo del INPC se utiliza un Índice de tipo Laspeyres, lo que implica que las cantidades de los rubros en la canasta permanecen fijas en el tiempo, mientras que los precios cambian constantemente.

La fórmula utilizada para el cálculo del INPC es:

$$I_{n+1} = \sum_{j=1}^n \frac{P_j Q_j}{P_0 Q_0} * 100,$$

donde P_0 son los precios del producto analizado (que variará dependiendo del rubro estudiado) en el tiempo 0, P_n los precios en el tiempo n, y Q_0 son las cantidades en el precio base.

3.3.4 Proceso de recaudación de los datos

Para recabar la información de precios de los 287 rubros que conforman la nueva canasta, mensualmente se visitan 2.000 establecimientos de diferentes tipos, obteniéndose un total de 32.000 precios. Para garantizar la veracidad de la información recolectada por los encuestadores del BCV, se les proporciona un cuestionario donde se identifican con precisión los productos cuyos precios deben recopilar y los establecimientos donde deben hacerlo, con el fin de evitar errores. Asimismo, los encuestadores son rotados regularmente, según zonas geográficas y tipo de encuestas.

3.3.5 Problemas de Medición del INPC

EL objetivo del índice de precios al consumidor es medir las variaciones del costo de la vida. Sin embargo, no es un indicador perfecto y presenta tres problemáticas de medición que son difíciles de resolver:

1. Efectos de sustitución.
2. Introducción de nuevos bienes o desaparición de viejos bienes.
3. Cambio no medidos en la calidad

1. Efecto Sustitución.

Cuando los precios varían de un año a otro, no varían todos proporcionalmente: unos varían más que otros. Así, los consumidores preferirán aquellos bienes y servicios que hayan subido menos. Sin embargo, el INPC se calcula suponiendo que la cesta de bienes y servicios es fija.

Dado esto, al no tener en cuenta la posibilidad que los consumidores sustituyan unos bienes por otros, sobreestima el aumento que experimenta el costo de la vida año tras año.

2. Introducción de nuevos bienes o salida de viejos bienes.

Cuando se introduce un bien los consumidores tienen más variedad para elegir. Esto significa que la moneda es más valiosa por lo que los consumidores necesitan menos dinero para mantener un determinado nivel de vida. Así mismo, cuando desaparecen del mercado un producto, o hay escasez, los consumidores naturalmente poseen menos variedad para escoger. Esto significa que la moneda es menos valiosa puesto que se requiere de una mayor cantidad de dinero para adquirir los mismos bienes y servicios que eran poseídos en el periodo de tiempo anterior.

Sin embargo, como el INPC se basa en una cesta fija no refleja el cambio en el poder adquisitivo de la moneda.

3. Cambio no medido en la calidad.

Si la calidad de un bien empeora de un año a otro, disminuye el valor del dinero, incluso aunque el precio del bien no varíe. Así mismo, si la calidad de un bien mejora, aumenta el valor del dinero. Como los precios son ajustados por el mercado, estas mejoras solo serán capturadas por el INPC cuando se hayan ajustado por el mercado.

Capítulo 4

Análisis de la Serie del INPC en Venezuela entre los años 2000 y 2012

En este capítulo presentaremos un análisis práctico de lo expuesto en los capítulos anteriores, es decir, se utilizará la teoría de Series Temporales para realizar el análisis de el Índice Nacional de Precios al Consumidor de diversos rubros en los cuales está subclasificado el mismo en Venezuela.

A través de la utilización de "R project", un software libre diseñado para realizar cálculos de estadística computacional y graficación, se analizarán los datos en la búsqueda de adaptarles modelos ARMA, ARIMA o SARIMA que nos permita la predicción de datos futuros.

4.1 Serie correspondiente a la variación acumulada del subgrupo "Alimentos" (Enero 2000 - Septiembre 2012)

Dentro de los subgrupos entre los cuales se encuentra dividida la recolección del INPC, los datos asociados a la inflación de los alimentos arrojan gran información con respecto al poder adquisitivo de los venezolanos.

El Índice Nacional de Precios al Consumidor está subdividido en 13

rubros, siendo "Alimentos y Bebidas no alcohólicas" el que cuenta con una ponderación significativamente mayor. Dentro de este rubro existen naturalmente 2 subclasificaciones, alimentos y bebidas no alcohólicas. En esta sección nos dedicaremos al análisis de la serie correspondiente al primero de las subdivisiones, con una data mensual desde Enero del año 2000 hasta el penúltimo trimestre del año 2012, y se realizaran predicciones para obtener de forma analítica los valores posibles de la variación correspondiente al último trimestre del año 2012.

En el siguiente cuadro se presentan los valores con los que fueron realizadas las predicciones, desde enero del 2000 hasta diciembre de 2007.

Mes / Año	2000	2001	2002	2003	2004	2005	2006	2007
Enero	16.3	17.8	21	30.3	41.5	52.2	60.6	79.4
Febrero	16.2	17.8	21.1	31.5	42.1	51.2	58.3	79.3
Marzo	16.1	18	22.4	30.3	42.5	51.3	58.4	75.6
Abril	16.2	18.5	23.1	30.8	43.3	51.7	58.7	77.4
Mayo	16.2	19.2	23.5	32.9	44.3	54.5	61.5	79.7
Junio	16.3	19.4	23.8	33.9	45.4	54.8	64.9	81.6
Julio	16.5	19.8	24.8	34.4	46.3	55.2	68.2	82.2
Agosto	16.7	19.8	25.7	34.8	47.6	56.2	71.1	84.1
Septiembre	16.9	19.8	26.6	35.6	47.4	56.6	72.1	85.7
Octubre	17.2	20.3	27.8	36.9	47.3	58.9	72.1	89.3
Noviembre	17.2	28.6	28.6	38.3	49.1	59.9	73.9	95.5
Diciembre	17.6	28.9	28.9	39.4	50.6	60.6	76.4	100

Tabla 4.1. Fuente: Banco Central de Venezuela (BCV)

A continuación, se presentan los datos del mismo subgrupo para los años 2008 al 2012, nótese que los último tres meses correspondientes al 2012 permanecen en blanco puesto que son los datos que calcularemos a través de las series temporales y el ajuste de un modelo.

Mes/Año	2008	2009	2010	2011	2012
Enero	104.8	150.6	180.4	248.3	323.1
Febrero	106.8	149.6	180.9	247.9	325.7
Marzo	107.8	149.7	188.2	250.6	327.7
Abril	109.5	149.8	211.8	253	333.2
Mayo	117.4	150.4	216.7	261	339.1
Junio	122.1	152.4	221.6	270	340.6
Julio	123.2	157.5	223.6	284	340.4
Agosto	126.7	161.8	225.7	292	342.2
Septiembre	131.4	167.8	225.5	294	350.9
Octubre	135.2	171	229.8	304	
Noviembre	140.3	174.2	231.8	313	
Diciembre	146.7	176.6	237.3	318	

Tabla 4.2. Fuente: Banco Central de Venezuela (BCV)

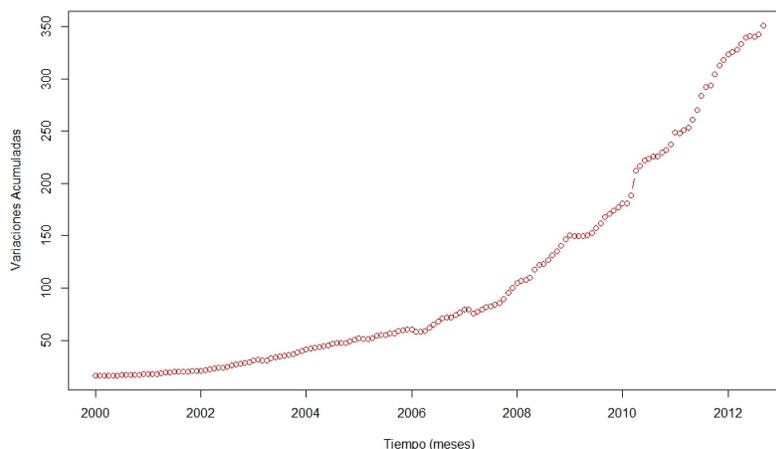
A partir de los valores antes presentados se constituirá la serie temporal x_t asociada, donde t variará entre los años 2000 y 2012 de forma mensual.

En particular, la data presentada corresponde a los valores acumulados de la variación porcentual del INPC con respecto a los alimentos, es decir, utilizando el año base 2000 y partiendo de este, cada punto no es más que la suma de la variación correspondiente a ese mes con los datos anteriores. Así, para Enero del 2008 el valor asignado corresponde a la variación correspondiente a ese mes más la suma de todos los datos anteriores.

Graficando los datos presentados anteriormente, podemos visualizar la serie asociada (Figura 4.1).

Esta gráfica permite observar lo explicado anteriormente, claramente presenta una tendencia creciente, es decir, cada valor incrementa con el paso de los años, consecuencia directa de la estructura de los datos que caracterizan la serie de tiempo x_t , puesto que al ser una suma acumulativa la tendencia será principalmente creciente. Sin embargo, es importante mencionar como a partir del año 2008 la velocidad de crecimiento se incrementa significativamente, lo cual representa una disminución del poder adquisitivo de los compradores, puesto que a partir de este punto adquirieron menos productos por el mismo valor. Así mismo, podemos ver la gráfica de Autocorrelación (ACF) y Autocorrelación Parcial asociada (PACF) para observar la presencia de un componente estacional (Figura 4.2)

Figura 4.1: INPC: Alimentos (Serie Enero 2000 - Septiembre 2012)



En la gráfica correspondiente a la función de autocorrelación (acf) se puede observar una tendencia correspondiente a los datos, aun cuando no permite deducir el orden del autoregresivo que puede ser ajustado, esto se ve como consecuencia de notar que todas las barras de la función salen de las bandas de confianza y decrecen lenta pero constantemente. Así mismo, en la función de autocorrelación parcial (pacf) las barras observadas son tan cercanas al cero que no permiten que se deduzca la existencia de un componente recurrente mensual.

A través de un comando específico de "R project" se puede extraer automáticamente tanto la tendencia que presentan los datos como el componente periódico. A continuación se presentará la gráfica de los distintos elementos que conforman la data.

Ahora bien, se puede observar en la Figura 4.4, como primer componente la serie temporal original, en la segunda parte de la gráfica se observa la tendencia que fue conseguida dentro de los datos, así mismo el componente estacional visualizado fue graficado en la tercera parte de esta gráfica, y los residuos restantes al eliminar la tendencia y el componente periódico de la serie Alimentos aparecen de último. En ella aun se puede observar cierto

Figura 4.2: Funciones de Autocorrelación y Autocorrelación Parcial, serie Alimentos

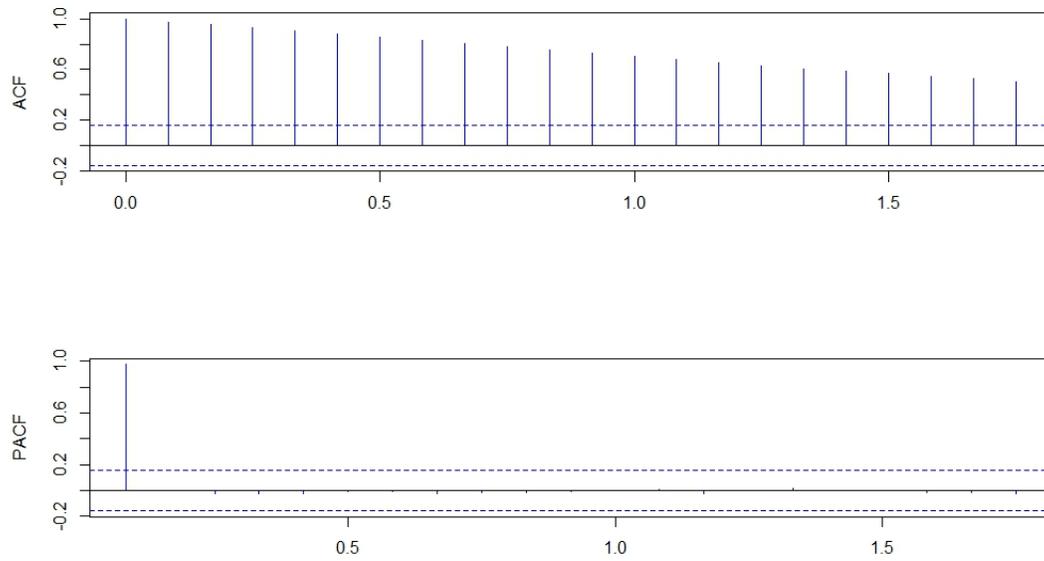
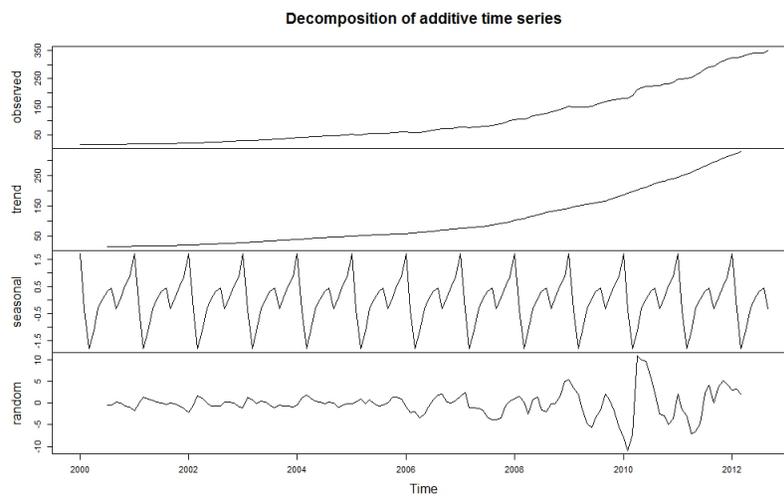
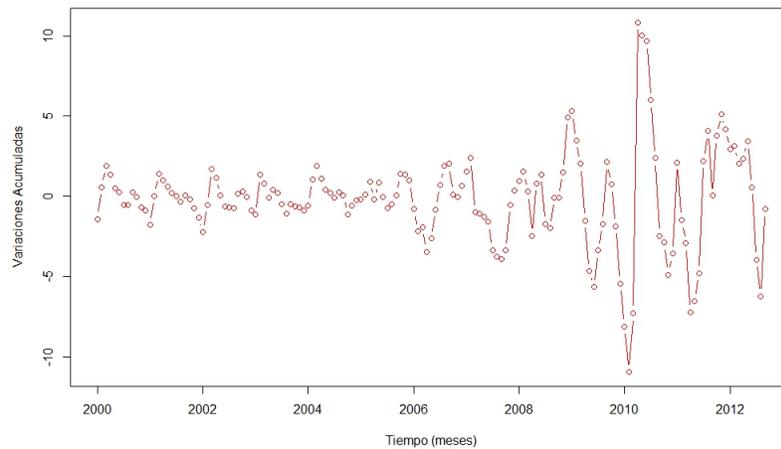


Figura 4.3: Descomposición de la serie "Alimentos"



patrón repetitivo, por lo que se procederá a la realización de diferencias entre los datos para la eliminación del mismo.

Figura 4.4: Serie "Alimentos" - sin tendencia o componente estacional



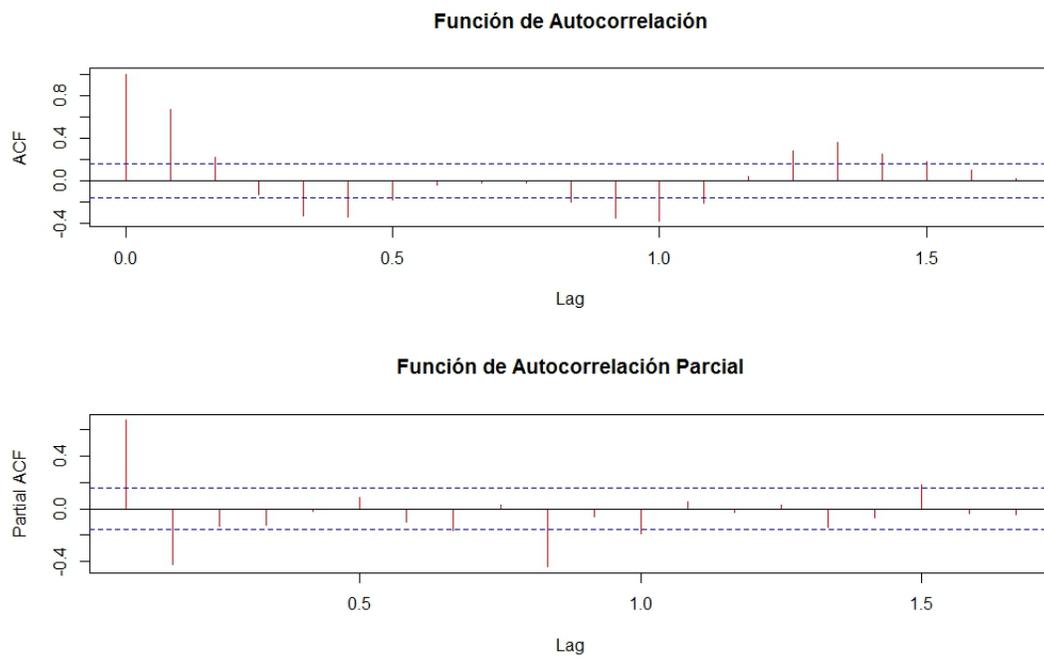
En la Figura 4.5, correspondiente a la función de autocorrelación se puede apreciar el patrón recursivo apreciado en la primera parte de la gráfica de los residuos, deducido por la forma en como se comportan las barras dentro de la gráfica, las cuales mantienen el mismo comportamiento cada cierto número de iteraciones, es por esto que se procedió a la realización de una serie de pruebas de diversos modelos SARIMA $(p,d,q) \times (P,D,Q)_s$ (Capítulo 2, Definición 36), motivados por el componente estacional, y se observó su coeficiente AIC (Criterio de Akaike), después de la comparación realizada el modelo seleccionado fue:

SARIMA(3,0,9)x(6,2,0)₄, cuyo AIC es 2.433

Como se vio en el Capítulo 1 (Definición 24), el Criterio de Akaike indica que los valores de los parámetros que minimizan el resultado obtenido especifica el mejor modelo.

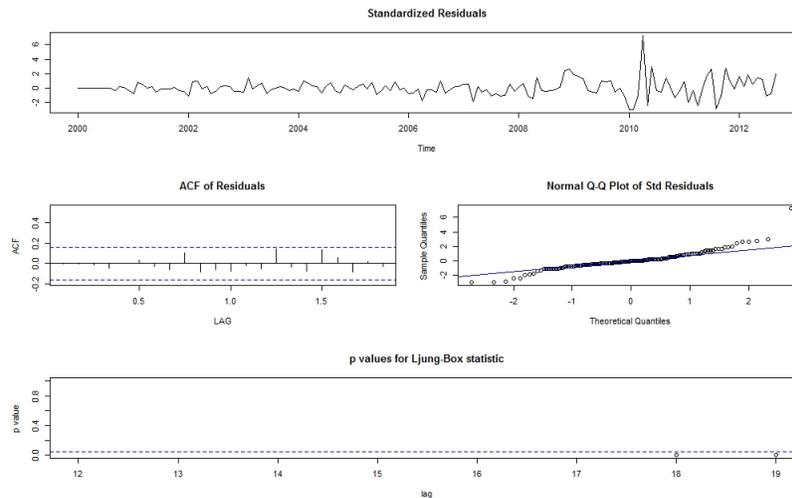
A continuación se presenta la gráfica correspondiente a diversas pruebas generadas con respecto a los residuos del modelo creado, estos incluyen la

Figura 4.5: Función de autocorrelación y autocorrelación parcial de los residuos



gráfica correspondiente a los residuos, la función de autocorrelación de los mismos, el gráfico Q-Q y la tabla de p-valores. (Figura 4.6), el cual asevera la gaussianidad de los residuos del ajuste del modelo. Esto indica como el modelo seleccionado es el más cercano a los datos presentados.

Figura 4.6: Gráfico de residuales



Así mismo, y para apreciar con detalle, se presentan las gráficas de las funciones de autocorrelación (acf) y autocorrelación parcial (pacf) de los residuos (Figura 4.7). En esta se puede apreciar como finalmente el patrón recursivo desapareció puesto que las líneas se encuentran por dentro del intervalo de confianza y no se puede apreciar ningún componente que se repita bajo ninguna variación de iteraciones.

Al crear el modelo SARIMA antes mencionado, obtenemos una nueva serie que se ajusta principalmente al ruido blanco w_t que se obtuvo luego de que la serie temporal x_t perdiese su tendencia y componente estacional, se deben agregar estos valores antes eliminados, para obtener un modelo que se ajuste a la serie del subgrupo "Alimentos" original x_t . En la siguiente gráfica se puede apreciar la comparación entre el modelo ajustado y la data original luego de realizar la suma de estos componentes.

Claramente se puede apreciar en la gráfica 4.8 lo cercano que se encuentra el modelo SARIMA construido con la serie original, esto permite deducir que

Figura 4.7: ACF y PACF de los residuos

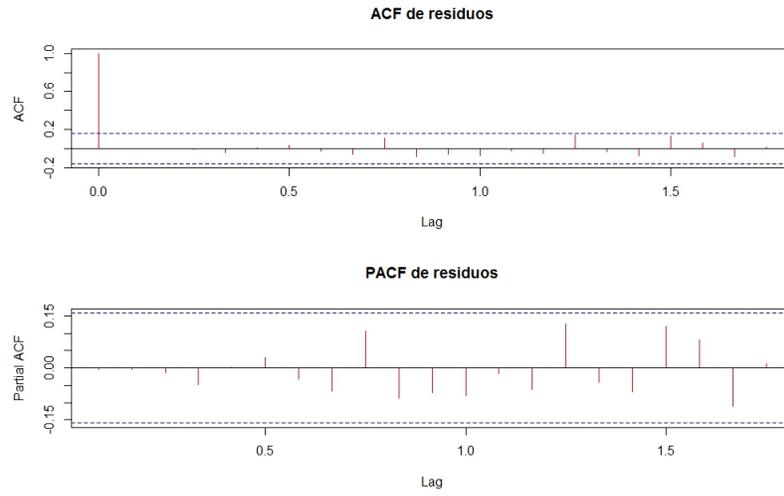
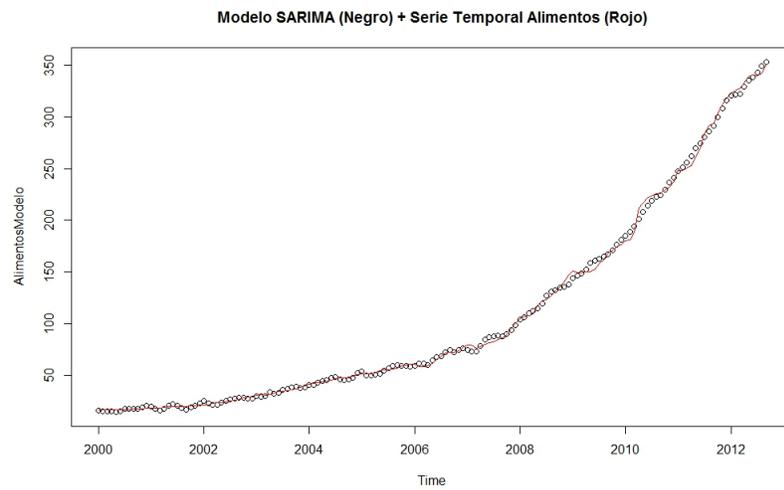


Figura 4.8: Modelo SARIMA y Serie original "Alimentos"



la estimación fue acertada.

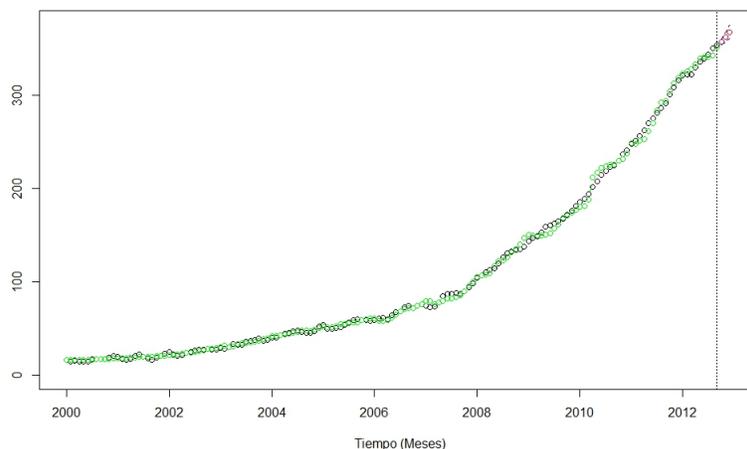
Para la realización de las predicciones correspondientes al último trimestre del año 2012, se calculan los siguientes tres valores de la serie de tiempo x_t , y se le añaden los valores correspondientes a la tendencia y al componente estacional respectivo a los meses necesarios.

Al realizarse la predicción se obtienen los siguientes resultados, los cuales son comparados con los valores emitidos por el Banco Central de Venezuela, y su diferencia es mostrada:

Fuente/Mes	Octubre	Noviembre	Diciembre
SARIMA (3,0,9)x(6,2,0)4	356.8461	361.7988	367.3351
BCV	344.8	355.8	376.0
Diferencia	12.0461	5.9988	8.6649

Para culminar, se puede observar la gráfica donde se encuentran tanto la serie asociada al subgrupo "Alimentos", el Modelo SARIMA (3,0,9)x(6,2,0)4 y las predicciones realizadas. (Figura 4.9)

Figura 4.9: Modelo SARIMA, Serie original "Alimentos", Predicciones



Para concluir, se puede señalar que las predicciones correspondientes al

último trimestre del año 2012 no fueron tan precisas como se esperaba, considerando lo preciso del resultado obtenido en la modelación de los 153 datos poseídos. Esto ocurre debido al tamaño de la data, la cual, en estadística computacional y en el análisis de Series Temporales es considerada pequeña. Sin embargo, podemos asumir que al tener una data mucho más extensa la predicción de los datos alcanza una precisión excelente.

4.2 Serie correspondiente a la variación mensual del subgrupo "Tabaco" (Enero 2008 - Septiembre 2012)

Uno de los rubros de clasificación del INPC, Bebidas Alcohólicas y Tabaco, representa la división con menor ponderación entre todas, puesto que son considerados bienes prescindibles dentro de la economía doméstica cotidiana.

Como parte de este grupo clasificador, podemos encontrar dos claras subdivisiones: Bebidas Alcohólicas y Tabaco. Sin embargo, nos dedicaremos al cálculo de predicciones, igual que en el ejemplo anterior, al segundo subgrupo.

Se presentan los datos correspondientes a las variaciones mensuales no acumuladas para el período Enero 2008 - Septiembre 2012, es decir, cada valor corresponde sólo a la variación porcentual observada correspondiente al mes en donde fue obtenido; esto presenta una clara diferencia con el ejemplo presentado anteriormente. En esta sección se espera calcular satisfactoriamente los datos correspondientes al último trimestre del año 2012.

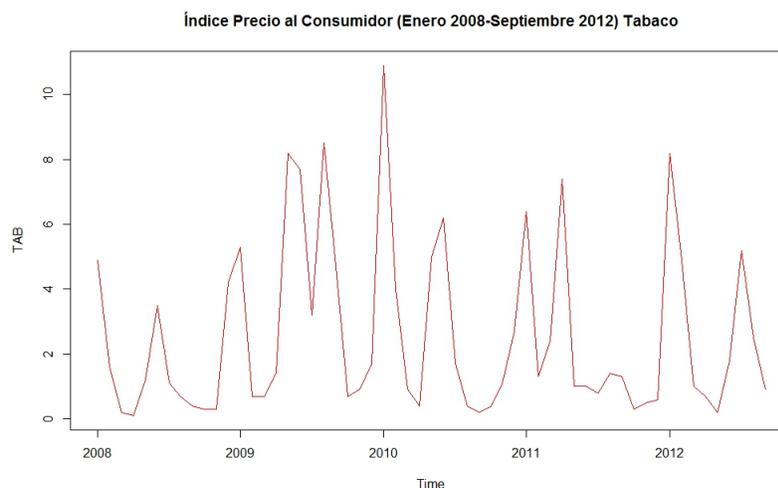
Mes/Año	2008	2009	2010	2011	2012
Enero	4.9	5.3	10.9	6.4	8.2
Febrero	1.6	0.7	4	1.3	4.9
Marzo	0.2	0.7	0.9	2.4	1
Abril	0.1	1.4	0.4	7.4	0.7
Mayo	1.2	8.2	5	1	0.2
Junio	3.5	7.7	6.2	1	1.8
Julio	1.1	3.2	1.7	0.8	5.2
Agosto	0.7	8.5	0.4	1.4	2.5
Septiembre	0.4	4.7	0.2	1.3	0.9
Octubre	0.3	0.7	0.4	0.3	
Noviembre	0.3	0.9	1.1	0.5	
Diciembre	4.2	1.7	2.7	0.6	

Tabla 4.4. Fuente: Banco Central de Venezuela.

Considerando los datos antes presentados, y asociándolos directamente a

una serie temporal y_t donde $t \in T$, variando de forma mensual desde Enero del 2008 hasta Septiembre 2012. Graficando y_t , podemos visualizar la serie.(Figura 4.10)

Figura 4.10: Serie "Tabaco"



Ahora bien, notoriamente es complicado deducir la presencia de tendencia o componente estacional en esta gráfica, por lo tanto se presentará su función de autocorrelación y autocorrelación parcial asociada (Figura 4.11).

Luego de visualizar las gráficas correspondientes a las funciones ACF y PACF, se sigue sin poder establecer claramente un patrón de repetición mensual o una tendencia clara, puesto que en ambas gráficas todas las barras están contenidas dentro del intervalo de confianza. Por lo tanto se utiliza una vez más el comando de "R Project" que permite la descomposición de la serie en sus elementos constitutivos. En la Figura 4.12 se pueden apreciar los distintos componentes de la serie.

Al descomponer la serie original, se puede visualizar en el primer gráfico la serie temporal asociada a subgrupo Tabaco, así mismo en el segundo y tercer gráfico, se pueden observar tanto la tendencia como la estacionalidad de la serie respectivamente; y por último se puede observar el residuo generado de eliminar los componentes antes mencionados de la serie original y_t .

Figura 4.11: ACF y PACF "Serie Tabaco"

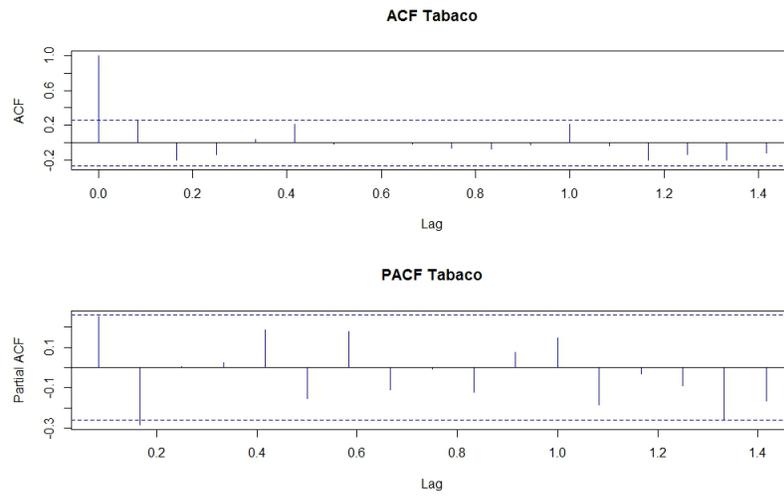
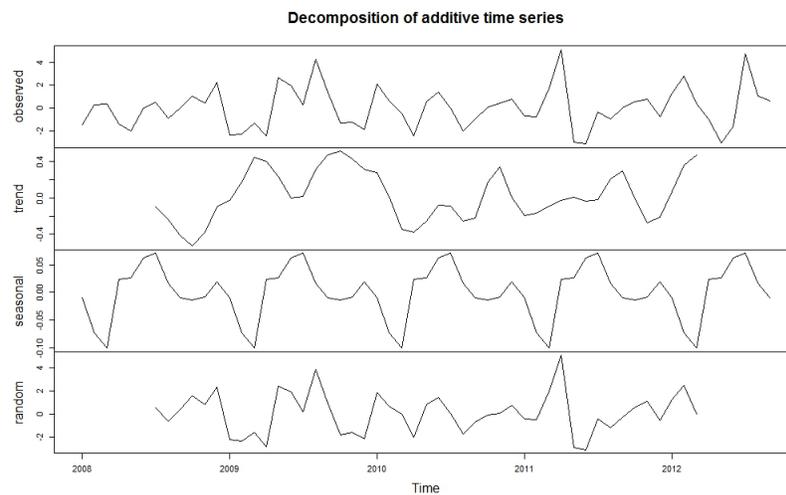
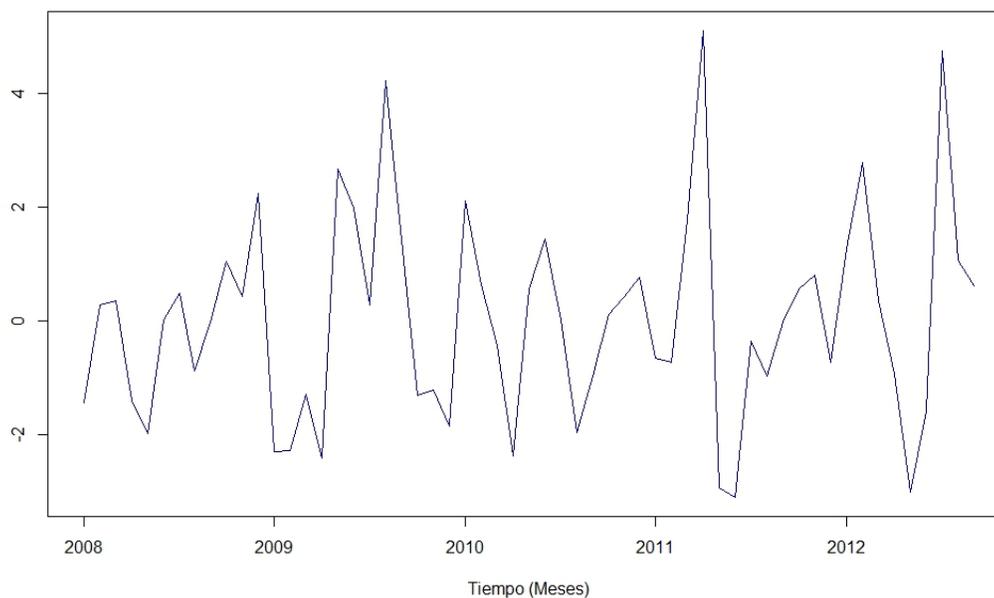


Figura 4.12: Gráfica de componentes de la serie Tabaco



(Figura 4.13).

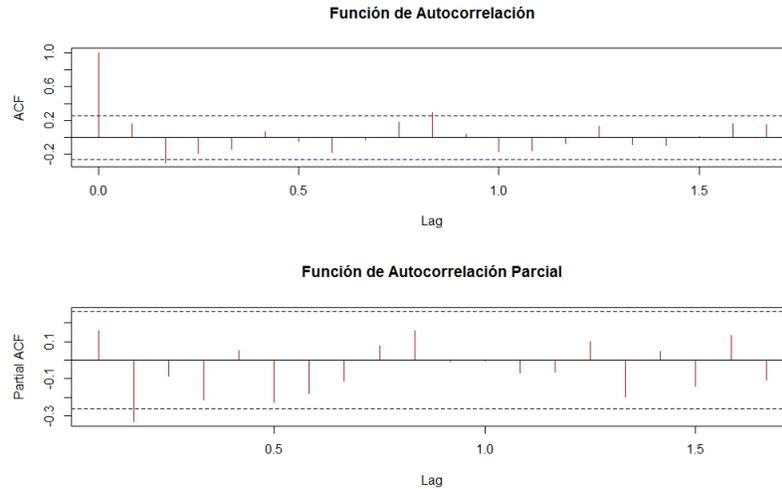
Figura 4.13: Serie Tabaco - Sin tendencia o componente estacional



Posteriormente a la eliminación de estos componentes, se grafican las funciones ACF y PACF de los residuos, con el fin de determinar el orden correspondiente al modelo ARMA que ajustará mejor los datos (Figura 4.14).

Al observar en la gráfica de la ACF la segunda barra como claramente sale de las bandas, y de igual forma la primera barra en la PACF, se puede advertir que el modelo que mejor se ajustará será un ARMA (2,1), sin embargo, se procederá a la prueba de una serie de variaciones de estos parámetros y al cálculo de su Criterio de Akaike (AIC) como valor referencial.

Figura 4.14: Funciones ACF y PACF de los residuos



Modelo	AIC
ARMA(2,1)	224
ARMA(1,1)	228.46
ARMA(0,1)	231.98
ARMA(2,2)	223.99
ARMA(2,0)	229.41

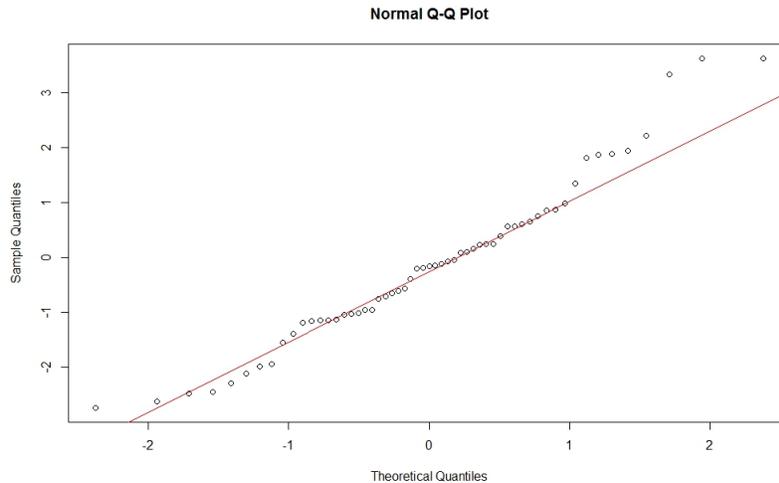
Tabla 4.5

De los resultados obtenidos y lo visto en el Capítulo 1, se selecciona como el modelo que mejor se ajuste a la serie original, al ARMA(2,2) cuyo valor AIC es 223.99, recordando una vez más la definición del Criterio de Akaike (Definición 24) que establece que mientras más pequeño sea el valor de AIC más ajustado y exacto será el modelo utilizado. De los parámetros seleccionados podemos decir que la serie temporal construida viene dada por:

$$x_t = 0.0625x_{t-1} + 0.3248x_{t-2} - 0.0046x_{t-1} - 0.9954x_{t-2}$$

Se puede observar el Gráfico Q-Q de los residuos del modelo ajustado para observar su gaussianidad, lo cual se deduce por lo ajustado que se encuentran los puntos a la barra en el centro de la gráfica. (Figura 4.15)

Figura 4.15: Gráfico Q-Q



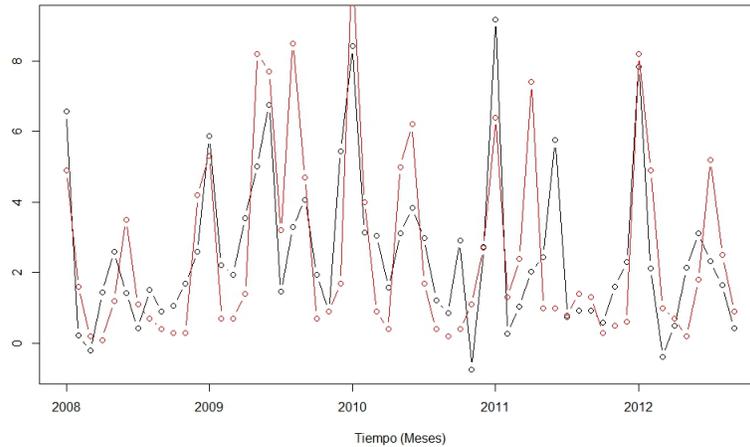
Al crear el modelo ARMA antes mencionado, obtenemos una función que ajusta el ruido blanco w_t obtenido luego de restar a la serie original componentes tanto estacionales como de tendencia, es por esto que se deben agregar los valores de periodicidad y la tendencia antes eliminados, así se logra obtener un modelo que se ajusta a la serie del subgrupo "Tabaco" original. En la siguiente gráfica se encuentran tanto los valores asignados a la serie original, como los hallados con la serie de tiempo construida, su objetivo es la comparación entre ambas (Figura 4.16).

Claramente se puede apreciar en la gráfica anterior que si bien el modelo ARMA construido no es precisamente igual en todos los puntos a la serie "Tabaco" original, es una buena aproximación de los valores de esta, puesto que con sólo 57 datos el error posible en los cálculos es muy alto.

Para la realización de las predicciones correspondientes al último trimestre del año 2012, se calcula los siguientes tres valores de la serie de tiempo x_t antes expuesta, y se le añaden los valores correspondientes a la tendencia y al componente estacional respectivo a los meses necesarios.

Al realizarse la predicción se obtienen los siguientes resultados, los cuales son comparados con los valores emitidos por el Banco Central de Venezuela,

Figura 4.16: Serie Tabaco (rojo), Modelo ARMA (2,2) (negro)



y su diferencia es mostrada:

Fuente/Mes	Octubre	Noviembre	Diciembre
ARMA (2,2)	0.69	0.187	5.132
BCV	1.1	3.3	4.1
Diferencia	0.41	3.113	1.032

Tabla 4.6

Para culminar, se puede observar la gráfica donde se encuentran tanto la serie asociada al subgrupo "Tabaco", el Modelo ARMA (2,2) y las predicciones realizadas (Figura 4.17).

Como conclusión se puede inferir, que una vez al presentar una data de tan corto tamaño a medida que se vuelven lejanas las predicciones, se debilita su exactitud. Sin embargo, en este ejemplo, a diferencia del anterior, no se presentaron diferencias considerables en los tres pasos en los que se realizó. Así mismo, podemos deducir, que los modelos ARMA son la mejor herramienta para predecir información basada en el Índice Nacional de Precios al Consumidor, con datos no acumulados.

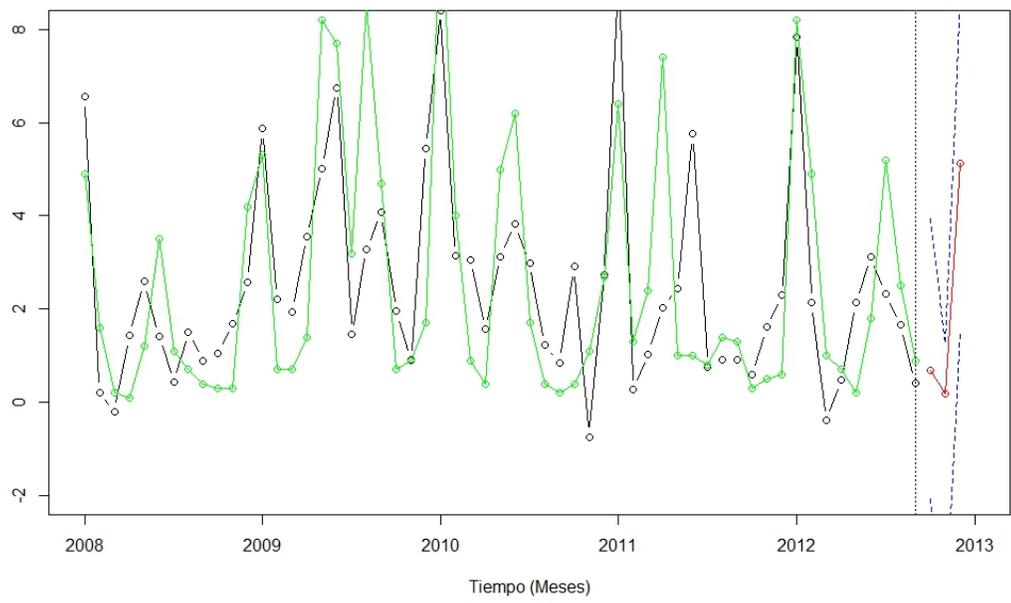


Figura 4.17: Serie "Tabaco" (verde), Modelo ARMA (negro), Predicciones (rojo)

4.3 Serie correspondiente a la variación mensual del subgrupo "Alquiler de Viviendas" (Enero 2008 - Septiembre 2012)

Con el fin de generalizar y obtener conclusiones, se presentará otro ejemplo corto correspondiente a otro subgrupo de rubros del INPC, de variaciones acumuladas. Sin embargo, no se presentará de forma detallada todos los pasos para la construcción del modelo.

La serie asociada a este subgrupo viene presentada a través de variaciones porcentuales acumuladas mes tras mes, de igual forma que la serie Alimentos del ejemplo anterior. A continuación se presentan los datos:

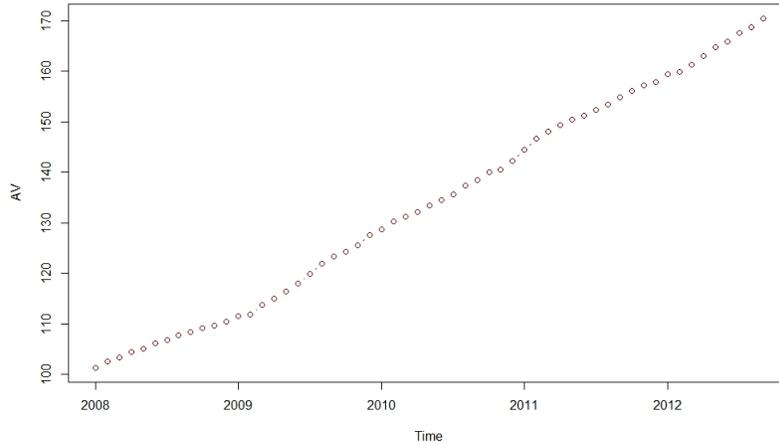
Mes/Año	2008	2009	2010	2011	2012
Enero	101.3	111.5	128.7	144.4	159.4
Febrero	102.6	111.9	130.3	146.6	159.9
Marzo	103.4	113.8	131.2	148.1	161.4
Abril	104.4	115.0	132.2	149.4	163.0
Mayo	105.1	116.5	133.4	150.4	164.8
Junio	106.2	118.0	134.6	151.3	165.9
Julio	106.8	119.9	135.7	152.3	167.6
Agosto	107.7	121.9	137.4	153.4	168.8
Septiembre	108.4	123.3	138.5	154.9	170.5
Octubre	109.2	124.3	140.0	156.2	
Noviembre	109.6	125.6	140.6	157.3	
Diciembre	110.4	127.6	142.2	157.9	

Tabla 4.7. Fuente: Banco Central de Venezuela.

A continuación se presenta la gráfica de la serie asociada al subgrupo "Alquiler de Viviendas" (Figura 4.18).

Se puede observar en la gráfica anterior una inflexión en fechas cercanas a junio del año 2009 y enero del 2011. El primero fue producto de la culminación del lapso para el estudio de la Ley de Arrendamientos Inmobiliarios (LAI) por parte de la Asamblea Nacional, así mismo en el 2011 fue aprobada la Ley para la Regularización y Control de los Arrendamientos Inmobiliarios; ambos instrumentos favorecen de sobremanera a los inquilinos lo cual

Figura 4.18: Serie "Alquiler de Viviendas"



generó un baja en la oferta de inmuebles, resultando en una disminución de los alquileres que es lo que finalmente refleja la gráfica en estos puntos.

Posterior al tratamiento usual de las series, como fue realizado en los ejemplos anteriores, la realización de las gráficas de ACF y PACF, eliminación de tendencia y componente periódico mensual, los modelos considerados y su AIC son los siguientes:

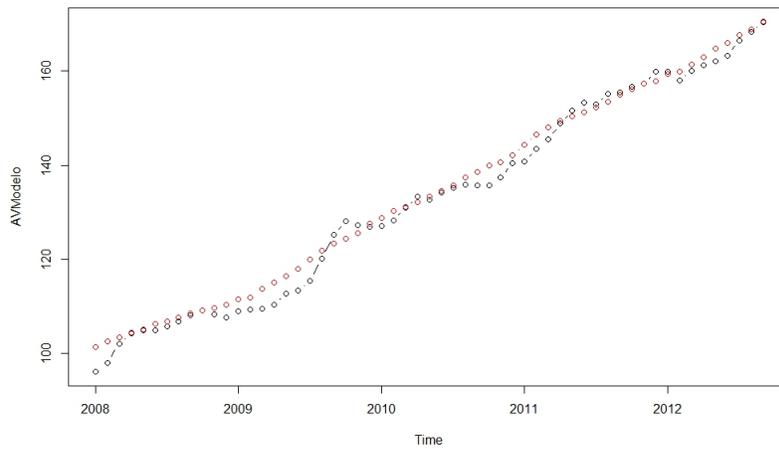
Modelo	AIC
ARMA (6,4)	59.98
ARMA (2,4)	58.9
ARMA (3,2)	57
ARMA (3,3)	59.41
ARMA (10,9)	63.06

Tabla 4.8.

Claramente podemos observar que el modelo cuyo AIC es menor corresponde a los parámetros ARMA (3,2). Al realizar la construcción del modelo pertinente y graficarlo podemos observar el siguiente resultado (Figura 4.19).

Y al comparar las predicciones obtenidas para el último trimestre del año

Figura 4.19: Serie "Alquiler de Viviendas" (rojo), Modelo ARMA (3,2) (negro)



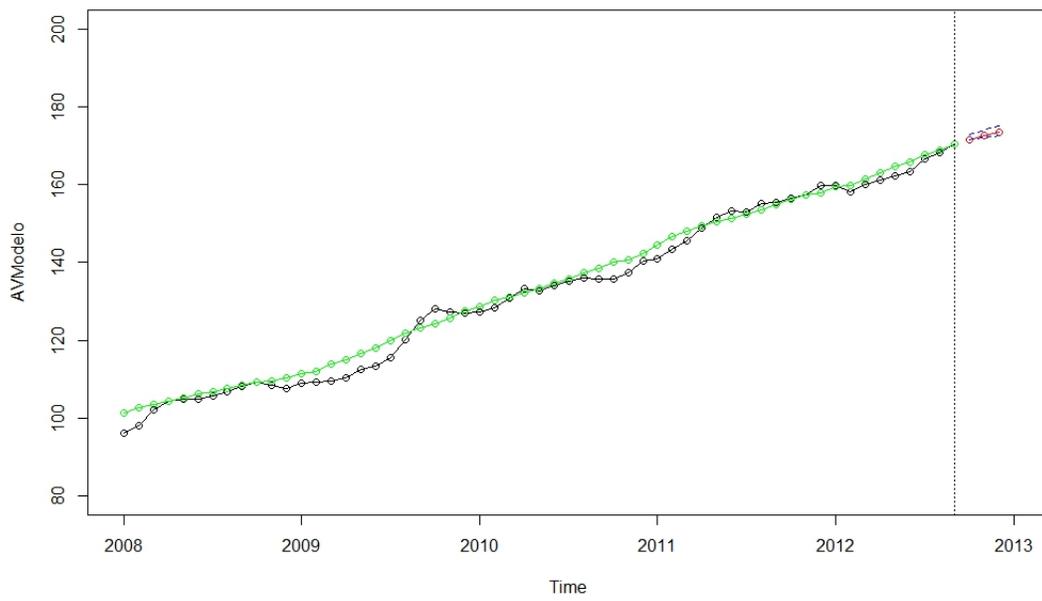
2012 con los datos obtenidos del BCV, tenemos:

Fecha/Fuente	BCV	ARMA (3,2)	Diferencia
Oct. 2012	172.2	171.641	0.5859
Nov. 2012	173.3	172.5661	0.7339
Dic. 2012	174.5	173.5227	0.97

Tabla 4.9

A continuación se presenta la gráfica final asociada a este subgrupo, la cual presenta tanto la serie original como la ajustada a través del modelo ARMA y para finalizar las predicciones. (Figura 4.20)

Figura 4.20: Serie "Alquiler de Viviendas" (verde), Modelo ARMA (negro), Predicciones (rojo)



Conclusión

Se puede observar a lo largo del Capítulo 4 de este trabajo la diferencia clave entre la presentación de las series asociadas a los distintos subgrupos del Índice Nacional de Precios al Consumidor. En la gráfica asociada al subgrupo "Alimentos" los datos presentados son de forma acumulada, en contraste con los presentados bajo la clasificación "Tabaco", la cual es de variación sencilla. Para variaciones acumuladas las gráficas muestran una tendencia creciente naturalmente, sin embargo mucho más notoria que en las de otra clase; así mismo, en ocasiones los saltos bruscos observados en las datas sencillas no son tan distintivos en las series de diferencias porcentuales acumuladas, y esto permite un mejor ajuste de modelos de series temporales.

Por otra parte, se puede señalar el hecho de que aun cuando para la serie "Alimentos" el modelo ajustado fue un SARIMA, mientras que el correspondiente a las otras dos series fue simplemente ARMA, estos modelos sencillos de realizar y cuyo valor computacional no es tan alto, generan predicciones estables y ajustadas a la realidad. Es decir, garantizan ser una buena herramienta para la predicciones de series, al menos en el caso del Índice Nacional de Precios al Consumidor.

Sin embargo aun cuando los modelos creados lograron ajustar las datas, se puede afirmar que mientras más datos originales se poseen, más preciso será el modelo construido. En los 3 ejemplos presentados, se poseían relativamente poca información puesto que en Venezuela la recolección de este indicador de inflación es relativamente reciente.

Para culminar se puede hacer referencia a las predicciones conseguidas por los modelos ajustados, y aun cuando los valores, al menos para el caso Alimentos, fueron cercanos a los valores reales emitidos por el BCV, es alarmante observar que en tan solo 12 años la inflación en este rubro haya alcanzado un 376 por ciento, o para la serie Alquiler de Viviendas se haya alcanzado, en sólo 5 años de diferencia con respecto al año base, 174 por ciento de aumento en los valores. Basándose en estos datos obtenidos, se puede corroborar el hecho de que Venezuela se encuentra entre los primeros países del mundo con mayor inflación.

Bibliografía

- [1] BROCKWELL, P.J.; DAVIS, R.A., Introduction to Time Series and Forecasting. *Springer-Verlag, New York Inc, New York.* (1996)
- [2] BROCKWELL, P.J.; DAVIS, R.A., Time Series: Theory and Methods. *2nd Edition. Springer-Verlag, New York Inc, New York* (2006)
- [3] SHUMWAY, R.H; STOFFER, D.S. Time Series Analysis and Its Applications with R examples. *2nd edition. Springer.* (2006)
- [4] JOHNSON, D.E., Métodos Multivariados Aplicados al Análisis de Datos. *International Thomson Editores.* (2000)
- [5] LÓPEZ C., Cálculo de Probabilidades e Inferencia Estadística. *Publicaciones UCAB.* (2006)
- [6] BANCO CENTRAL DE VENEZUELA, Departamento Económico, Boletín Económico Informativo. *Disponible en: <http://www.bcv.org.ve=Upload=Publicaciones=bcvozecon042010:pdf>*
- [7] BANCO CENTRAL EUROPEO, Material Didáctico, ¿Qué es la Inflación? *Disponible en: <http://www.ecb.europa.eu/ecb/educational/hicp/html/index.es.html>*
- [8] E-CONOMIC, Glosario de términos contables, Inflación. *Disponible en: <http://www.e-conomic.es/programa/glosario/definicion-inflacion>*
- [9] INSTITUTO NACIONAL DE ESTADÍSTICA, Índice Nacional de Precios al Consumidor. *Disponible en: <http://www.ine.gob.ve/index>*
- [10] FINANCIAL RED, IPC: Blog sobre el IPC, ¿Qué es el IPC?. *Disponible en: <http://www.ipcblog.es/que-es-el-ipc/>*